

T.C.
ONDOKUZ MAYIS ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
HARİTA MÜHENDİSLİĞİ ANA BİLİM DALI



UYDU GÖRÜNTÜLERİNİN SINIFLANDIRILMASINDA ÇOK
DEĞİŞKENLİ UYARLANABİLİR REGRESYON EĞRİLERİ
ALGORİTMASINA DAYALI HİBRİT MODELLERİN
GELİŞTİRİLMESİ

Doktora Tezi

Azize UYAR

Danışman

Prof. Dr. Derya ÖZTÜRK

SAMSUN
2024

TEZ KABUL VE ONAYI

Azize UYAR tarafından, Prof. Dr. Derya ÖZTÜRK danışmanlığında hazırlanan “UYDU GÖRÜNTÜLERİNİN SINIFLANDIRILMASINDA ÇOK DEĞİŞKENLİ UYARLANABİLİR REGRESYON EĞRİLERİ ALGORİTMASINA DAYALI HİBRİT MODELLERİN GELİŞTİRİLMESİ” başlıklı bu çalışma, jürimiz tarafından 11.09.2024 tarihinde yapılan sınav sonucunda oy birliği ile başarılı bulunarak Doktora Tezi olarak kabul edilmiştir.

	Unvanı Adı Soyadı Üniversitesi Ana Bilim/Ana Sanat Dalı	Sonuç
Başkan	Prof. Dr. Erdem Emin MARAŞ Samsun Üniversitesi Pilotaj Ana Bilim Dalı	☒ Kabul ☐ Ret
Üye	Prof. Dr. Derya ÖZTÜRK Ondokuz Mayıs Üniversitesi Harita Mühendisliği Ana Bilim Dalı	☒ Kabul ☐ Ret
Üye	Doç. Dr. Fatih SARI Konya Teknik Üniversitesi Harita Mühendisliği Ana Bilim Dalı	☒ Kabul ☐ Ret
Üye	Doç. Dr. Mehmet Ali DERELİ Giresun Üniversitesi Harita Mühendisliği Ana Bilim Dalı	☒ Kabul ☐ Ret
Üye	Dr. Öğr. Üyesi Nükhet KONUK Ondokuz Mayıs Üniversitesi Harita Mühendisliği Ana Bilim Dalı	☒ Kabul ☐ Ret

Bu tez, Enstitü Yönetim Kurulunca belirlenen ve yukarıda adları yazılı jüri üyeleri tarafından uygun görülmüştür.

Prof. Dr. Faik Ahmet SESLİ
Enstitü Müdürü

BİLİMSEL ETİĞE UYGUNLUK BEYANI

Hazırladığım Doktora tezinin bütün aşamalarında bilimsel etiğe ve akademik kurallara riayet ettiğimi, çalışmada doğrudan veya dolaylı olarak kullandığım her alıntıya kaynak gösterdiğimi ve yararlandığım eserlerin Kaynaklar'da gösterilenlerden oluştuğunu, her unsurun enstitü yazım kılavuzuna uygun yazıldığını ve TÜBİTAK Araştırma ve Yayın Etiği Kurulu Yönetmeliği'nin 3. bölüm 9. maddesinde belirtilen durumlara aykırı davranılmadığını taahhüt ve beyan ederim.

Etik Kurul Gerekli mi ?

Evet ☐ (Gerekli ise ekler kısmına ekleyiniz)

Hayır ☒

26 /06 / 2024

Azize UYAR

TEZ ÇALIŞMASI ÖZGÜNLÜK RAPORU BEYANI

Tez Başlığı : UYDU GÖRÜNTÜLERİNİN SINIFLANDIRILMASINDA ÇOK DEĞİŞKENLİ UYARLANABİLİR REGRESYON EĞRİLERİ ALGORİTMASINA DAYALI HİBRİT MODELLERİN GELİŞTİRİLMESİ

Yukarıda başlığı belirtilen tez çalışması için şahsım tarafından 26.06.2024 tarihinde intihal tespit programından alınmış olan özgünlük raporu sonucunda;

Benzerlik oranı : % 11

Tek kaynak oranı : % 3 çıkmıştır.

26 /06 / 2024

Prof. Dr. Derya ÖZTÜRK

ÖZET

UYDU GÖRÜNTÜLERİNİN SINIFLANDIRILMASINDA ÇOK DEĞİŞKENLİ UYARLANABİLİR REGRESYON EĞRİLERİ ALGORİTMASINA DAYALI HİBRİT MODELLERİN GELİŞTİRİLMESİ

Azize UYAR

Ondokuz Mayıs Üniversitesi

Lisansüstü Eğitim Enstitüsü

Harita Mühendisliği Ana Bilim Dalı

Doktora, Eylül/2024

Danışman: Prof. Dr. Derya ÖZTÜRK

Uzaktan algılama verilerinin kullanım alanı arttıkça, farklı özelliklere sahip doğal ve yapay nesnelerin tespit edilmesi, arazi örtüsü/kullanımı türlerinin belirlenmesi ve tematik haritaların üretilmesi, mevcut durum ve değişimlerin belirlenmesi, planlama ve modelleme gibi birçok çalışma için kullanılan görüntü sınıflandırma algoritmaları büyük önem kazanmaktadır. Bu amaç doğrultusunda sınıflandırma doğruluğunun artırılması ve daha güvenilir sonuçların elde edilmesi için yeni yöntem ve yaklaşımlar araştırılmaya devam etmektedir. Bu kapsamda hibrit modelleme yaklaşımlarıyla, farklı yöntemlerin avantajlarının birleştirilerek analizlerin doğruluğunun ve güvenilirliğinin artırılması amaçlanmaktadır.

Bu tez çalışmasında, parametrik olmayan bir sınıflandırma tekniği olan Çok Değişkenli Uyarlanabilir Regresyon Eğrileri (Multivariate Adaptive Regression Splines – MARS) algoritmasının yüksek performanslı bir sınıflandırıcı ve optimizasyon teknikleri ile entegre edildiği yeni hibrit modeller önerilmektedir. Çalışma kapsamında MARS algoritmasına dayalı üç model geliştirilmiştir. İlk modelde, görüntü sınıflandırmada yüksek performansa sahip Rastgele Orman (Random Forest – RF) algoritması ile MARS-RF hibrit modeli oluşturulmuştur. İkinci ve üçüncü modellerde ise MARS algoritmasının parametrelerini optimize etmek amacıyla Parçacık Sürü Optimizasyonu (Particle Swarm Optimization – PSO) ve Bayes Optimizasyonu (Bayesian Optimization – BO) algoritmaları ile entegre MARS-PSO ve MARS-BO hibrit modelleri oluşturulmuştur.

MARS ve RF tekil algoritmaları ve MARS-RF, MARS-PSO ve MARS-BO hibrit modellerin oluşturulması ve uygulanması MATLAB ve RStudio yazılımları ortamında gerçekleştirilmiştir. MARS, RF ve üç farklı hibrit model (MARS-RF, MARS-PSO, MARS-BO), iki farklı uydu görüntüsü kullanılarak (Landsat-8 OLI ve Sentinel-2A MSI) ve iki farklı çalışma alanında (Samsun-Bafra ve Samsun-Atakum) yapılan arazi örtüsü/kullanımı sınıflandırma sonuçlarıyla test edilmiştir. Sınıflandırma sonuçları, toplam doğruluk, kappa istatistiği ve F1 skor doğruluk ölçütleri ile karşılaştırılmıştır. Sonuçlar karşılaştırıldığında tüm hibrit modellerin tüm uydu görüntüsü ve çalışma alanlarında MARS ve RF yöntemlerinden daha yüksek doğruluk gösterdiği belirlenmiştir. Önerilen hibrit modellerin performansı karşılaştırıldığında ise doğruluk sıralaması MARS-BO, MARS-PSO ve MARS-RF şeklindedir. Elde edilen sonuçlar, hibrit modellerin sınıflandırma performansını arttırmada önemli bir rol oynadığını ve özellikle optimizasyon teknikleriyle entegrasyonun sonuçların iyileştirilmesinde etkisinin yüksek olduğunu göstermiştir. Bu bağlamda çalışmanın sonuçları, hibrit modellerin uydu görüntülerinin sınıflandırılmasında önemli avantajlar sağlayarak daha doğru ve güvenilir sonuçların elde edilmesine katkıda bulunduğunu göstermektedir.

Anahtar Sözcükler: Görüntü Sınıflandırma, Hibrit Sınıflandırma, MARS, RF, Optimizasyon, PSO, BO

ABSTRACT

DEVELOPMENT OF HYBRID MODELS BASED ON MULTIVARIATE ADAPTIVE REGRESSION SPLINES ALGORITHM FOR CLASSIFICATION OF SATELLITE IMAGES

Azize UYAR

Ondokuz Mayıs University

Institute of Graduate Studies

Department of Geomatics Engineering

Ph.D., September/2024

Supervisor: Prof. Dr. Derya ÖZTÜRK

With the expanding scope of remote sensing applications, image classification algorithms are becoming increasingly essential across diverse research areas, including the detection of natural and artificial objects, land cover/use classification, thematic map production, monitoring of current conditions and changes, and applications in planning and modeling. To enhance classification accuracy and yield more reliable results, there is a continuous focus on developing novel methods and approaches. In this context, hybrid modeling approaches are particularly significant as they combine the strengths of various methods to improve the accuracy and reliability of analyses.

This thesis proposes innovative hybrid models that integrate the Multivariate Adaptive Regression Splines (MARS) algorithm, a non-parametric classification technique, with high-performance classifiers and optimization strategies. Three MARS-based models were developed as part of this study. The first model combines MARS with the Random Forest (RF) algorithm—recognized for its robust performance in image classification—resulting in the MARS-RF hybrid model. The second and third models enhance the MARS algorithm by optimizing its parameters using Particle Swarm Optimization (PSO) and Bayesian Optimization (BO), leading to the development of the MARS-PSO and MARS-BO hybrid models, respectively.

The standalone MARS and RF algorithms, along with the MARS-RF, MARS-PSO, and MARS-BO hybrid models, were implemented and tested within MATLAB and RStudio environments. These models were evaluated on land cover/use classification tasks using two distinct satellite datasets (Landsat-8 OLI and Sentinel-2A MSI) and two geographic study areas (Samsun-Bafra and Samsun-Atakum). Classification performance was assessed through overall accuracy, kappa statistics, and F1-score metrics. Comparative analysis revealed that all hybrid models exhibited superior accuracy across all satellite images and study areas relative to the standalone MARS and RF methods. Among the hybrid models, the MARS-BO model achieved the highest accuracy, followed by MARS-PSO and MARS-RF. The findings underscore the effectiveness of hybrid models in substantially enhancing classification performance, particularly through the integration of optimization techniques. Overall, the study demonstrates that hybrid models significantly advance satellite image classification, contributing to the generation of more accurate and dependable results.

Keywords: Image Classification, Hybrid Classification, MARS, RF, Optimisation, PSO, BO

ÖN SÖZ VE TEŞEKKÜR

Akademik eğitim ve öğretim sürecimin bir üst noktası olan doktora tez çalışmalarım boyunca beni yönlendirerek destek olan, çalışmam esnasında oldukça özverili davranan, bilgi ve deneyimi ile çalışmamı yönlendiren danışman hocam Sayın Prof. Dr. Derya ÖZTÜRK’e teşekkür ederim.

Tez izleme komitesi hocalarım Sayın Prof. Dr. Aziz ŞİŞMAN, Sayın Prof. Dr. Erdem Emin MARAŞ, ayrıca tez savunma sınavı jüri üyeleri hocalarım Sayın Doç. Dr. Fatih SARI, Sayın Doç. Dr. Mehmet Ali DERELİ ve Sayın Dr. Öğr. Üyesi Nükheth KONUK’a katkılarından dolayı teşekkür ederim.

Öğrenimim boyunca çok değerli bilgiler öğrendiğim ve üzerimde emeği olan tüm hocalarıma, manevi desteğini ve fikirlerini esirgemeyerek daima yanımda olan değerli arkadaşlarıma teşekkür ederim.

Sevgilerini her an hissettiren, her zaman yanımda olan aileme sonsuz teşekkürlerimi sunarım.

Azize UYAR

İÇİNDEKİLER

TEZ KABUL VE ONAYI	i
BİLİMSEL ETİĞE UYGUNLUK BEYANI	ii
TEZ ÇALIŞMASI ÖZGÜNLÜK RAPORU BEYANI	ii
ÖZET	iii
ABSTRACT	iv
ÖNSÖZ VE TEŞEKKÜR	v
İÇİNDEKİLER	vi
SİMGELER VE KISALTMALAR	vii
ŞEKİLLER DİZİNİ	viii
TABLolar DİZİNİ	ix
1. GİRİŞ	1
1.1. Tezin Amacı ve Özgün Değer	4
1.2. Motivasyon	5
1.3. Tezin Yapısı ve Metodoloji	6
1.4. Literatür Özeti	7
2. KURAMSAL TEMELLER	14
2.1. Uydu Görüntülerinin Sınıflandırılması	14
2.2. Hibrit Sınıflandırma	16
2.2.1. Hibrit Sınıflandırma Teknikleri	18
2.2.2. Hibrit Sınıflandırma Performansını Etkileyen Faktörler	20
2.2.3. Hibrit Sınıflandırmanın Avantajları	21
2.3. Çok Değişkenli Uyarlanabilir Regresyon Eğrileri (Multivariate Adaptive Regression Splines – MARS)	22
2.3.1. Düğüm Noktası	25
2.3.2. Temel Fonksiyonlar	26
2.3.3. İleri Adım	28
2.3.4. Geri Adım	29
2.3.5. MARS Algoritmasının Avantaj ve Kısıtları	31
2.3.6. Mekânsal Analizlerde MARS Algoritmasının Kullanımı	32
2.4. Rastgele Orman Algoritması (Random Forest – RF)	33
2.4.1. Topluluk Öğrenmesi	33
2.4.2. RF Algoritması	35
2.4.3. RF Algoritmasının İşlem Adımları	36
2.4.4. RF Algoritmasının Avantaj ve Kısıtları	37
2.5. Optimizasyon	38
2.5.1. Optimizasyon Problemleri	39
2.5.2. Parçacık Sürü Optimizasyonu (Particle Swarm Optimization – PSO)	42
2.5.3. Bayes Optimizasyonu (Bayesian Optimization – BO)	45
2.6. Sınıflandırma Doğruluğunun Değerlendirilmesi	49
2.6.1. Toplam Doğruluk	50
2.6.2. Kappa İstatistiği	51
2.6.3. F1 Skor	52
3. METODOLOJİ	54
3.1. MARS Algoritmasına Dayalı Hibrit Sınıflandırma Modellerinin Geliştirilmesi	54
3.1.1. MARS-RF Hibrit Sınıflandırma Modeli	55
3.1.2. MARS-PSO Hibrit Sınıflandırma Modeli	56
3.1.3. MARS-BO Hibrit Sınıflandırma Modeli	58
3.2. Kodlama	59

3.3. Uygulama ve Modellerin Testi	64
3.3.1. Test Alanlarının Belirlenmesi	65
3.3.2. Uydu Görüntüleri ve Bantların Seçimi	67
3.3.2.1. Sentinel-2A/B.....	67
3.3.2.2. Landsat-8 OLI	68
3.3.3. Sınıflandırmada Kullanılacak Arazi Örtüsü/Kullanımı Sınıflarının Belirlenmesi	68
3.3.4. Görüntü Ön İşleme	69
3.3.5. Alıştırma Örneklerinin Seçilmesi	70
3.3.6. Görüntü Sınıflandırma	71
3.3.7. Sınıflandırma Doğruluk Analizi	72
4. BULGULAR VE TARTIŞMA.....	74
4.1. Test Alanı 1: Bafra İlçesi Sınıflandırma Sonuçları	74
4.1.1. Test Alanı 1: Bafra İlçesi MARS Sınıflandırma Sonuçları	74
4.1.2. Test Alanı 1: Bafra İlçesi RF Sınıflandırma Sonuçları.....	76
4.1.3. Test Alanı 1: Bafra İlçesi MARS-RF Hibrit Sınıflandırma Sonuçları	78
4.1.4. Test Alanı 1: Bafra İlçesi MARS-PSO Hibrit Sınıflandırma Sonuçları	80
4.1.5. Test Alanı 1: Bafra İlçesi MARS-BO Hibrit Sınıflandırma Sonuçları.....	82
4.2. Test Alanı 2: Atakum İlçesi Sınıflandırma Sonuçları	84
4.2.1. Test Alanı 2: Atakum İlçesi MARS Sınıflandırma Sonuçları	84
4.2.2. Test Alanı 2: Atakum İlçesi RF Sınıflandırma Sonuçları.....	86
4.2.3. Test Alanı 2: Atakum İlçesi MARS-RF Sınıflandırma Sonuçları	88
4.2.4. Test Alanı 2: Atakum İlçesi MARS-PSO Sınıflandırma Sonuçları.....	90
4.2.5. Test Alanı 2: Atakum İlçesi MARS-BO Sınıflandırma Sonuçları	92
4.3. Sınıflandırma Sonuçlarının Karşılaştırılması	94
4.4. MARS Algoritmasına Dayalı Hibrit Sınıflandırma Modellerinin Katkıları ve Kısıtları.....	109
5. SONUÇ VE ÖNERİLER	112
KAYNAKLAR	114
ÖZ GEÇMİŞ.....	125

SİMGELER VE KISALTMALAR

BO	: Bayes Optimizasyonu – Bayesian Optimization
BOF	: Uyum Kötülüğü – Badness of Fit
CBS	: Coğrafi Bilgi Sistemleri
CNN	: Evrişimli Sinir Ağları – Convolutional Neural Network
CORINE	: Çevresel Bilginin Koordinasyonu – Coordination of Information on the Environment
DT	: Karar Ağacı – Decision Tree
ELM	: Aşırı Öğrenme Makinesi – Extreme Learning Machine
ESA	: Avrupa Uzay Ajansı – European Space Agency
FAO	: Gıda ve Tarım Örgütü – Food and Agriculture Organization
FMV	: Bulanık Çoğunluk Oylaması – Fuzzy Majority Voting
GCV	: Genelleştirilmiş Çapraz Geçerlilik – Generalized Cross Validation
ISODATA	: Yinelemeli Kendi Kendini Düzenleyen Veri Analizi Tekniği – Iterative Self Organizing Data Analysis Technique
LaSRC	: Arazi Yüzey Yansıma Kodu – Land Surface Reflectance Code
LOF	: Modeldeki Uyuşumsuzluk – Lack Of Fit
MARS	: Çok Değişkenli Uyarlanabilir Regresyon Eğrileri – Multivarite Adaptive Regression Splines
MCS	: Çoklu Sınıflandırma Sistemi – Multiple Classifier System
MLP	: Çok Katmanlı Algılayıcı – Multilayer Perceptron
MR	: Maksimum Kural – Maximum Rule
OLI	: Operational Land Imager
OOB	: Out Of Bag
PSO	: Parçacık Sürü Optimizasyonu – Particle Swarm Optimization
RF	: Rastgele Orman – Random Forest
SAE	: Stacked Auto Encoder
SNAP	: Sentinel Application Platform
SVM	: Destek Vektör Makineleri – Support Vector Machines
TIRS	: Thermal Infrared Sensor
USGS	: Birleşik Devletler Jeolojik Araştırma – United States Geological Survey
WS	: Ağırlıklı Toplam – Weighted Sum

ŞEKİLLER DİZİNİ

Şekil 2.1 Hibrit sınıflandırma (Ranawana and Palade, 2006).....	18
Şekil 2.2 Sınıflandırma yöntemlerinin ardışık kombinasyonu (Du et al., 2012)	19
Şekil 2.3 Sınıflandırma yöntemlerinin paralel kombinasyonu (Du et al., 2012)	19
Şekil 2.4 MARS modelindeki düğüm noktaları (Anonim 2019)	25
Şekil 2.5 MARS modelinde düğüm noktaları arasındaki temel fonksiyonlar (Ünal, 2009). .	26
Şekil 2.6 Temel fonksiyonlar (yansıma çifti) (Özmen, 2010)	27
Şekil 2.7 İkili ağaç diyagramı (Hastie et al., 2001).....	29
Şekil 2.8 MARS modeli işlem adımları	31
Şekil 2.9 Torbalama yöntemi (Lee et al., 2020).....	35
Şekil 2.10 RF sınıflandırması (Özdarıcı Ok vd., 2011)	35
Şekil 2.11 Optimizasyon problemleri kategorizasyonu (Sarker and Newton, 2007).....	40
Şekil 2.12 PSO işlem adımları (Hasanoğlu, 2019)	45
Şekil 2.13 1 boyutlu tasarım probleminde BO algoritmasının kullanımı (Brochu et al., 2010).....	47
Şekil 2.14 BO işlem adımları.....	48
Şekil 3.1 MARS-RF işlem adımları.....	56
Şekil 3.2 MARS-PSO işlem adımları	57
Şekil 3.3 MARS-BO hibrit model işlem adımları.....	59
Şekil 3.4 RStudio arayüzü.....	60
Şekil 3.5 MATLAB yazılımı	60
Şekil 4.1 Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS sınıflandırma sonucu.....	74
Şekil 4.2 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS sınıflandırma sonucu	75
Şekil 4.3 Bafra ilçesi Landsat-8 OLI uydu görüntüsü RF sınıflandırma sonucu	76
Şekil 4.4 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü RF sınıflandırma sonucu	77
Şekil 4.5 Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS-RF sınıflandırma sonucu.....	78
Şekil 4.6 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS-RF sınıflandırma sonucu...	79
Şekil 4.7 Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS-PSO sınıflandırma sonucu	80
Şekil 4.8 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS-PSO sınıflandırma sonucu	81
Şekil 4.9 Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS-BO sınıflandırma sonucu.....	82
Şekil 4.10 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS-BO sınıflandırma sonucu	83
Şekil 4.11 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS sınıflandırma sonucu.....	84
Şekil 4.12 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS sınıflandırma sonucu ...	85
Şekil 4.13 Atakum ilçesi Landsat-8 OLI uydu görüntüsü RF sınıflandırma sonucu	86
Şekil 4.14 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü RF sınıflandırma sonucu	87
Şekil 4.15 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS-RF sınıflandırma sonucu	88

Şekil 4.16 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-RF sınıflandırma sonucu.....	89
Şekil 4.17 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS-PSO sınıflandırma sonucu.....	90
Şekil 4.18 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-PSO sınıflandırma sonucu.....	91
Şekil 4.19 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS-BO sınıflandırma sonucu	92
Şekil 4.20 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-BO sınıflandırma sonucu.....	93



TABLÖLAR DİZİNİ

Tablo 2.1 Sınıflandırma yöntemlerinin kategorilendirilmesi (Lu and Weng, 2007)	14
Tablo 2.2 Hata matrisi.....	51
Tablo 3.1 Sentinel-2A MSI uydu görüntüsünün teknik özellikleri.....	67
Tablo 3.2 Landsat-8 OLI uydu görüntüsünün teknik özellikleri.....	68
Tablo 3.3 Eğitim ve test veri sayıları	71
Tablo 4.1 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS sınıflandırma sonucu hata matrisi.....	75
Tablo 4.2 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS sınıflandırma sonucu hata matrisi.....	76
Tablo 4.3 Bafra ilçesi Landsat-8 OLI uydu görüntüsü RF sınıflandırma sonucu hata matrisi.....	77
Tablo 4.4 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü RF sınıflandırma sonucu hata matrisi.....	78
Tablo 4.5 Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS-RF sınıflandırma sonucu hata matrisi	79
Tablo 4.6 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS-RF sınıflandırma sonucu hata matrisi	80
Tablo 4.7 Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS-PSO sınıflandırma sonucu hata matrisi	81
Tablo 4.8 Bafra ilçesi Sentinel-2A uydu görüntüsü MARS-PSO sınıflandırma sonucu hata matrisi	82
Tablo 4.9 Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS-BO sınıflandırma sonucu hata matrisi	83
Tablo 4.10 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS-BO sınıflandırma sonucu hata matrisi	84
Tablo 4.11 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS sınıflandırma sonucu hata matrisi	85
Tablo 4.12 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS sınıflandırma sonucu hata matrisi	86
Tablo 4.13 Atakum ilçesi Landsat-8 OLI uydu görüntüsü RF sınıflandırma sonucu hata matrisi.....	87
Tablo 4.14 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü RF sınıflandırma sonucu hata matrisi.....	88
Tablo 4.15 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS-RF sınıflandırma sonucu hata matrisi	89
Tablo 4.16 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-RF sınıflandırma sonucu hata matrisi	90
Tablo 4.17 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS-PSO sınıflandırma sonucu hata matrisi	91
Tablo 4.18 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-PSO sınıflandırma sonucu hata matrisi	92

Tablo 4.19 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS-BO sınıflandırma sonucu hata matrisi	93
Tablo 4.20 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-BO sınıflandırma sonucu hata matrisi	94
Tablo 4.21 Bafra İlçesi Kullanıcı Doğruluğu / Kesinlik, Üretici Doğruluğu / Duyarlılık ve F1 Skor sonuçları.....	95
Tablo 4.22 Atakum İlçesi Kullanıcı Doğruluğu / Kesinlik, Üretici Doğruluğu / Duyarlılık ve F1 Skor sonuçları.....	102
Tablo 4.23 Bafra ve Atakum ilçeleri için Landsat-8 OLI ve Sentinel-2A MSI veri setlerinin sınıflandırılmasında Toplam Doğruluk, Kappa ve makro ortalama ile hesaplanan F1 skor sonuçları.....	107



1. GİRİŞ

Uzaktan algılama, geniş kapsama alanı, tekrar algılama yapabilme ve çoklu spektral bantlar gibi avantajlara sahiptir. Bu teknoloji, geniş ya da erişilmesi zor alanlarda güncel bilgi elde edilmesini, küresel ve bölgesel gözlemler yapılmasını, ayrıca arazi örtüsü/kullanımı değişim analizlerinin gerçekleştirilmesini sağlar (Lu and Weng, 2007; Li et al., 2014). Uzaktan algılama ile elde edilen veri zenginliği sayesinde, yeryüzünün mevcut durumunu gözlemlemek ve görüntülerden bilgi çıkararak geleceğe dair tahminlerde bulunmak mümkündür. Bu kapsamda, uzaktan algılama verilerinin analizi önemli bir çalışma alanı oluşturmuş olup, edinilen bilgilerin daha doğru, hızlı ve güvenilir olmasını sağlamak amacıyla araştırmalar devam etmektedir (Herold et al., 2002; Lillesand et al., 2015).

Uzaktan algılamada veri analizi ile ilgili çalışmaların başında uydu görüntülerinin sınıflandırılması gelmektedir. Arazi örtüsü/kullanımını belirlemek, alanları hesaplamak, değişimleri tespit etmek, geleceğe yönelik tahmin ve modellemeler yapabilmek için uydu görüntülerinde sınıflandırma analizi yapılır. Görüntü sınıflandırma, arazi örtüsü/kullanımı bilgisi, şehir planlama, tarım, ormancılık, jeolojik uygulamalar ve askeri alanlar gibi birçok alanda kullanılmaktadır (Mather and Koch, 2011; Lillesand et al., 2015).

Arazi örtüsü/kullanımı bilgileri, sürdürülebilir yönetim, doğal kaynakların geliştirilmesi ve kullanılması, çevrenin korunması, planlama, bilimsel analiz, modelleme ve izleme konularında önemli bir rol oynamaktadır. 1980'ler ve 1990'lar boyunca çoğu görüntü sınıflandırma yöntemi, görüntü pikselini temel analiz birimi olarak kullanmıştır (Blaschke, 2010; Lillesand et al., 2015). Piksel tabanlı sınıflandırma yöntemleri olarak adlandırılan bu yaklaşımlarda, her piksel tek bir arazi örtüsü/kullanımı sınıfı olarak belirlenir. Piksel tabanlı sınıflandırma yöntemleri, kontrolsüz (ISODATA, kümeleme vb.) ve kontrollü (maksimum olabilirlik, paralelyüz vb.) sınıflandırma yöntemleri olarak iki ana kolda geliştirilmiştir. Ancak bir piksel, yalnızca tek bir arazi örtüsü veya kullanım türünü değil, aynı zamanda çeşitli arazi örtüsü veya kullanım türlerinin bir karışımını da içerebilir; bu durum, literatürde karışık piksel problemi olarak geçmektedir. Karışık piksel sorununu çözmek için 1980'li yılların sonuna doğru bulanık sınıflandırma yöntemleri geliştirilmiştir. Ayrıca IKONOS, QuickBird ve WorldView gibi yüksek çözünürlüklü uzaktan algılama sensörlerinin üretilmesiyle nesne tabanlı sınıflandırma yöntemleri

geliştirilmiştir (Dribault et al., 2012; Wilson and Oreopoulos, 2013).

Uydu görüntüleri, elektromanyetik spektrumun görünür, kızılötesi ve mikrodalga kısımlarından elde edilebilir ve bu farklı türdeki uydu görüntüleri yeryüzünün farklı özelliklerini tespit eder. Ayrıca dokusal bilgiler, indeksler, Sayısal Yükseklik Modeli veya diğer yardımcı veri kaynakları gibi orijinal uzaktan algılama verilerinden türetilen bilgiler de sınıflandırma doğruluğunun artırılmasına katkıda bulunabilir (Pal and Mather, 2005). Ancak, girdi verisi miktarındaki artış her zaman sınıflandırma performansında artış anlamına gelmemektedir. Böyle bir yaklaşım, daha gürültülü ve gereksiz veriler içeren büyük bir veri hacmi yaratabilir ve bu da sınıflandırma doğruluğunu azaltabilir. Bu nedenle, sınıflandırma algoritmalarının uygun olarak seçilmesi sınıflandırma sonuçlarının kalitesinde önemli bir iyileşme sağlayabilir (Lu and Weng, 2007).

Yaygın olarak kullanılan klasik sınıflandırma algoritmalarının en büyük sınırlaması, uzaktan algılama verilerini yeterince modelleyemeyen varsayımlara dayanmalarıdır. Son zamanlarda, makine öğrenme teorisine dayanan parametrik olmayan sınıflandırma teknikleri geliştirilmiştir. Klasik sınıflandırıcıların aksine, Rastgele Orman (Random Forest – RF), Yapay Sinir Ağı (Artificial Neural Networks – ANN), Destek Vektör Makineleri (Support Vector Machines – SVM) ve Çok Değişkenli Uyarlanabilir Regresyon Eğrileri (Multivariate Adaptive Regression Splines – MARS) gibi ileri sınıflandırma teknikleri varsayımlara dayanmaz ve bu nedenle karmaşık bir veri kümesini daha etkili bir biçimde ele alabilirler. Bu özellikleri nedeniyle, bu sınıflandırıcılar klasik yöntemlere önemli alternatifler olarak görülmekte ve özellikle regresyon yöntemlerinin sınıflandırıcı olarak kullanılması ilgi çekici hale gelmektedir (Lu and Weng, 2007; Waske et al., 2007; Kuter, 2014; Li et al., 2014).

Görüntü sınıflandırmada kullanılan regresyon analizi değişkenler arasındaki ilişkiyi açıklayabilecek doğru modele karar vererek, bilinen verilerden hareketle bilinmeyen durumlar hakkında tahminler yapma olanağı sağlar. Ancak görüntü sınıflandırmada kullanılan klasik regresyon yöntemleri, gerçek olguları tam anlamıyla modelleyememektedir. Bunun nedeni, gerçek olgularda doğrusallık, değişkenler arası etkileşim, birden fazla bağımlı ve bağımsız değişkenin bulunması gibi problemlerin bulunmasıdır. Bu sebeplerden dolayı, parametrik olmayan regresyon modelleri kullanılmaya başlanmıştır (Breiman et al., 1984; Kuter, 2014;

Bui et al., 2019). Friedman (1991) tarafından Parametrik olmayan regresyon analizlerinde kullanılmak üzere MARS algoritması geliştirilmiştir. Mevcut doğrusal ve doğrusal olmayan regresyon modellerinin aksine, MARS bağımsız değişkenlerin çeşitli aralıkları için farklı katsayılar hesaplayarak ve etkileşim terimlerini dahil ederek, gerçek dünya yapılarını daha doğru yansıtan modeller oluşturur. Bu sayede, MARS algoritması karmaşık veri kümelerini daha etkili modelleyebilir ve piksel tabanlı sınıflandırma yöntemlerinin çözümleyemediği doğrusal olmayan ilişkileri daha iyi yansıtabilir. Ancak, MARS algoritmasının da bazı kısıtları bulunmaktadır. Aşırı uyum sorunu, modelin eğitim verisine fazla uyum sağlamasına ve test verisindeki performansının düşmesine neden olabilir. Ayrıca, MARS algoritmasının karmaşıklığı ve parametrelerinin belirlenmesindeki zorluklar da önemli kısıtlar arasındadır (Friedman, 1993).

Görüntü sınıflandırmasının tarihsel sürecine bakıldığında, daima bir gelişim ve değişim sürecinin yaşandığı görülmektedir. Bu gelişim, sınıflandırma doğruluğunu arttırmaya yönelik çalışmalarla devam etmektedir. Görüntü sınıflandırma analizinde performansı arttırmaya yönelik çalışmalar yeni yöntemlerin geliştirilmesi, klasik yöntemlerde değişiklikler yapılması ve ilave yöntemler eklenmesi çerçevesinde gerçekleştirilmektedir (Li et al., 2014; Mather and Tso, 2016). Özellikle günümüzde yapılan çalışmalar, sınıflandırma performansını arttırmak için mevcut yöntemlerin avantajlarını birleştirmeye odaklanmıştır. Mevcut yöntemlerin birleştirilmesi sınıflandırma algoritmalarının hibrit kullanımıyla mümkündür. Hibrit modeller, farklı algoritmaların güçlü yönlerini bir araya getirerek daha yüksek doğruluk sağlar. Bu yöntemler, farklı sınıflandırma algoritmalarının birlikte kullanılması sayesinde, her bir algoritmanın zayıf yönlerini telafi eder ve daha iyi bir performans elde edilmesini sağlar (Ranawana and Palade, 2006; Foody et al., 2007; Lv et al., 2017; Saini and Ghosh, 2017).

Hibrit sınıflandırma, özellikle karmaşık veri setlerinde ve farklı koşullarda yüksek doğruluk oranları elde etmek için büyük önem taşır. Bu sayede, arazi örtüsü/kullanımı değişikliklerinin izlenmesi, çevresel etkilerin değerlendirilmesi ve doğal kaynak yönetimi gibi alanlarda daha etkin ve güvenilir analizler yapılabilir. Son zamanlarda, sınıflandırma algoritmalarının yanı sıra optimizasyon algoritmalarının kombinasyonunu içeren hibrit sınıflandırma modelleri geliştirilmektedir. Bu hibrit modeller, sınıflandırma sürecini optimize etmek ve

iyileştirmek için yaygın olarak kullanılmaktadır (Omran et al., 2004; Bo and Wang, 2005; Wang et al., 2017; Hoang et al., 2018).

Uydu görüntülerinin sınıflandırılmasında, son zamanlarda kullanılmaya başlanan ve regresyon yöntemi olarak öne çıkan MARS yönteminin yukarıda bahsedilen kısıtlarını gidermek için hibrit modeller kullanılabilir. Hibrit modeller, farklı sınıflandırma algoritmalarının birlikte kullanılması sayesinde, her bir algoritmanın zayıf yönlerini telafi eder ve daha iyi bir performans elde edilmesini sağlar. MARS algoritması ile diğer sınıflandırma algoritmalarının hibrit kullanımı, hem aşırı uyum sorununu azaltabilir hem de modelin genel performansını artırabilir. Aynı zamanda MARS algoritmasının parametrelerinin belirlenmesinde optimizasyon yöntemlerinin kullanılmasıyla hesaplama zorluğu azaltılarak, performans artırılabilir (Omran et al., 2004; Bo and Wang, 2005; Wang et al., 2017).

1.1. Tezin Amacı ve Özgün Değer

Bu tezin amacı, piksel tabanlı parametrik olmayan bir sınıflandırma tekniği olan MARS'ın, arazi örtüsü/kullanımı haritalarının doğruluğunu arttırmak için performansı yüksek bir sınıflandırıcı veya optimizasyon yöntemi ile entegre edilerek yeni hibrit modellerin geliştirilmesidir. Bu amaç doğrultusunda, i) girdi veri setlerinin elde edilmesi ve sınıflandırma öncesi ön işlemlerin yapılması, ii) MARS algoritmasının RF, Parçacık Sürü Optimizasyonu (Particle Swarm Optimization – PSO) ve Bayes Optimizasyonu (Bayesian Optimization – BO) teknikleri ile entegre edildiği hibrit sınıflandırma süreci ve tematik harita üretimi, iii) sonuçların karşılaştırılması olmak üzere üç ana işlem adımı takip edilmiştir.

Önerilen hibrit yöntemler uydu görüntülerinin sınıflandırılmasında daha önce uygulanmamıştır. Bu nedenle tez çalışması kapsamında aşağıdaki araştırma hedeflerine ulaşılması hedeflenmektedir.

- Parametrik olmayan sınıflandırıcıların gücünün ve uygunluğunun gösterilmesi (Bu algoritmalar yüksek sınıflandırma doğruluğuna sahiptir ve genellikle klasik parametrik sınıflandırıcılardan daha iyi performans göstermektedir.)
- Uydu görüntülerinin sınıflandırılması alanında mevcut sınıflandırma yöntemlerinin kısıtlamalarının en aza indirgenmesi ve doğruluğun artırılması için yeni yaklaşımların geliştirilmesi.
- MARS algoritmasının performansını arttırmak için sınıflandırma

yöntemlerinden yine yüksek performansa sahip RF algoritması ile hibrit bir modelin geliştirilmesi.

- MARS sınıflandırma algoritması parametrelerinin bulunması sorunlarının optimizasyon tekniği ile çözülmesi. Bu kapsamda PSO ve BO optimizasyon yöntemleriyle entegre hibrit modellerin geliştirilmesi.
- MARS algoritmasının performansını arttırmak için RF sınıflandırma, PSO ve BO optimizasyon tekniklerini kullanarak oluşturulan hibrit modellerle tahmin doğruluğunun ve modelin genel performansının artırılması.

Bu çalışmada, sınıflandırma yöntemlerinin kısıtlamalarını en aza indirmek ve doğruluğu arttırmak amacıyla uydu görüntülerinin sınıflandırılması alanında MARS algoritmasına dayalı yeni hibrit sınıflandırma yaklaşımları geliştirilerek, daha doğru ve kapsamlı bilgi elde etmek için araştırmacılara ve uygulayıcılara önemli bir aracın sağlanması amaçlanmıştır. MARS algoritmasına dayalı yeni modellerin geliştirilmesi amacıyla, MARS algoritmasının performansını arttırmak için yüksek performansa sahip bir sınıflandırıcı ile birleştirilmesi ve optimizasyon tekniklerinin kullanılarak algoritma parametrelerinin optimize edilmesini kapsayan yaklaşımlar önerilmektedir. Bu modeller, MARS algoritmasının tahmin doğruluğunu, hesaplama hızını ve/veya modelin genel performansını arttırmayı hedeflemektedir. Yaygın kullanılan ileri sınıflandırma yöntemlerinde değişkenlerin öneminin belirlenememesi, yorumlama güçlüğü olması doğruluğu etkileyen dezavantajlardandır. Bu tezde, bu dezavantajları parametrik olmayan regresyon eğrileri çerçevesinde "akılcı" ve "yenilikçi" bir şekilde ele alarak görüntü sınıflandırması için yeni alternatif yaklaşım araştırılacaktır. MARS algoritmasına dayalı hibrit yaklaşımların sınıflandırma doğruluğu farklı alan ve farklı veri setlerinde irdelenecektir.

1.2. Motivasyon

MARS algoritmasının hibrit modellerde kullanımı oldukça nadirdir ve literatürde sınırlı sayıda örnek bulunmaktadır. Ancak, MARS algoritmasına dayalı hibrit model oluşturma çalışmaları, sağlık ve mühendislik gibi farklı disiplinlerde giderek artan bir ilgi görmektedir (Yao et al., 2013; Santoso and Wulandari, 2018; Benemaran and Esmaeili-Falak, 2020; Nguyen et al., 2022; Hanteh et al., 2023; Hai et al., 2023). Bu bağlamda, MARS algoritmasına dayalı hibrit modellerin uydu görüntüleri sınıflandırma analizinde de benzer başarılar elde edebileceği öngörülmektedir. Ancak, bu alandaki araştırmaların sınırlı olması, MARS algoritmasının potansiyelini tam anlamıyla ortaya koyabilmek için daha fazla çalışma

yapılması gerektiğini göstermektedir.

Uzaktan algılama teknolojisi, sağladığı yüksek çözünürlük, geniş kapsama alanı ve tekrarlanabilir görüntüleme gibi teknik avantajlar sayesinde arazi örtüsü/kullanımı sınıflandırması ve tematik haritalamada yaygın olarak kullanılmaktadır. Bu teknoloji, mevcut durumun tespit edilmesinin yanı sıra, iklim değişikliği, kentsel alanların genişlemesi ve sosyoekonomik eşitsizlik gibi faktörlerin yol açtığı arazi örtüsü/kullanımındaki büyük değişimlerin izlenmesinde de önemli bir rol oynamaktadır (Lillesand and Kiefer 1994; Lu and Weng, 2007; Lillesand et al., 2015; Öztürk, 2002).

Arazi örtüsü/kullanımı haritalamak, değişimleri izlemek ve modellemek için yeni sınıflandırma yöntemlerinin geliştirilmesinin yanı sıra mevcut yöntemlerin hibrit olarak birleştirilmesi de literatürde öne çıkmaktadır. Son yıllarda giderek daha fazla tercih edilen hibrit sınıflandırma yöntemleri, uzaktan algılama görüntülerinde sınıflandırma doğruluğunu artırma konusunda oldukça etkili olmuş ve başarılı sonuçlar elde edilmiştir (Lu and Weng, 2007; Du et al., 2012; Zhang and Tang, 2013; Phiri and Morgenroth, 2017; Wang et al., 2018; Tan et al., 2019; Hanteh et al., 2023). Tekil sınıflandırıcıları bir araya getirerek sınıflandırıcı toplulukları oluşturan hibrit modeller, son yirmi yılda hızla gelişmiş ve görüntü işleme, örüntü tanıma ve hedef tanımlama gibi alanlarda geniş bir uygulama yelpazesi bulmuştur. Doğruluğu ve verimliliği artırma potansiyelleri sayesinde hibrit yöntemler, günümüzde oldukça ilgi çeken bir araştırma konusu haline gelmiştir (Phyu, 2009; Chi et al., 2009; Rozenstein and Karnieli, 2011).

Bu doktora tezinin temel motivasyonu, uydu görüntüsü sınıflandırma analizlerinde mevcut yöntemlerin güçlü yanlarını birleştirerek sınıflandırma doğruluğunu ve performansını artırmaktır. Hızla değişen çevresel koşullarda, arazi örtüsü/kullanımı değişikliklerini etkin bir şekilde izlemek ve anlamak için en doğru ve güvenilir haritaların üretilmesi gerekmektedir. Bu hedef doğrultusunda, MARS algoritmasına dayalı üç farklı hibrit model geliştirilmiştir.

1.3. Tezin Yapısı ve Metodoloji

Tez kapsamında, arazi örtüsü/kullanımının sınıflandırılmasında MARS algoritmasına dayalı hibrit bir yaklaşımın ortaya konulması amaçlanmaktadır. Tez çalışması beş bölümden oluşmaktadır:

Birinci bölüm, teze giriş ve genel bakış niteliğinde olup çalışmanın amaç ve özgün değerini, motivasyonunu, mevcut tezin yapısını ve literatür incelemesini

ortaya koymaktadır.

Çalışma için işlevsel bir arka plan sunmak amacıyla, ikinci bölüm uydu görüntülerinin sınıflandırılması, hibrit sınıflandırma, MARS, RF, optimizasyon ve sınıflandırma doğruluğunu değerlendirme konularını ele almaktadır.

Üçüncü bölüm, MARS algoritmasına dayalı hibrit sınıflandırma modellerinin geliştirilmesi, kodlama ve uygulama/modellerin testlerini kapsayan metodolojiyi sunmaktadır.

Dördüncü bölüm, farklı test alanı ve uydu görüntüleri için MARS ve RF tekil sınıflandırma yöntemleri ile MARS-RF, MARS-PSO ve MARS-BO hibrit model sonuçlarını, sonuçların yorumlanması/karşılaştırılmasını ve MARS algoritmasına dayalı hibrit modellerin katkılarını içermektedir.

Beşinci bölümde tezin çıktıları, literatüre katkısı ve gelecek çalışmalar için öneriler bulunmaktadır.

1.4. Literatür Özeti

Tekil sınıflandırıcıların sonuçlarının birleştirilmesi, sınıflandırma doğruluğunu arttırmakta, eğitim süresini kısaltmakta ve sınıflandırma performansını geliştirmektedir. Bununla birlikte, birçok çalışma hibrit modellerin ancak tekil sınıflandırıcılar doğru ve çeşitliyse, yani düşük hata oranları sergiliyorsa ve farklı hatalar yapıyorsa veya hataları bağımsızsa etkili olduğunu göstermiştir.

Yapılan literatür araştırması sonucunda MARS algoritmasının tek başına kullanılarak hem uydu görüntülerinin sınıflandırılmasında hem de diğer alanlarda performansını kanıtladığı görülmüştür. MARS algoritmasının kullanımını diğer bilim dallarında incelediğimizde, sınırlı sayıda da olsa, MARS algoritması ile hibrit modellerin oluşturulduğu ve bu hibrit modellerin diğer yöntemlerle karşılaştırıldığında performanslarının daha iyi sonuç verdiği anlaşılmaktadır.

Bu bölümde literatür özeti MARS algoritmasının uydu görüntülerinin sınıflandırılmasında kullanımı, hibrit sınıflandırma ve MARS algoritmasına dayalı hibrit modelleme olarak üç alt bölümde ele alınmıştır. İlk olarak MARS algoritmasının uydu görüntülerinin sınıflandırılmasında kullanımı ile ilgili çalışmalar aşağıda özetlenmiştir.

Ouiros vd. (2009), multispektral uydu görüntülerinin sınıflandırılmasında yeni bir yöntem olarak MARS algoritmasını önermiştir. Çalışmada karşılaştırma için Maksimum Benzerlik ve Paralelyüz yöntemleri de kullanılmıştır. Doğruluk karşılaştırma sonuçlarına göre, MARS yöntemiyle, 17 sınıfın 14'ünde daha iyi

sonuçlar elde etmiştir. Bu sonuç MARS yönteminin multispektral sınıflandırma için uygunluğunu doğrulamaktadır.

Kuter (2014), görüntü sınıflandırmasında, kar haritalaması amacına yönelik MARS algoritmasının performansını değerlendirmiştir. Sınıflandırma işlemi için R yazılımında çalışan “earth” modülü kullanılmış, su, bulut, kar ve kara olmak üzere dört adet sınıf oluşturulmuştur. Sınıflandırma sonuçları hata matrisleri kullanılarak karşılaştırılmış ve dört sınıftan üçünde MARS algoritmasının daha yüksek doğruluk verdiği sonucuna ulaşılmıştır.

Veri madenciliğinden kanser tanı çalışmalarına kadar birçok veri analizinde son zamanlarda ön plana çıkan hibrit modelleme yaklaşımı, görüntü sınıflandırma alanında da ilgi görmeye başlamıştır. Geniş bir perspektif katması amacıyla çeşitli veri analizi çalışmalarında geliştirilen hibrit modellerle ilgili literatür özeti aşağıda yer almaktadır. Burada sunulan hibrit modeller tez çalışmasının amaç ve kapsamına uygun olarak iki veya daha fazla sınıflandırıcının birleştirilmesi veya optimizasyon yöntemi kullanılması ile ilgili örneklerle sınırlandırılmıştır.

Friedl ve Brodley (1997), karar ağacı (Decision Tree – DT) tekniklerinden yararlandığı çalışmada tek değişkenli, çok değişkenli ve hibrit karar ağacı yöntemleriyle sınıflandırma gerçekleştirmiş ve hibrit karar ağacının diğer iki yöntemden daha iyi performansa sahip olduğunu göstermişlerdir.

Omran vd. (2004), PSO algoritmasına dayanan kontrolsüz görüntü kümeleme yaklaşımı sunmuşlardır. Yeni görüntü kümeleme algoritması, geniş uygulanabilirliğini göstermek için sentetik, MR ve uydu görüntüleri olarak üç farklı görüntüye başarıyla uygulanmıştır. Önerilen algoritma ISODATA sınıflandırma yöntemi ile karşılaştırılmış ve daha iyi sonuçlar verdiği belirlenmiştir.

Bo ve Wang (2004), Landsat TM görüntülerinin sınıflandırılması için hibrit Oylayan Öznitelik Aralıkları (Voting Feature Intervals – VFI) ve Bayes algoritmalarını kullanmıştır. Önerilen hibrit yöntem ile VFI ve Bayes algoritmaları toplam doğruluk sonucuna göre karşılaştırıldığında hibrit yöntemin daha yüksek toplam doğruluğa sahip olduğu sonucuna ulaşılmıştır.

Fauvel vd. (2006), bulanık kümeler teorisini kullanarak bir sinir ağından ve bir bulanık sınıflandırıcıdan elde edilen sınıflandırma sonuçlarının birleştirilmesi ile elde edilen hibrit bir yöntem önermiştir. Önerilen hibrit yöntemin, farklı sınıflandırıcıların ayrı kullanımına kıyasla sınıflandırma doğruluğunu iyileştirdiği sonucuna ulaşılmıştır.

Foody vd. (2007), biyoçeşitlilik haritalaması için yaptıkları çalışmada, görüntü sınıflandırmada DT, diskriminant analizi, SVM, MLP ve radyan tabanlı fonksiyon yöntemlerini kullanmıştır. Ayrıca birden çok sınıflandırma yöntemini entegre etmek için oy çokluğu seçim kuralını (majority voting rule) kullanmışlardır. Entegre sınıflandırmanın, sınıflandırma kesinliğini belirleme ve sınıflandırma sonrası iyileştirme için kullanılabileceği sonucuna ulaşmışlardır.

Phyu (2009), veri madenciliğinde sınıflandırma teknikleri araştırmasında, DT'nin ve Bayes ağının çoğu durumda, biri çok doğru olduğunda diğerinin doğru olmadığı ve tam tersi şekilde önemli farklı özelliğe sahip olduğu sonucuna varmıştır. İki farklı yöntemin güçlü taraflarını ele alarak, algoritmaları topluluk yöntemiyle birleştirerek hibrit bir yöntem önermiştir.

Chi vd. (2009), hiperspektral uzaktan algılama görüntülerinin sınıflandırılmasındaki nicelik problemini elimine edebilmek için bir topluluk sınıflandırma algoritması uygulamıştır. Önerilen yöntemde, sonuçlar artan sınıflandırma doğruluğunun yanı sıra artan robustluk sağlamıştır.

Rozenstein ve Karnieli (2011), Landsat-TM görüntüleri kullanarak ISODATA, Maksimum Olabilirlik ve bu iki yöntemin kombinasyonundan oluşan hibrit sınıflandırma sonuçlarını karşılaştırmıştır. Elde edilen sonuçlar hibrit sınıflandırmanın diğer iki yöntemle göre daha yüksek doğruluğa sahip olduğunu göstermiştir.

Banos vd. (2013), hem aktivite hem de sensör sınıflandırma seviyelerinde verilen tekil kararların potansiyelini hesaba katan bir hibrit sınıflandırma metodolojisi önermiştir. Ağırlıklı toplam yaklaşımı kullanılarak oluşturulan yöntemin, benzer bir performans elde etmek için yüksek boyutlu bir özellik alanı gerektirebilecek klasik modellerle elde edilen sonuçlardan sistematik olarak daha iyi performans gösterdiği sonucuna ulaşılmıştır.

Zhang ve Tang (2013), arazi örtüsü sınıflandırmasını geliştirmek için Boosting sınıflandırıcısını ve SVM yöntemlerini birleştirmiştir. Sonuç, sadece Boosting ve sadece SVM ile karşılaştırıldığında sınıflandırma doğruluğunda önemli bir gelişme olduğunu göstermiştir.

Jaison vd. (2015), DNA mikroarray teknolojisinin gelişmesi ile kanser sınıflandırmasındaki ilerlemeyi teşvik etmek amacıyla DNA mikroarray verilerini analiz etmek için en yakın komşu, SVM ve Naive Bayes sınıflandırıcıdan oluşan hibrit bir sınıflandırma yöntemi önermişlerdir. Önerilen yöntemin performansı SVM,

en yakın komşu ve Naive Bayes sınıflandırıcılarla karşılaştırılmıştır. Sonuçlar hibrit yaklaşımının klasik sınıflandırıcılardan daha yüksek sınıflandırma doğruluğu ürettiğini göstermiştir.

Piao vd. (2015), yüksek boyutlu verilerin sınıflandırılmasında, her bir sınıflandırıcı için alıştırma örneklerinin bölümlere ayrılmasıyla belirlenen farklı bir dizi alıştırma örneklerinden oluşturulduğu bir topluluk yöntemi önermiştir. Önerilen yöntemde her sınıflandırıcının sonuçları çoğunluk oylamasıyla birleştirilir. Yöntemin performansı diğer yöntemler ile karşılaştırılmış ve sonuçlar, önerilen yöntemin diğer yöntemlerden daha iyi performans gösterdiğini ortaya çıkarmıştır.

Chinnaswamy ve Srinivasan (2016), kanser teşhisinde kullanılan mikroarray gen ekspresyon verilerinde, özellik seçimi ve sınıflandırma için korelasyon katsayısı ile PSO birleştiren hibrit bir özellik seçimi yaklaşımı önermiştir. Çalışma kapsamında, özellik seçimi ve sınıflandırma üç çok sınıflı veri seti üzerinde gerçekleştirilmiştir. Özellik seçimi sürecinden sonra seçilen genler aşırı öğrenme makinesi (Extreme Learning Machines – ELM) sınıflandırıcısına tabi tutulmuştur. Deney sonuçları, önerilen hibrit yaklaşımın daha yüksek sınıflandırma doğruluğu sağladığını ve klasik sınıflandırıcılar ile yapılan deneylere kıyasla daha az özellik kullandığını göstermiştir.

Tan vd. (2016), alıştırma örnekleri seçiminde yeni bir strateji ve daha iyi bir performans elde edebilmek amacıyla çoklu sınıflandırma sistemini kullanmışlardır. Çoklu sınıflandırıcılar topluluğu kullanarak yüksek çözünürlüklü uzaktan algılama görüntülerinde değişiklik tespiti için otomatik yöntem önermişlerdir.

Lv vd. (2017) ELM ve yığınlanmış otomatik kodlayıcı (Stacked Auto Encoder – SAE) yöntemleri kullanarak bir uzaktan algılama görüntü sınıflandırma algoritması önermiştir. Sınıflandırma doğruluğunu iyileştirmek için önerilen yöntemde iki sınıflandırıcının güçlü yönleri ele alınmıştır. SAE, yüksek çeşitlilik oluşturmak için; ELM ise algoritmanın öğrenme hızını iyileştirmek için temel sınıflandırıcı olarak seçilmiştir.

Wang vd. (2018), ağırlıklı oylama (weighted voting) kuralıyla entegre edilen ve üstün bir performans elde eden çoklu nesne özellikleri ve çoklu sınıflandırıcılar kullanan nesne tabanlı bir değişiklik tespit yöntemi önermiştir. Yüksek çözünürlüklü uzaktan algılama veri setinin deneysel sonuçları, önerilen yaklaşımın toplam doğruluk açısından klasik yöntemlerden daha iyi performansa sahip olduğunu

göstermiştir.

Zhang (2018) ve Zhang vd. (2018), yaptıkları çalışmalarda MLP ve Evrişimli Sinir Ağları (Convolutional Neural Network – CNN) yöntemleriyle farklı birleştirme yaklaşımı kullanarak bir hibrit sınıflandırıcı önermiştir. İlk çalışmada (Zhang, 2018) kural tabanlı bir karar birleştirme yaklaşımı kullanılarak, yöntemler kısa ve etkili bir şekilde entegre edilmiştir. İkinci çalışmada ise (Zhang et al., 2018), sınıflandırma doğruluğunu arttırmak için bölgesel bir karar füzyonu ile entegre etmek önerilmiştir. Önerilen her iki hibrit sınıflandırıcı da sınıflandırma doğruluğu açısından MLP ve CNN'den daha iyi performans göstermiştir.

Tan vd. (2019), yüksek çözünürlüklü uydu görüntülerinde çoklu sınıflandırıcılar kullanılarak otomatik değişiklik tespiti için nesne tabanlı bir yaklaşım önermiştir. SVM, en yakın komşu ve ekstra ağaçlar (ExT) sınıflandırıcıları temel sınıflandırıcı olarak seçilmiştir. Değişikliklerin belirlendiği görüntüyü sınıflandırmak için çok ölçekli nesne tabanlı kanıt teorisi füzyonu ve belirsizlik analizi kullanılmıştır. Sonuçlar, piksel tabanlı ve nesne tabanlı yöntemlerin ilgili avantajlarını birleştiren önerilen yaklaşımın etkinliğini ve üstünlüğünü doğrulamıştır.

MARS algoritması, görüntü sınıflandırma alanında nispeten yeni bir yöntem olup, bu alanda yapılan çalışmalar sınırlıdır. MARS algoritmasının hibrit modellerde kullanımı ise oldukça nadirdir ve literatürde sınırlı sayıda örnek bulunmaktadır. Ancak, MARS algoritmasına dayalı hibrit model oluşturma çalışmaları, sağlık ve mühendislik gibi farklı disiplinlerde giderek artan bir ilgi görmektedir. Bu çalışmalarda, MARS algoritması diğer sınıflandırma ve optimizasyon teknikleriyle birleştirilerek daha yüksek doğruluk, genelleme kabiliyeti ve model performansı elde edilmeye çalışılmıştır. Sağlık alanında MARS tabanlı hibrit modeller, tıbbi teşhis ve hastalık öngörüsünde kullanılarak klinik karar destek sistemlerinin geliştirilmesine katkı sağlamıştır. Mühendislik alanında ise, bu hibrit modeller, sistemlerin performansını optimize etmek, arıza teşhisleri yapmak ve karmaşık mühendislik problemlerini çözmek için uygulanmıştır. Çeşitli sağlık ve mühendislik alanlarında kullanılan MARS algoritmasına dayalı hibrit model oluşturma çalışmaları aşağıda özetlenmektedir.

Yao vd. (2013), hastalık tahmini ve teşhisi için veri madenciliği tekniklerini araştırmış ve hastalık tahmin modeli oluşturmak için RF ve MARS yöntemlerinin entegrasyonuna dayalı yeni bir yaklaşım önermiştir. İlk olarak değişkenlerin ön taramasını yapmak ve önem sıralarını elde etmek için RF yöntemi kullanılmıştır. RF

yöntemi sonucu değişkenler elimine edilmiş ve yeni veri kümesi elde edilmiştir. Elde edilen yeni veri kümesi ile hastalığa sahip insanların hayatta kalabilirliğini tahmin etmek için kolay yorumlanabilir MARS yöntemi kullanılmıştır. Elde edilen sonuçlar doğrultusunda önerilen yöntemin daha yüksek bir doğruluk ve uygulanması daha kolay bir model sağladığı belirtilmiştir.

Liao vd. (2018) yaptıkları çalışmada köprülerin sel ve depreme karşı güvenliğini etkileyen faktörleri incelemişlerdir. Köprülerdeki oyulma derinliği etkisini ölçmek için Swarm Optimization (SO) ve MARS yöntemlerinin bir entegrasyonuna dayanan SO-MARS öğrenme algoritmasını önermişlerdir. SO-MARS algoritması, MARS algoritmasının optimum parametreleri bulmak için SO algoritmasının kullanıldığı bir modeldir.

Santoso ve Wulandari (2018), yaptıkları çalışmada erken doğum tahmini için SVM ve MARS yöntemlerinin hibrit bir modelini sunmuşlardır. SVM uygulanırken özellik seçimi en önemli problemlerdendir. Bu doğrultuda sınıflandırma doğruluğunu arttırmak için, MARS algoritması girdi özelliklerini seçmek için kullanılmıştır. Oluşturulan MARS-SVM hibrit modeli sonuçları SVM ve MARS yöntemleri ile karşılaştırılmış ve hibrit MARS-SVM modelinin daha yüksek doğruluk oranına sahip olduğu sonucuna ulaşılmıştır.

Alade vd. (2019), glikol nano akışkanların özgül ısı kapasitesini tahmin etmek için SVM yöntemine BO tekniğini uygulayarak, Bayes Destek Vektör Regresyon modelinin uygulanabilirliğini incelemişlerdir. Araştırma sonuçları, önerilen modelin %99,95 korelasyon katsayısı ve 0,0047'lik karesel ortalama hata ile son derece yüksek bir doğruluk sağladığını göstermektedir. Bu model, güneş kolektörlerinin termal verimliliğinin daha doğru hesaplanabilmesi için etkili bir yöntem sunmaktadır.

Benemaran ve Esmaeili-Falak (2020), beton numunelerinin uzun vadeli basınç dayanımının tahmin edilmesinde MARS yöntemini uygulanmış ve MARS yöntemi ile elde edilen denklemleri PSO algoritması ile optimize etmişlerdir. Son model, daha düşük maliyetli olması ve daha uzun çalışma süresi ile betonun nihai ürününün maliyetini düşürmüş ve mekanik özelliklerini iyileştirmiştir.

Nguyen vd. (2022), yaptıkları çalışmada, madencilik patlamalarında zemin titreşimini tahmin etmek için yenilikçi bir hibrit model olan MARS-PSO-MLP modelinin kullanımını araştırmışlardır. Araştırmada, farklı parametreler ile çeşitli MARS modelleri geliştirilmiş ve bu modellerin çıktı tahminleri yeni bir veri seti

olarak birleştirilmiştir. Önerilen model, tek başına kullanılan MARS ve MLP modellerine kıyasla daha yüksek doğruluk ve güvenilirlik sağlamış, karesel ortalama hata ve ortalama mutlak hata değerlerinde önemli iyileşmeler göstermiştir.

Hanteh vd. (2023), kolonların basınç dayanımını tahmin etmek için MARS ve PSO algoritmalarının birleştirilmesiyle MARS-PSO modeli oluşturmuşlardır. Bu model, eğitim ve test aşamalarında sırasıyla 0.9972 ve 0.9961 korelasyon katsayılarına sahip olmasıyla diğer modellere göre daha iyi performans göstermiştir. Bu yöntemlerin kullanımı, model verimliliğini arttırırken sonuç elde etme hızını da yükseltmiştir.

Hai vd. (2023), orman yangın duyarlılık haritalaması için MARS algoritmasında optimizasyon uygulanan hibrit bir yöntem önermişlerdir. Haritalama için 5 yıllık bir periyodun ele alındığı çalışmada Kedi Sürüsü Optimizasyonu algoritması (Cat Swarm Optimization – CSO), duyarlılık haritaları oluşturmak için MARS parametrelerini ayarlamaktadır. MARS-CSO modelinin, model ağacı, azaltılmış hata budama ağacı ve MARS algoritmasından daha iyi performansa sahip olduğu sonucuna ulaşılmıştır.

2. KURAMSAL TEMELLER

2.1. Uydu Görüntülerinin Sınıflandırılması

Sınıflandırma algoritmaları farklı kriterlere göre kategori edilebilir, Tablo 2.1 farklı sınıflandırma kriterleri ve kategorilerine göre sınıflandırma yöntemlerini göstermektedir.

Tablo 2.1 Sınıflandırma yöntemlerinin kategorilendirilmesi (Lu and Weng, 2007)

Sınıflandırma Kriteri	Sınıflandırma Kategorisi	Sınıflandırma Yöntemleri
Piksel bilgisi kullanımına göre	Piksel tabanlı sınıflandırma	Maksimum Olabilirlik, Minimum Mesafe, ANN, RF, MARS
	Nesne tabanlı sınıflandırma	eCognition
Eğitim örneklerinin kullanılmasına göre	Kontrollü sınıflandırma	Maksimum Olabilirlik, Minimum Mesafe, ANN, DT, RF, MARS
	Kontrolsüz sınıflandırma	ISODATA, kümeleme
Veri dağılımı ve parametre kullanımına göre	Parametrik sınıflandırma	Maksimum Olabilirlik, Doğrusal Diskriminant Analizi (Linear Discriminant Analysis)
	Parametrik olmayan sınıflandırma	Destek Vektör Makineleri, ANN, DT, RF, MARS
Çıktı sayısına göre	Kesin (Hard) sınıflandırma	Maksimum Olabilirlik, Minimum Mesafe, SVM, ANN, RF, MARS
	Yumuşak (Soft) sınıflandırma	Bulanık Küme Sınıflandırıcıları, Alt Piksel Sınıflandırıcı, Spektral Karışım Analizi
Algoritma sayısına göre	Tekil algoritma	
	Hibrit sistem	

Uydu teknolojisinin ilk dönemlerinde yapılan çalışmalar, görüntüdeki her pikseli bir tematik sınıfa atayarak görüntüleri piksel düzeyinde sınıflandırma üzerine yoğunlaşmıştır. Görüntü sınıflandırmada kullanılan ilk yöntemler arasında, maksimum olabilirlik, minimum mesafe ve paralelyüz gibi istatistiksel yöntemler bulunmaktadır. Piksel tabanlı sınıflandırma olarak adlandırılan bu yöntemler piksel yansıtım değerlerine göre görüntüyü sınıflandırmaktadır. Her nesne farklı radyometrik özelliklere sahiptir; bu nedenle farklı nesneler farklı spektral imzalara sahiptir. Bu spektral imzalar, piksel yansıtım değerlerinin benzersiz kombinasyonları

olarak tanımlanabilir ve bu da piksel tabanlı sınıflandırma yöntemlerinde her pikselin farklı sınıflara ayrılmasını sağlar (Lu and Weng, 2007; Tuia et al., 2011; Li et al., 2014). Piksel tabanlı sınıflandırma kontrolsüz ve kontrollü sınıflandırma olarak ikiye ayrılmaktadır. Kontrolsüz sınıflandırma, günümüzde çoğunlukla çalışma alanı hakkında herhangi bir ön bilgi olmadığında veya arazi hakkında ön bilgi elde etmek için kullanılan bir yöntemdir. Verilerin temel yapısının araştırıldığı bu yöntemde, sınıfları temsil eden bir dizi ağırlık merkezi üretilir ve pikseller bu merkezlere göre otomatik olarak sınıflara ayrılır. Kontrolsüz sınıflandırmanın avantajları arasında hızlı olması, insan hatalarını barındırmaması ve ayrıntılı ön bilgi gerektirmemesidir; kısıtı ise azami ölçüde ayrılabilir sınıflar oluşturmastır. Kontrollü sınıflandırma yönteminde ise sınıflandırma, tüm görüntüdeki nesneleri temsil eden ve ait oldukları sınıflar belli olan pikseller kullanılarak gerçekleştirilir. Alıştırma örneklemeleri olarak adlandırılan bu piksel grupları, görüntüdeki gerekli bilgilerin küçük bir örneğini içerir ve her sınıfın özellik açıklamasını genişletmek için kullanılır. Başka bir deyişle, kontrollü sınıflandırma tekniğinde, sınıflandırıcıya karar sınırlarını öğretmek için alıştırma örneklemeleri kullanılır. Elde edilen özellikler her piksel için araştırılarak ve benzerliği değerlendirilerek görüntü sınıflandırılır. Kontrollü sınıflandırma tekniğinin avantajı, kullanıcının deneyimiyle kontrolsüz sınıflandırmaya kıyasla daha doğru sonuç vermesi, kısıtı ise eğitim aşamasının zaman alıcı ve maliyetli olmasıdır (Xia et al., 2017; Sharma et al., 2018).

Uydu teknolojisinde çözünürlüklerin artmasıyla, görüntüleri piksel düzeyinde yorumlamak, bir pikselin tematik anlamını ve farklı arazi örtülerini ayırmak için ek bilgilerin gerekebileceği yönünde çalışmalar yapılmıştır (Geneletti and Gorte, 2003; Frohn et al., 2009). Piksel yansıtım değerlerinin yanı sıra görüntüdeki doku, ton ve büyüklük gibi benzer özelliklere sahip piksellerin gruplandırılması ve bu pikselleri temsil eden nesnelerin oluşturulması ile yapılan sınıflandırma yöntemi, nesne tabanlı sınıflandırma olarak adlandırılır. Bu yöntem sayesinde nesne düzeyinde analizler yapılmaya başlanmış ve nesne tabanlı sınıflandırma, yüksek çözünürlüklü uydu görüntülerinde iyi performans sağlamıştır (Lu and Weng, 2007; Kalkan ve Maktav, 2010; Xia et al., 2017).

Veri dağılımı ve parametre kullanımı göz önüne alındığında, sınıflandırma algoritmaları parametrik ve parametrik olmayan olmak üzere iki grupta incelenmektedir. Parametrik algoritmalar verilerin belirli bir istatistiksel dağılımda olduğunu varsaymaktadır. Parametrik olmayan algoritmalar ise istatistiksel dağılım

hakkında varsayım yapmamaktadır (Huang et al., 2002; Mather and Koch, 2011). Uydu görüntülerinin sınıflandırılmasında DT, ANN ve SVM gibi güçlü parametrik olmayan sınıflandırma yöntemleri, son zamanlarda literatürde geniş yer bulmaktadır (Lu and Weng, 2007; Mather and Koch, 2011).

Görüntü sınıflandırma yöntemleri, çıktı sayısına göre kesin (hard) ve yumuşak (soft) sınıflandırma yöntemleri olarak kategorize edilmektedir. Kesin sınıflandırma yöntemlerinde, her piksel yalnızca tek bir tematik sınıfa atanır, yani pikseller kesin sınıf etiketleri alır. Bu yöntemler, genellikle belirli bir sınıfa ait olma olasılığı en yüksek olan sınıfı seçerek pikselleri sınıflandırır. Buna karşılık, yumuşak sınıflandırma yöntemlerinde, her piksel birden fazla sınıfa atanabilir ve her sınıfa ait olma olasılığı belirtilir. Yumuşak sınıflandırma, piksellerin birden çok sınıfa ait olabileceği karmaşık ve belirsiz alanlarda daha esnek ve ayrıntılı sonuçlar sunar (Lu and Weng, 2007; Foody, 2009).

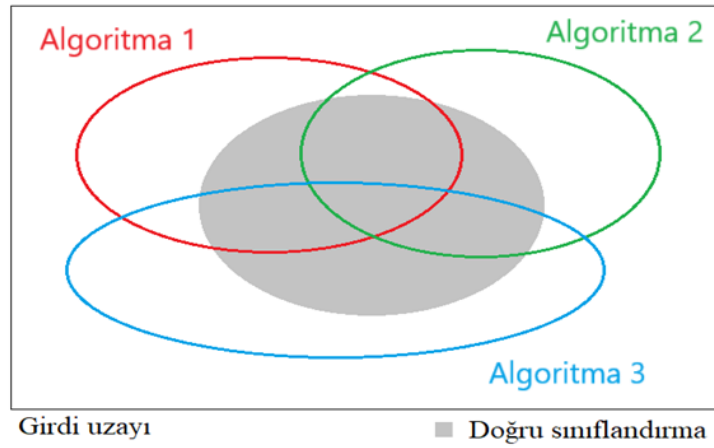
Görüntü sınıflandırma yöntemleri, kullanılan algoritma sayısına göre ele alındığında tekil ve hibrit sınıflandırma yöntemleri olarak kategorize edilebilir. Tekil sınıflandırma yöntemleri, her pikseli sınıflandırmak için yalnızca bir algoritma kullanırken, hibrit sınıflandırma yöntemleri birden fazla algoritmanın kombinasyonunu kullanır. Hibrit sınıflandırma yöntemleri farklı algoritmaların güçlü yanlarını birleştirerek daha yüksek doğruluk ve güvenilirlik elde etmektedir. Bu yöntemler, veri çeşitliliğine ve karmaşıklığına daha iyi uyum sağlar, parametrik ve parametrik olmayan algoritmaların avantajlarını bir araya getirir ve sınıflandırma hatalarını azaltarak daha kapsamlı ve doğru sonuçlar sunar (Hansen and Salamon, 1990; Du et al., 2012; Piao et al., 2015).

2.2. Hibrit Sınıflandırma

Görüntü sınıflandırma yöntemleri, birçok disiplin alanında kullanılmaktadır. Günümüze kadar birçok görüntü sınıflandırma yöntemi önerilmiş, geliştirilmiş ve bu değişim süreci devam etmektedir. Herhangi bir yöntemin diğerleri arasında en iyi yöntem olduğu söylenemez, çünkü doğruluk bir dizi faktöre bağlıdır. Uydu görüntüsünün çözünürlüğü, elde edilecek sınıf sayısı, kullanılan bant sayısı ve çalışma alanının özellikleri sınıflandırma doğruluğunu etkilemektedir (Sharma et al., 2016; Thakur and Maheshwari, 2017). Bu faktörlerin bir arada değerlendirilmesiyle oluşan karmaşık yapı nedeniyle en ideal sınıflandırma yöntemini seçmek zordur. Çünkü her bir sınıflandırma yönteminin kendine özgü güçlü yanları ve sınırlamaları vardır. Örneğin, kontrolsüz sınıflandırma yöntemleri, hızlı ve ön bilgi gerektirmeyen

bir çözüm sunarken, bu yöntemler belirli bir sınıfa ait pikselleri yeterince ayırt edemeyebilir ve bu da doğruluğun düşmesine neden olabilir. Benzer şekilde, maksimum olabilirlik gibi kontrollü sınıflandırma yöntemleri, yeterli eğitim noktaları ve normal dağılımlı görüntü değerleri ile iyi performans gösterirken karmaşık görüntülerde güvenilir sonuçlar elde edilemeyebilir (Lu and Weng, 2007; Phiri and Morgenroth, 2017).

Farklı algoritmalar, aynı çalışma alanı için farklı sonuçlar verebilir fakat tek bir algoritma tüm sınıflar için en iyi performansı gösteremeyebilir (Şekil 2.1). Sınıflandırıcılar, birbiriyle ilişkisiz sınıflandırma hatalarına sahiptirler. Bu nedenle, farklı sınıflandırıcıların birleştirilmesiyle daha yüksek sınıflandırma doğrulukları elde edilebilir. Başka bir deyişle, her bir sınıflandırma yönteminin kendine özgü güçlü yanları ve kısıtları vardır, farklı sınıflandırıcıların güçlü yanları kullanılarak sınıflandırma doğruluğu artırılabilir. Bu yöntem hibrit sınıflandırma yöntemi olarak adlandırılır. Bir hibrit sınıflandırma modeli, tek bir sınıflandırıcı kullanılarak elde edilebilecekten daha iyi bir sınıflandırma performansı elde etmek için bir dizi algoritmadan oluşan bir bileşik modeldir (Hansen and Salamon, 1990; Opitz and Shavlik, 1996; Liu and Mason, 2009). Hibrit sınıflandırma, sınıflandırma alanında nispeten yeni bir kullanım alanına sahiptir ve ilk olarak 1990'larda farklı sınıflandırıcıları birleştirerek daha yüksek doğruluk elde etme amacıyla kullanılmaya başlanmıştır. İki doğrusal regresyon modelinden oluşan bir hibrit model fikri ilk olarak Tukey (1977) tarafından önerilmiştir. Birden fazla sınıflandırıcının birlikte kullanımının yanı sıra tek bir sınıflandırıcı ile sınıflandırma algoritmasının performansını arttırmak amacıyla ek bir optimizasyon adımının dahil edilmesi son zamanlarda uygulanan hibrit yöntemlerdendir. Bu yöntemler, farklı algoritmaların güçlü yanlarını bir araya getirerek, sınıflandırma doğruluğunu arttırmakta ve daha güvenilir sonuçlar elde edilmesine olanak tanımaktadır (Hashem, 1994; Krogh and Vedelsby, 1995; Aziz et al., 2017).



Şekil 2.1 Hibrit sınıflandırma (Ranawana and Palade, 2006)

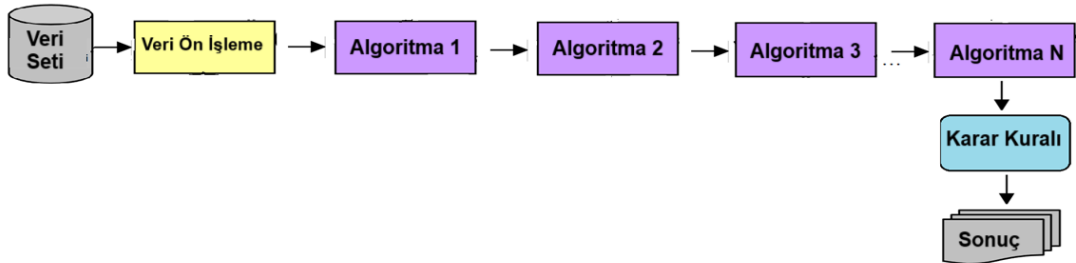
Hibrit yöntem, makine öğrenmesi alanında önemli bir kavramdır ve karar vermek için birden fazla modeli birleştirme yaklaşımını ifade eder. Bu çerçevede her model, bir temel öğrenici olarak kabul edilebilir. Hibrit yöntemin performansının genellikle tekil bir yöntemden daha iyi olduğu yaygın olarak kabul edilmektedir. Bunun nedeni, farklı modellerin aynı veri seti üzerinde her zaman aynı yanlış tahminleri yapmamasıdır; başka bir deyişle, modeller birbirlerinin hatalarını telafi eder. Bu durum, bir dizi yöntemin hibrit olarak çalışmasının, tek bir yöntem tarafından üretilen hataları düzeltebileceği ve böylece daha iyi genel performansa katkıda bulunabileceği şeklindeki temel önermeyi tanımlar. Hibrit yöntemler, sağlamlık ve doğruluğun iyileştirilmesinin yanı sıra, sınırlı örneklem boyutu, yüksek boyutlu görevler, sınıf dengesizliği ve veri gürültüsü gibi zorluklarla başa çıkamadaki etkinliği nedeniyle de araştırma konusu olmaktadır (Aziz et al., 2017; Fallah et al., 2018).

2.2.1. Hibrit Sınıflandırma Teknikleri

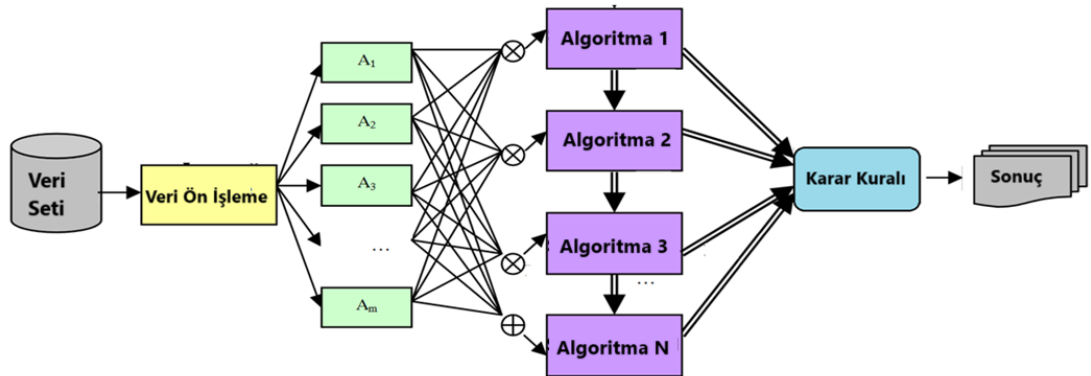
Sınıflandırma performansını yükseltmek amacıyla yapılan araştırmalar, mevcut yöntemlerin özelliklerini birleştirerek sınıflandırma doğruluğunu arttırmaya odaklanmıştır. Farklı yaklaşımlar kullanılarak hibrit sınıflandırma yöntemleri (çoklu sınıflandırma sistemleri) oluşturulabilir. Hibrit yöntemin ilk kullanıldığı zamanlarda görüntü sınıflandırması yöntemlerinin hibrit kullanımı, Çoklu Sınıflandırma Sistemi (Multiple Classifier System – MCS) olarak adlandırılmaktaydı. Çoklu Sınıflandırma Sisteminin ana amacı, birden çok sınıflandırıcı arasındaki tamamlayıcılığı ve doğruluğu arttırmaktır. Daha sonra yapılan çalışmalarla, sadece farklı sınıflandırıcıların değil, farklı algoritmaların birlikte kullanımı ile Hibrit Sınıflandırma Sistemi oluşturulması amaçlanmıştır. Hibrit Sınıflandırma Sistemi,

doğruluk ve verimliliği artırma yeteneği nedeniyle son yirmi yılda hızla geliştirilmiş ve örüntü tanıma, görüntü işleme ve hedef tanımlama gibi çeşitli alanlarda yaygın olarak kullanılmaktadır. (Chi et al., 2009; Liu and Mason, 2009; Du et al., 2012; Tan et al., 2016).

Hibrit Sınıflandırma Sistemi üç ana başlıkta incelenmektedir; alıştırma örneklemeleri aşamasında oluşturulan kombinasyon, ardışık kombinasyon, paralel kombinasyon. (Rozenstein and Karnieli 2011). Bir algoritma tarafından oluşturulan sınıflandırma sonucunun, bir sonraki algoritmada girdi olarak kullanılması ardışık kombinasyon olarak adlandırılmıştır. Şekil 2.2’de gösterilen bu yöntemle oluşturulan hibrit modellerde, her algoritmanın elde ettiği sonuç, zincirdeki son algoritmaya ulaşana kadar sırayla bir sonraki algoritmaya aktarılır. Bu zincirde amaç tüm sistemin doğruluğunu iyileştirmek ve hata kontrollerini dahil etmektir. Şekil 2.3’te gösterilen paralel kombinasyonda ise, çoklu sınıflandırıcılar karşılıklı etkileşim olmadan bağımsız olarak tasarlanır ve çıktıları karar kuralına göre birleştirilir. En yaygın kullanılan karar kuralı her bir algoritmanın belirli bir sınıf için tahmin oluşturmasıdır (Xu et al., 1992; Rahman and Fairhurst, 1999; Rozenstein and Karnieli 2011; Du et al., 2012).



Şekil 2.2 Sınıflandırma yöntemlerinin ardışık kombinasyonu (Du et al., 2012)



Şekil 2.3 Sınıflandırma yöntemlerinin paralel kombinasyonu (Du et al., 2012)

Yaygın olarak uygulanan hibrit sınıflandırma sisteminde karar kuralları;

- a) Maksimum Kural (Maximum Rule – MR), birden çok sınıflandırıcı tarafından sağlanan olasılıkları birleştirmek için basit bir yöntemdir. Her sınıf üyeliği, k sınıftan biri için bir oy olarak yorumlanır ve her bir sınıflandırıcı için en yüksek sınıf üyeliğini alan sınıf, o sınıflandırıcının sınıf etiketi olarak atanır. Daha sonra N sınıflandırıcılardan gelen sınıf etiketleri tekrar karşılaştırılır ve bir pikselin, f_i sınıflandırıcı tarafından verilen bir C_k sınıfına ait olan bir pikselin sınıf üyeliği P_{MR} olasılığı hesaplanır (Eşitlik 2.1). En yüksek sınıf üyeliğini alan sınıf son sınıflandırma olarak belirlenir. Maksimum kuralın en büyük sorunu, güvenilirliklerine bakılmaksızın tüm sınıflandırıcıların aynı yetkiye sahip olmasıdır (Foody, 2009; Salah, 2017).

$$P_{MR} = \max[\max pp(C_k|f_i)] \quad (2.1)$$

- b) Ağırlıklı Toplam (Weighted Sum – WS) yönteminde, ilk olarak her sınıflandırıcının sonucundaki sınıf üyeliği, sınıflandırıcıların her sınıf için ($0 \leq \alpha_{ci} \leq 1$) güvenilirliğine (doğruluğuna) göre ağırlıklandırılır. Bundan sonra her sınıf için tüm sınıflandırıcıların sonucundaki sınıf üyelikleri toplanır ve maksimum toplamı alan sınıf, Eşitlik 2.2'deki gibi son sınıf etiketi olarak belirlenir. Bu denklemde, P_{WS} ağırlıklı toplama dayalı olasılıktır, α_{ci} her sınıflandırıcının ağırlığıdır, pp_i i 'nci için elde edilen sınıf üyelik değeridir, N ise sınıflandırıcıların sayısıdır (Le et al., 2007; Salah, 2017).

$$P_{WS} = \sum_{i=1}^N \alpha_{ci} pp_i \quad (2.2)$$

- c) Bulanık Çoğunluk Oylaması (Fuzzy Majority Voting – FMV) ağırlıkların değerlerinin doğrudan sağlanabilmesi için ağırlıklara bazı anlamlar veren yöntemdir. Ağırlıklar için bulanık dilsel niceleyicilere dayalı anlambilim kullanılır ve Eşitlik 2.3 ile hesaplanabilir. Bu denklemde pp_i i 'nci pikselin Markov olasılığı, k sınıfların sayısıdır (Zadeh, 1983).

$$P_{FMV} = \arg \max_k \left[\sum_{i=1}^N w_{pi} pp_i \right] \quad (2.3)$$

2.2.2. Hibrit Sınıflandırma Performansını Etkileyen Faktörler

Hibrit modeller, farklı modellerin veya algoritmaların güçlü yönlerinden yararlanma avantajı sunarak gelişmiş doğruluk ve genelleme sağlar. Farklı veri türlerini etkili bir şekilde ele alabilir ve sağlam tahminler veya sınıflandırmalar

sağlayabilirler.

Hibrit modellerin kısıtları karmaşıklık, artan hesaplama gereksinimleri ve aşırı uyum riskidir (Aziz et al., 2017; Fallah et al., 2018);

- Karmaşıklık: Hibrit modellerin geliştirilmesi ve uygulanması tek modellere kıyasla daha karmaşık olabilir. Birden fazla algoritma veya modelin entegre edilmesi dikkatli bir tasarım ve koordinasyon gerektirir, bu durum genel sistemin karmaşıklığını arttırabilir.
- Artan hesaplama gereksinimleri: Hibrit modeller genellikle tekil modellere kıyasla daha fazla hesaplama kaynağı gerektirir. Birden fazla model birleştirildiğinden, eğitim ve çıkarımın hesaplama yükü daha yüksek olabilir ve modelin ölçeklenebilirliğini sınırlayabilir.
- Aşırı uyum riski: Hibrit modeller, özellikle farklı modellerin entegrasyonu dikkatli yapılmazsa, aşırı uyum sağlamaya daha yatkındır. Aşırı uyum, bir modelin eğitim verilerine çok yakından uyması durumunda ortaya çıkar ve bu da tüm veriye uygulandığında düşük performansa neden olur. Bu riski azaltmak için uygun düzenleme teknikleri ve doğrulama prosedürleri gereklidir.

2.2.3. Hibrit Sınıflandırmanın Avantajları

Her temel sınıflandırma algoritması, yeni bir örneğin sınıf değerini tahmin etmek için bir eğitim setinden bir model oluşturmak üzere kendi öğrenme prosedürüne sahiptir ve her yeni örneği sınıflandırmak için kullanılacak olan verilerden bir model oluşturur. Hibrit sınıflandırma ise yeni bir örneğin sınıf değerini belirlemek için bir dizi model oluşturur. Yapılan çalışmalar hibrit sınıflandırma algoritmasının yalnızca daha yüksek tahmin doğruluğu elde etmediğini, aynı zamanda öğrenme sonuçlarının kolay yorumlanmasını da sağladığını göstermektedir (Hansen and Salamon, 1990; Hashem, 1994; Krogh and Vedelsby, 1995).

Her bir uygulamaya özel sınıflandırıcı yöntemi veya optimum parametre setini seçmek için evrensel bir öneri yoktur. Hibrit modeller, güçlü yönlerini birleştirerek tekil modellerin sınırlamalarını ele almak için tasarlanmıştır. Hibrit modelin her bir bileşeni kendine özgü yetenekleriyle katkıda bulunarak daha güçlü ve çok yönlü bir model ortaya çıkarır (Aziz et al., 2017; Fallah et al., 2018);

- Daha iyi performans: Hibrit modeller, tekil modellere kıyasla daha yüksek doğruluk ve daha iyi performans elde edebilir. Hibrit modeller, farklı

algoritmaların güçlü yönlerinden yararlanarak tek bir modelin sınırlamalarının üstesinden gelebilir ve daha doğru tahminler veya sınıflandırmalar sağlayabilir.

- Geliştirilmiş sağlamlık: Hibrit modeller genellikle verilerdeki gürültü ve aykırı değerlere karşı daha dayanıklıdır. Farklı algoritmalar farklı veri türlerine karşı farklı hassasiyetlere sahip olduğundan, hibrit bir model daha geniş bir veri varyasyon ile başa çıkabilir ve daha güvenilir sonuçlar üretebilir.
- Artan genelleştirilebilirlik: Hibrit modeller, verilere iyi genelleme yapabilme yeteneğine sahiptir. Her biri kendi genelleme yeteneklerine sahip birden fazla modeli birleştirerek, verilerdeki daha geniş bir desen ve ilişki yelpazesini yakalayabilen hibrit yöntemler daha iyi genelleme performansı sağlar.

2.3. Çok Değişkenli Uyarlanabilir Regresyon Eğrileri (Multivariate Adaptive Regression Splines – MARS)

Regresyon analizi, bir veya daha fazla bağımsız değişken kullanarak bir bağımlı değişken hakkında tahminlerde bulunmak için kullanılan bir analiz yöntemidir. Başka bir ifadeyle, regresyon, bağımlı ve bağımsız değişkenler arasındaki ilişkileri belirleyebilmek için yapılan istatistiksel analizdir ve hem tanımlayıcı hem de çıkarımsal istatistik sağlamaktadır. Regresyon analizinin amacı bağımlı değişken ile bağımsız değişken arasında anlamlı bir ilişki olup olmadığını araştırmak bir ilişki varsa bunun gücünü belirlemek ve ileriye yönelik tahminleri belirlemeye çalışmaktır. Değişkenler arasında tespit edilen ilişki, matematiksel bir fonksiyon şeklinde yazılır ve bu fonksiyona regresyon denklemi denmektedir (Orhunbilge, 1996; Sevimli, 2009).

Regresyon analizinde bir değişkenin başka bir veya daha fazla değişkene olan bağımlılığı fonksiyonel veya deterministik ilişkilerle değil istatistiksel ilişkilerle incelenmektedir. Değişkenlerin bulunduğu tüm sistemlerde, bazı değişkenlerin diğerleri üzerindeki etkilerinin incelenmesi regresyon analizi ile yapılabilmektedir. Regresyon analizi veri tanımlama, parametre kestirimi, ön kestirim ve denetleme amaçlarıyla kullanılabilir. Regresyon mühendislik, fizik, iktisat ve biyoloji gibi farklı bilim dallarının konusu olmuştur (Kutner et al., 2005; Weisberg, 2005; Öztürkcan, 2009; Erar, 2013).

Bazı uygulamalarda bağımlı değişkeni etkileyen bir bağımsız değişken varken

bazı uygulamalarda birden fazla değişken olabilir, bu durum regresyon modeli kurulmasında farklı yaklaşımların kullanılmasına neden olmaktadır. Bir değişkeni etkileyen iki ve daha fazla bağımsız değişken arasındaki neden-sonuç ilişkilerini doğrusal bir modelle açıklamak ve bu bağımsız değişkenlerin etki düzeylerini belirlemek için yararlanılan yöntemle çok değişkenli regresyon analizi denir. Çok değişkenli regresyonda, bağımlı değişkeni etkileyen birden çok bağımsız değişken bulunmaktadır ve Eşitlik 2.4'te gösterildiği gibi ifade edilir (Alpar, 1997).

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_i x_i + \varepsilon \quad (2.4)$$

β_i regresyon katsayısını ifade etmektedir ve çoklu doğrusal regresyon modellerinde, yaygın olarak regresyon katsayısı değerleri en küçük kareler yöntemiyle bulunur. Çoklu regresyon modellerinde veriler matris şeklinde gösterilir, bu nedenle regresyon katsayıları tahmininde kullanılan en küçük kareler yöntemi aşağıdaki gibi matris formunda yazılabilir (Eşitlik 2.5) (Alpar, 1997).

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \dots \\ \beta_i \end{bmatrix} = \left(\begin{bmatrix} 1 & x_{11} & \dots & x_{1i} \\ 1 & x_{21} & \dots & x_{2i} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n1} & \dots & x_{ni} \end{bmatrix}^I \begin{bmatrix} 1 & x_{11} & \dots & x_{1i} \\ 1 & x_{21} & \dots & x_{2i} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n1} & \dots & x_{ni} \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & x_{11} & \dots & x_{1i} \\ 1 & x_{21} & \dots & x_{2i} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n1} & \dots & x_{ni} \end{bmatrix}^I \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} \quad (2.5)$$

Çok değişkenli regresyon modeli varsayımları;

- Model doğrusal kurulmuştur.
- Bağımsız değişken değerleri yinelenen örneklemelerde değişmez.
- Hataların toplamı sıfırdır.
- Hatalar arasında ilişki yoktur.
- Parametreler normal dağılımdadır.

Çok değişkenli regresyon analizinde bağımlı değişkeni etkileyen bağımsız değişkenler modele eklenmektedir, bu ekleme işlemi üç şekilde gerçekleştirilebilir (Anonim 2019);

- Hiyerarşik ekleme (Blockwiseentry): Literatür araştırmalarına dayanarak en önemli tahmin değişkeni önce girilir.
- Zorla ekleme (Enter): Tüm tahmin değişkenleri eşzamanlı olarak modele girilir.
- Adım adım ekleme: Tamamen matematiksel ölçütlere göre girilir. Bağımlı değişkendeki değişimi en fazla açıklayan tahmin değişkenini bulunur, sonra geri kalan değişimi en fazla açıklayan tahmin değişkenini bulunur.

Doğadaki olgular, olaylar doğrusal olmayan bir eğilim gösterdiğinden doğrusal modellerle temsil edilemezler. Bu modellerde genellikle parametrik olmayan yöntemler kullanılır, fakat oluşturulacak modelde kullanılacak bağımsız değişken sayısı fazla sayıda ise çoğu parametrik olmayan regresyon yöntemleri yeterli olmamaktadır. Çünkü büyük veri setlerinden bilgi çıkarımı yapmak yani modellemek zordur. Son yıllarda, gerçek olguların daha etkili yansıtılabilmesi amacıyla birçok bağımsız değişkenin kullanılması yaygın hale gelmiştir. Bu bağlamda, 1991 yılında Jerome Friedman tarafından geliştirilen MARS, çok sayıda bağımsız değişkenin bir arada değerlendirilmesini mümkün kılan algoritmaların kullanımını sağlamıştır. MARS, regresyon analizlerindeki karmaşıklığı azaltarak, veri setlerindeki ilişkileri daha hassas modellemeye imkân tanımıştır.

MARS yöntemi, klasik doğrusal ve doğrusal olmayan yöntemlerden farklı bir yaklaşım sergiler. Çünkü MARS bağımlı ve bağımsız değişkenler için herhangi bir özel varsayıma ihtiyaç duymaz. Bu yöntem, belirlenen aralıklarda farklı katsayılar üretir ve modelin yapısını daha iyi yansıtmak adına etkileşim terimlerini dahil eder. MARS algoritması, bağımlı bir değişken ile açıklayıcı değişken seti arasındaki olası ilişkiyi belirlemek amacıyla "düzleştirme eğrileri (uzanımları – smoothing splines)" kullanma prensibine dayanır. Hem yüksek boyutlu sınıflandırma hem de tahmin problemleri için kullanılacak esnek bir modelleme tekniği olan MARS bağımsız değişkenlerin oluşturduğu uzayı birbirleriyle örtüşen birçok bölgeye ayırır ve bu bölgelerde fonksiyonlar oluşturur (Friedman, 1993; Leathwick et al., 2006; Toprak, 2011; Weber et al., 2012). Bu bölgeler, düğüm noktaları ile birbirinden ayrılır ve iki düğüm arasındaki fonksiyon, temel fonksiyon (basis function) olarak adlandırılır. Temel fonksiyonlar, bağımlı değişken ile parçalı doğrusal (piecewise linear) bir ilişki içindedir. Farklı katsayılar ve etkileşim terimleri kullanılarak belirlenen aralıklarda model oluşturulması, veri setinin gerçek yapısını daha hassas bir şekilde yansıtarak daha doğru sonuçlar elde edilmesine olanak tanır. Bu nedenle, MARS yöntemi, analizlerdeki esneklik ve geniş kapsamlı modele uyum yeteneği açısından değerli bir araç olarak kabul edilir. MARS algoritması bağımsız değişkenler arasındaki etkileşimi de dikkate alarak bağımsız değişkenlerin modele girip girmeyeceğini belirlemektedir ve robust modeller üretir (Friedman, 1993; Stevens, 1991; Özfalci, 2008; Temel vd, 2010; Sevgenler, 2019).

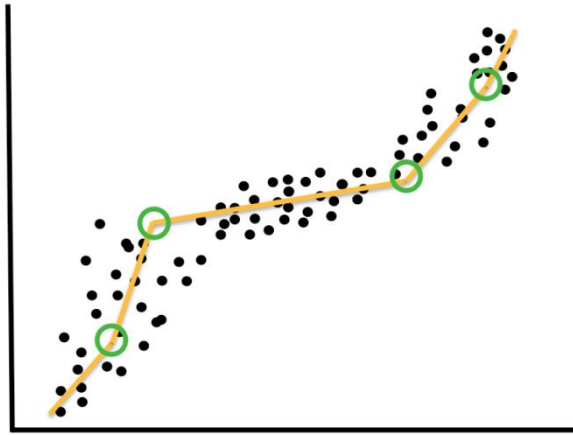
MARS, yüksek boyutlu parametrik olmayan regresyon modelleri oluşturabilen esnek bir yaklaşım sunar. Bu parametrik olmayan regresyon yöntemi, bağımlı ve

bağımsız değişkenler arasındaki temel fonksiyonel ilişki konusunda özel bir varsayım yapmaz. Yani, model oluşturulurken belirli bir matematiksel formül veya ilişki yapısal kısıtlamalara tabi değildir.

2.3.1. Düğüm Noktası

MARS yöntemi, değişkenler arasındaki ilişkileri dikkate alarak değişimlerin olduğu belirli noktaları belirler ve bu noktalar arasında uygun dönüşümlerle düz bir doğru elde etmeye çalışır. Modeldeki değişimin meydana geldiği, yani regresyon doğrusundaki eğimin değiştiği noktalara "düğüm noktası" denir. Her bir bağımsız değişken için düğüm noktalarının belirlenmesi, hızlı ancak çok yoğun bir arama prosedürü kullanılarak gerçekleştirilir.

Şekil 2.4'te, bağımsız değişken x ve bağımlı değişken y arasındaki ilişkinin dağılımını göstermektedir. Bu örnekte dört adet düğüm noktası bulunmaktadır ve her düğüm noktasında fonksiyonun eğiminin değiştiği gözlemlenmektedir. MARS, bu düğüm noktalarını tespit ederek, veri setindeki karmaşık ilişkileri yakalamaya ve modelin esnek bir şekilde uyarlanmasına olanak tanır. Bu sayede, regresyon doğrusundaki eğimin belirli aralıklarda değişim göstermesi, gerçek veri setinin özgün yapısını daha iyi yansıtabilir ve daha doğru tahminler elde edilebilir. (Friedman, 1991; Özfalcı, 2008; Ünal, 2009).



Şekil 2.4 MARS modelindeki düğüm noktaları (Anonim 2019)

Düğüm noktası belirleme işlemi için ilk olarak sabit fonksiyon seçilir (Eşitlik 2.6);

$$f_j = c_j \quad (2.6)$$

c_j sabiti bağımlı değişkenin ortalaması olarak alınır (Eşitlik 2.7).

$$c_j = \frac{1}{N} \sum y \quad (2.7)$$

İlk düğüm noktası bulunduktan sonra Lack of Fit (LOF) kriteri geçerli olur (Eşitlik 2.8);

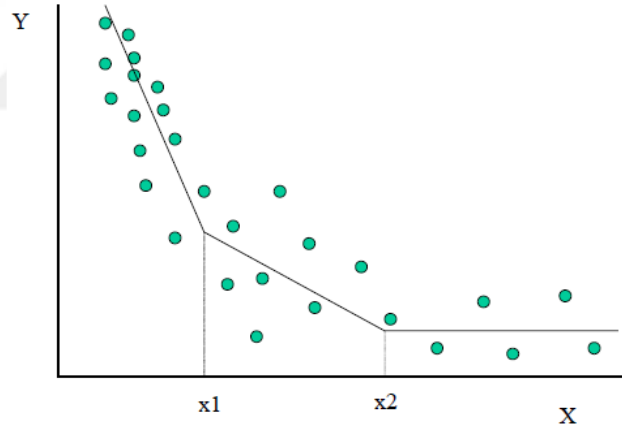
$$LOF[f_j(x)] = \min \sum (y_i - c)^2 \quad (2.8)$$

LOF kriteri toplamı minimum yapan c noktasının bulunması işlemidir. Bu iterasyon istenilen düğüm noktası sayısına ulaşılan kadar devam etmektedir (Friedman, 1993).

2.3.2. Temel Fonksiyonlar

Birçok gerçek dünya problemi, parçalı doğrusal olan sürekli fonksiyonları içerir. Şekil 2.5'te gösterildiği gibi üç doğru parçasına sahip parçalı doğrusal bir eğriyi modellemek için genellikle ikili değişkenler kullanılır.

MARS algoritması veriler arasındaki ilişkiyi parçalı fonksiyonlarla belirler, böylece tüm değerlere en yakın geçecek şekilde yani ilişkiyi en güçlü şekilde modelleyen temel fonksiyonlar oluşturur. Bağımsız değişken değerlerini bölgelere ayırarak (Şekil 2.5) her bölgeyi bir temel fonksiyon ile ifade etmektedir (Akyol, 2011; Özkurt, 2013).

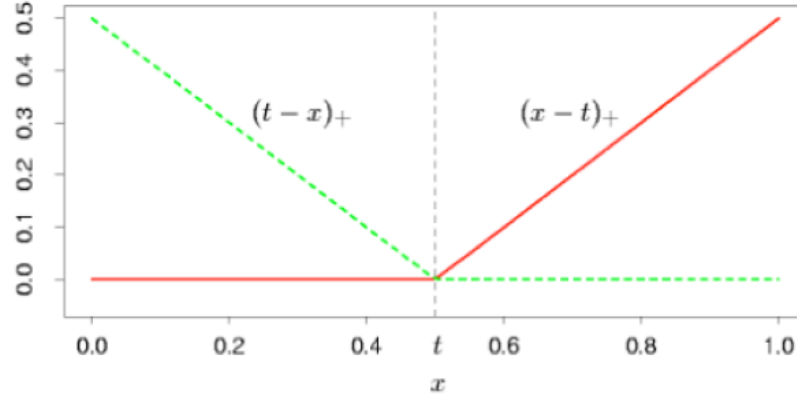


Şekil 2.5 MARS modelinde düğüm noktaları arasındaki temel fonksiyonlar (Ünal, 2009).

MARS modeli $[x - t]_+$ ve $[t - x]_+$ formundaki temel fonksiyonlar kullanılır (Eşitlik 2.9). Bu temel fonksiyonlar bağımsız değişkenlerin bağımlı değerleri en iyi düğüm noktalarında bölen parçalı regresyon eğrileridir ve yansıması olan çiftler olarak adlandırılır. Alt indis olarak gösterilen “+” pozitif tarafı temsil etmektedir.

$$[x - t]_+ = \begin{cases} x - t, & x > t \\ 0, & x < t \end{cases} \quad [t - x]_+ = \begin{cases} t - x, & x < t \\ 0, & x > t \end{cases} \quad (2.9)$$

Örnek olarak düğüm noktası 0.5 olan $(x-0.5)_+$ ve $(0.5-x)_+$ temel fonksiyonları Şekil 2.6’da gösterilmiştir.



Şekil 2.6 Temel fonksiyonlar (yansıma çifti) (Özmen, 2010)

Şekil 2.6'da gösterilen parçalı fonksiyonlar, MARS algoritmasında "yansıma çifti (reflected pair)" olarak adlandırılır. Her bir düğüm noktasında, MARS algoritması tarafından bir yansıma çifti oluşturulur. Bu yansıma çiftleri, düğüm noktaları arasında parçalı fonksiyonları temsil eder. Her yansıma çifti, belirli bir aralıktaki değişkenler arasındaki ilişkiyi yakalayarak, modelin karmaşıklığını düşürmeye ve veri setinin doğasını daha iyi anlamaya yönelik bir adaptasyon sağlar (Stevens, 1991; Goepp et al., 2018)

Temel fonksiyonların gösterimi Eşitlik 2.10'da gösterilmektedir.

$$B = \{[X_j - t]_+, [t - X_j]_-, t \in \{x_{1j}, x_{2j}, x_{3j}, \dots, x_{Nj}\}, j = 1, 2, 3, \dots, p\} \quad (2.10)$$

Eşitlik 2.10'da N veri sayısı, p girdi boyutunu temsil etmektedir. Tüm veri seti kullanılarak toplam temel fonksiyon sayısı ise $2Np$ formülü ile bulunmaktadır. Temel fonksiyonları bulunmasında belirli bir temel fonksiyonu bulmak için Eşitlik 2.11 kullanılmaktadır.

$$B_m(X) = \prod_{k=1}^{K_m} H[s_{km}(x_{v(k,m)} - t_{km})]_+ \quad (2.11)$$

burada K yansıma çifti sayısı

$x_{v(k,m)}$ kesişen lineer fonksiyona karşılık gelen bağımsız değişken

t_{km} düğüm noktası

$s_{km} = \pm 1$ olarak alınır ve temel fonksiyonu pozitif yapmak için kullanılır.

MARS modelinin genel fonksiyon gösterimi ise Eşitlik 2.12'de gösterilmektedir.

$$f(x) = \beta_0 + \sum_{m=1}^M \beta_m B_m(x) + \varepsilon \quad (2.12)$$

Bu eşitlikte;

$B_m(x)$ m'inci temel fonksiyonu

M temel fonksiyon sayısını temsil etmektedir.

MARS modelinde ilk temel fonksiyon 1 olarak hesaplanır (Eşitlik 2.13) (Friedman, 1993).

$$B_0(X) = 1 \quad (2.13)$$

Temel fonksiyonları olası yapısı Eşitlik 2.14'te gösterilmektedir.

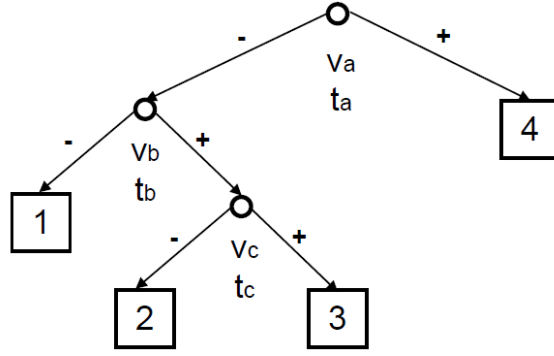
- 1
 - x_k
 - $[x_k - t_i]_+$
 - $x_k x_l$
 - $[x_k - t_i]_+ x_l$
 - $[x_k - t_i]_+ [x_l - t_j]_+$
- (2.14)

MARS çok sayıda değişkeni bir arada bağımsız ve etkileşimli olarak modele dahil edebilen, güçlü algoritmalarla parametre tahmin edilmesine imkan sağlayan parametrik olmayan yöntemlerdendir. MARS yöntemi regresyon modelini oluştururken ileri ve geri adım algoritması diye adlandırılan iki aşamalı bir algoritma kullanmaktadır (Friedman et al., 2001; SPM Guide Help, 2018).

2.3.3. İleri Adım

MARS algoritması, tüm değişkenler arasındaki potansiyel bütün etkileşimleri içeren her düğüm konumunu belirler. Bu nedenle, ilk aşama olan ileri adımda, değişkenleri ve potansiyel düğümleri test ederek aşırı doymuş (overfit) bir modele ulaşılabilir. Bu aşırı doymuş model, veri setindeki her bir değişken ve düğüm kombinasyonunu içerir, ancak bu durum modelin genelleme yeteneğini zayıflatır. İleri adım sürecinde, bağımsız değişkenlerin en uygun düğüm noktalarının tespit edilip aralıklara bölünmesiyle oluşturulan parçalı doğrusal regresyon eğrilerine "temel fonksiyon" denir. Bu temel fonksiyonlar, değişkenler arasındaki etkileşimleri yakalamak ve her bir düğüm noktasında regresyon eğrisini modellenmesini sağlamak amacıyla kullanılır. İleri adım sürecini göstermek için ikili ağaç diyagramı çizilebilir (Şekil 2.7) (Breiman et al., 1984; Hastie et al., 2001; Ünal, 2009).

Temel fonksiyonların sürekli çift olarak eklenmesi karmaşık bir model oluşturur ve iterasyon halinde devam eden bu işlem kullanıcının belirlediği temel fonksiyon sayısına ulaşıncaya kadar devam etmektedir. En üst seviyeye ulaşıncaya kadar devam eden iterasyon ardından ikinci adım olan geri adıma geçilir (Hastie et al., 2001).



Şekil 2.7 İkili ağaç diyagramı (Hastie et al., 2001)

2.3.4. Geri Adım

İkinci aşama olan geri adımda gereksinim olmayan faktörler ve bileşenler elenmektedir. Hata kareler toplamı minimum olması amacıyla; önemli bağımsız değişkenler ve bu değişkenlerin etkileşimleri belirlenir. Optimum modeli elde edebilmek için bazı temel fonksiyonlar modelden çıkarılmaktadır. İleri adımda temel fonksiyonlar modele çift olarak eklenirken geri adımda çiftin bir tarafı yok sayılır, yani son modelde temel fonksiyonlar çift değildir. Bu eleme yöntemi, Craven ve Wabha (1978) tarafından geliştirilen “genelleştirilmiş çapraz geçerlilik” (generalized cross validation – GCV) kriterini temel almaktadır. GCV kaç tane temel fonksiyon olacağını, aynı zamanda modeldeki uyumsuzluğu (Lack Of Fit – LOF) gösterir (Eşitlik 2.15). Regresyon analizinde, gözlemlenen veriler ile ortalama fonksiyon arasındaki fark rastgele olmayan bir bileşene sahiptir ve bu bileşen LOF olarak adlandırılır (Friedman, 1993; Akyol, 2011; Özkurt, 2013)

$$LOF(f) = GCV(\alpha) = \frac{1}{N} \frac{\sum_{i=1}^N (y_i - f_{\alpha}(x_i))^2}{(1 - M(\alpha)/(N))^2} \quad (2.15)$$

$$M(\alpha) = u + dk$$

k = seçilen düğüm sayısı

d= optimal temel fonksiyonları elde etmek için kullanılan bir ceza parametresi $2 < d < 5$ en ideali 3.

U= temel fonksiyonların sayısı

N= veri setindeki bağımsız değişken sayısı

$GCV(\alpha)$ pay hata kareler toplamını, payda ise modelin karmaşıklığını hesaplamaktadır. Geriye doğru adım algoritmasında bağımsız değişkenler ve değişkenlerin etkileşimleri belirlenerek $GCV(\alpha)$ sonucunu kendi minimumuna ulaşana kadar iterasyon devam eder ve GCV değerini minimum yapan model seçilir (Friedman, 1993).

Ayrıca Friedman (1991) GCV haricinde R^2 kriterini de karşılaştırma ölçütü olarak önermektedir ve bir model, R^2 değeri önemli ölçüde yüksekse tercih edilmektedir. Optimum düğüm konumları ve temel fonksiyon sayısı belirlendiğinde, seçilen temel fonksiyonlar üzerinden uygun modelin tahminlerini yapabilmek için en küçük kareler yöntemi kullanılır. Tek bir bağımsız x_i değişkeni ile y_i bağımlı değişkeni arasındaki ilişki modellendiğinde Eşitlik 2.16 oluşturulur (Sephton 2001);

$$Y_i = \beta_0 + \sum_{k=1}^K \beta_k B_k(X_i) + \epsilon_i \quad (2.16)$$

Burada

$B_k(X_i)$ = X_i 'nin k 'inci temel fonksiyonudur.

β_0 = modeldeki sabit terim

β_k = temel fonksiyon katsayısı

K = temel fonksiyon sayısı

İleri ve geri adımları tamamlandıktan sonra MARS genel formülü Eşitlik 2.17'de gösterilmektedir.

$$f(x) = \beta_0 + \sum_{K_m=1} f_i(x_i) + \sum_{K_m=2} f_{ij}(x_i, x_j) + \sum_{K_m=3} f_{ijk}(x_i, x_j, x_k) + \dots \quad (2.17)$$

Burada birinci toplam tek değişkenli bütün temel fonksiyonları içerir. İkinci toplam ise iki değişkeni ve varsa bu değişkenler arasındaki etkileşimleri içeren bütün temel fonksiyonları içerir. Bu mantık benzer şekilde devam eder. Fonksiyonların yapısını ise Eşitlik 2.18'de gösterilmektedir (Friedman, 1991);

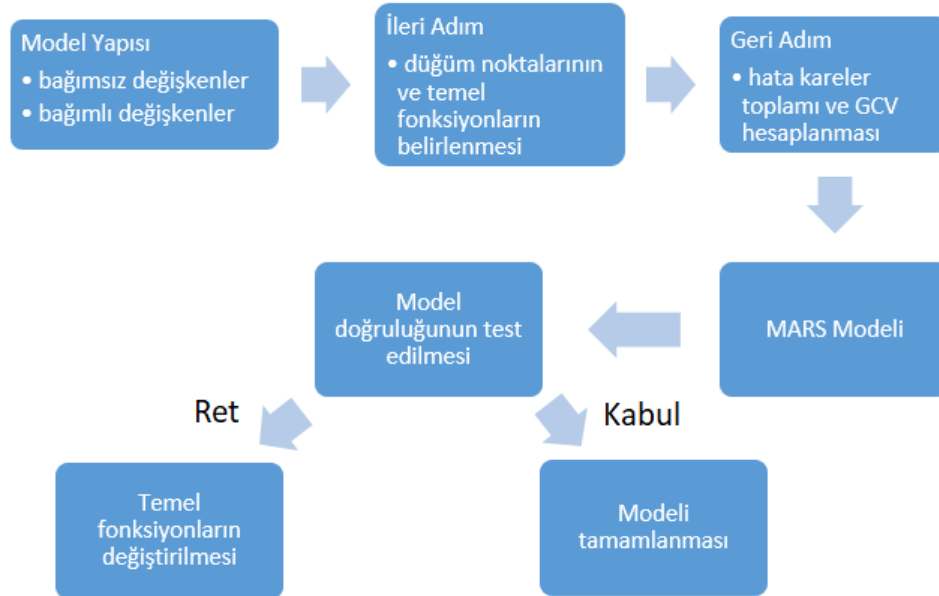
$$\begin{aligned} f_i(x_i) &= \sum_{K_m=1} \beta_m B_m(x_i) \\ f_{ij}(x_i, x_j) &= \sum_{K_m=2} \beta_m B_m(x_i, x_j) \\ f_{ijk}(x_i, x_j, x_k) &= \sum_{K_m=3} \beta_m B_m(x_i, x_j, x_k) \end{aligned} \quad (2.18)$$

MARS algoritmasında düğüm nokta sayısı ve bu noktaların konumları önceden belirlenmez. İleri ve geri adımsal süreç kullanılarak düğüm noktalarını belirlemek için en iyi noktalar araştırılmaktadır. İleri adımda aşırı doymuş model, geri adımda GCV kullanarak sadeleştirilir ve düğüm noktaları belirlenir (Friedman, 1991; Koç, 2012; Sevgenler, 2019).

MARS algoritmasının adımları özetlenecek olursa;

- İleriye adım
 1. En basit modelle yani sadece bir sabit katsayı ile başlanır.
 2. Her açıklayıcı/bağımsız değişken için temel fonksiyon araştırılır.
 3. Tahmin hatasını minimum yapan temel fonksiyon yansıma çiftleri belirlenir ve modele dahil edilir.
 4. Modelin önceden kullanıcı tarafından belirlenen karmaşıklığa ulaşana kadar adım 2'ye gidilir.
- Geri adım
 5. Sabit fonksiyon hariç temel fonksiyonlar araştırılır ve GCV kısıtı kullanılarak modele en az katkıda bulunan temel fonksiyonlar silinir.
 6. GCV sonucu kendi minimumuna ulaşana kadar 5. adım tekrarlanır.

MARS algoritmasının işlem adımları Şekil 2.8'de görselleştirilmiştir. Bağımlı ve bağımsız değişkenler kullanılarak düğüm noktaları ve temel fonksiyonlar belirlenir, sonrasında hata kareler toplamı ve GCV ile geri adım gerçekleştirilir. Oluşturulan model test edildikten sonra ret edilirse bağımlı değişkenler tam ve doğru şekilde açıklanamamıştır demektir. Kabul gören model ise tahmin veya sınıflandırma için yeterli doğruluğa sahiptir.



Şekil 2.8 MARS modeli işlem adımları

2.3.5. MARS Algoritmasının Avantaj ve Kısıtları

MARS algoritmasının temel kısıtı parametreleri belirlerken arama uzayını rastgele örnekleyerek çalıştığı için önceki iterasyonlardan elde edilen bilgilerin

kullanılmamasıdır. Bu durum özellikle fonksiyonun bilinmediği parametre seçim problemlerinde sorun yaratmaktadır, ayrıca parametre tahmini için yapılan iterasyonlardan dolayı analiz süresi uzun olmaktadır. Optimizasyon yöntemleri bu noktada devreye girer ve parametrelerin belirlenmesinde ilk iterasyonda rastgele olsa dahi sonraki iterasyonlarda ilk iterasyon sonucunu kullanarak en uygun parametre değerine ulaşır.

MARS algoritmasının avantajları (Friedman, 1991; Yazıcı, 2011);

- Bağımlı ve bağımsız değişkenler arasındaki ilişkide var olabilecek doğrusal olmayan durumların tespitini sağlar.
- Bağımlı ve bağımsız değişkenler yani veri dağılımları hakkında herhangi bir varsayım bulunmamaktadır.
- Bağımsız değişkenler arasında uygun etkileşimleri otomatik olarak arar ve çok sayıda etkileşimli değişken olduğunda tercih edilir.
- Herhangi bir bağımsız değişken modele yeterince katkıda bulunmuyorsa, model karmaşıklığını azaltmak için bağımsız değişkeni modelden çıkarabilir.
- Çoklu bağlantı probleminde diğer regresyon yöntemlerine göre çok az etkilenir.
- Orijinal verilerin alt kümelerini temel alır, bu nedenle modelin yorumlanması kolaydır.
- Kayıp değerlerden ve/veya aşırı uç değerlerden diğer yöntemlere göre daha az etkilenmektedir.
- Bağımsız değişkenleri parçalara ayırarak, farklı aralık değerleri için farklı katsayılar türeterek yüksek tahmin performanslı modeller geliştirir.

MARS algoritmasının kısıtları (Yazıcı, 2011; Friedman, 1991);

- Modelin yorumlanması zorlayıcıdır.
- Büyük veri setlerinde daha başarılı olduğu için, genellikle büyük veri setlerine ihtiyaç duyar.
- Optimum temel fonksiyon sayısı ve ceza parametresi belirleme zorlayıcıdır.

2.3.6. Mekânsal Analizlerde MARS Algoritmasının Kullanımı

Regresyon ve sınıflandırma yeryüzü ile ilgili çeşitli bilgileri elde etmek için kullanılan Coğrafi Bilgi Sistemleri (CBS) ve uzaktan algılama teknolojilerinde önemli bir rol oynamaktadır. Uzaktan algılama verilerinden bilgi çıkarımı için birçok

çeşit regresyon ve sınıflandırma yöntemi uygulanmaktadır. Böylece mekânsal olgular daha iyi anlaşılır ve karar vermek için modellenir. Ayrıca, olaylar başka bölgelerde veya zamanlarda tahmin edilebilecek şekilde modellenir. Gerçek yaşam problemleri ve çoğu etkinin genellikle doğrusal olmayan davranış sergilediği doğal olayları açıklayan temel veri madenciliği araçları parametrik olmayan regresyon ve sınıflandırma yöntemleridir (Kuter, 2014). Regresyon yöntemlerinin mekânsal analizlerde kullanılmasının sebepleri;

- Olayları daha iyi anlamak ve karar vermek için modellemektir.
- Olayları başka bölgelerde veya zamanlarda tahmin edilebilecek şekilde modellemektir.
- Hipotezleri belirlemektir.

Sınıflandırma yöntemleri parametrik ve parametrik olmayan olmak üzere ikiye ayrılmaktadır. Parametrik sınıflandırma yöntemlerinde Gauss dağılımı gibi başlangıç koşulları uzaktan algılama görüntülerinde bulunmaz. Eğer veriler yüksek bir boyutluluğa sahipse, parametrik yöntemlerin öğrenme aşaması için birçok örnek gereklidir. Yapay sinir ağları ve karar ağaçları gibi parametrik olmayan sınıflandırma yöntemleri ise veri dağılımı hakkında herhangi bir varsayım yapmadan analiz yapmaktadır. Birçok çalışma parametrik olmayan yöntemlerin daha iyi sınıflandırma sonuçları sağladığını göstermiştir (Friedl and Brodley, 1997; Foody et al., 2007; Lu and Weng, 2007; Mather and Koch, 2011). Yapılan literatür incelemelerinde, küçük eğitim örnekleriyle bile, parametrik olmayan sınıflandırma algoritmalarının parametrik olanlardan daha iyi sonuçlar sağladığı görülmüştür. Parametrik olmayan regresyon yöntemi olan MARS algoritması tıp, biyoloji, ekonomi, ziraat alanlarında, ayrıca simülasyon ve tahmin modellenmesinde özellikle veri madenciliği alanında sıklıkla kullanılmaktadır, yapılan yeni çalışmalar mekânsal analizlerde de kullanılabilirliğini kanıtlamıştır (Quirós et al., 2009; Kuter, 2014; Kuter vd, 2015; Park et al., 2017). Bu sav literatür araştırması bölümündeki (Bölüm 1.4) örneklerle desteklenmektedir.

2.4. Rastgele Orman Algoritması (Random Forest – RF)

2.4.1. Topluluk Öğrenmesi

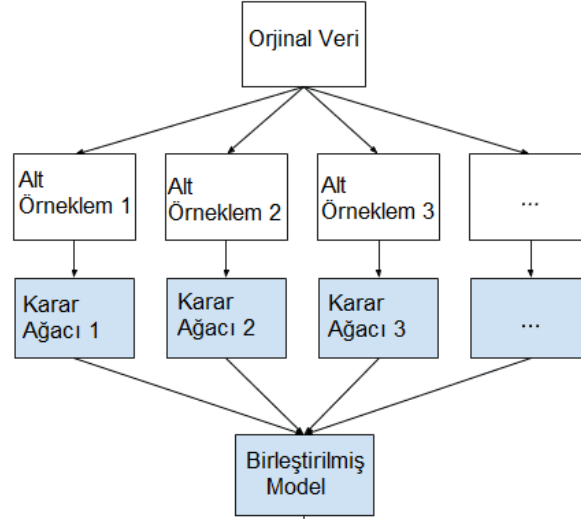
Bir hibrit model ile bir topluluk modelinin kullanılması, probleme ve mevcut verilerin özelliklerine bağlıdır. Farklı model türlerinin entegre edilmesini veya farklı verilerin işlenmesini gerektiriyorsa, hibrit bir model daha uygun olabilir. Öte yandan, amaç varyansı azaltmak ve istikrarı arttırmaksa, bir topluluk modeli daha iyi bir

seim olabilir. Topluluk modelleri, aynı modelin birden fazla rneęini birleřtirerek varyansı azaltma ve kararlılıęı arttırır. Byk veri kmelerini verimli iřleyebilir ve ařırı uyum riskini azaltabilirler (Breiman, 1996; Lee et al., 2020).

Topluluk ęrenimi, iki veya daha fazla modelden elde edilen tahminleri birleřtirerek tahminlerde doęruluęu ve esneklięi arttıran bir makine ęrenimi teknięidir. Topluluk ęrenimi modelin genel performansını arttırır ve bu yaklařım yalnızca doęruluęu arttırmakla kalmaz, aynı zamanda verilerdeki belirsizliklere karřı dayanıklılık da saęlar (Breiman, 2001; Pal and Mather, 2005).

Birok farklı topluluk ęrenmesi yntemi bulunmaktadır; maksimum oylama, ortalama, aęırlıklı ortalama, yıęma (stacking), torbalama (bagging), ykseltme (boosting). Gnmzde bu yntemlerden torbalama yntemi uydu grnts sınıflandırma analizlerinde kullanılan DT yntemine uygulanarak RF algoritması ortaya ıkmıřtır (Breiman, 1996; Breiman, 2001).

1996 yılında Breiman tarafından geliřtirilmiř olan Torbalama (bootstrap aggregating - bagging) yntemi, orijinal veri setinden elde edilen nykleme rneklerine tahmincilerin uygulanmasıyla bir topluluk oluřturulur. nykleme iřlemi, iadeli rastgele seim yntemiyle alt rneklem oluřturmak iin kullanılır. Bu alt rneklem, orijinal veri setindeki gzlem sayısıyla aynı olacak řekilde oluřturulur. Bu srete, bazı gzlemler alt rneklemlerde hi yer almazken, bazıları birden fazla kez seilebilir. Tahminlerin birleřtirilmesi ařamasında, regresyon aęaları iin tahminlerin ortalaması alınırken, sınıflandırma aęalarında sonular oylama yntemiyle belirlenir. Torbalama yntemi ile DT ynteminden uyarlanan RF yntemi řekil 2.9’da gsterildięi gibi, alt rneklemlerin oluřturulması, bu alt rneklemlerden karar aęalarının oluřturulması ve orman oluřturan karar aęalarının birleřtirilmesidir. Torbalama, tutarsız tahminci deęiřkenlerin tahmin doęruluęunu arttırabilir. Dřk yanlılık miktarına sahip ancak yksek varyanslı deęiřkenleri kullanarak, bu deęiřkenleri daha elveriřli hale getirir. Ayrıca, deneysel sonular torbalama ynteminin, tekil aęalara gre daha etkin sonular verdięini gstermiřtir (Breiman, 1996; Lee et al., 2020).

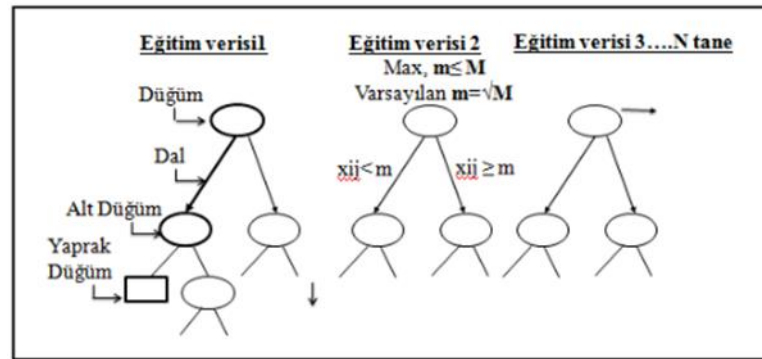


Şekil 2.9 Torbalama yöntemi (Lee et al., 2020)

2.4.2. RF Algoritması

DT yöntemi, sınıfları birbirinden ayırt etmek için sınıflandırma problemini katmanlı bir yapıya dönüştürmektedir. RF algoritması DT algoritmasından türetilmiş ve eğitim aşamasında birden çok karar ağacının kullanılmasıyla karar ormanı oluşturulur. Matematiksel olarak kolay ifade edilebilen ve programlanabilen yapısı nedeniyle RF, birden fazla algoritmanın hibrit kullanıldığı sınıflandırma modellerinin oluşturulmasında tercih edilen algoritmaların başında gelmektedir (Breiman, 2001; Pal and Mather 2005;).

Makine algoritması olan RF, Şekil 2.10’da gösterildiği gibi her bir karar ağacı bir başlangıç düğümü içerir. Bu başlangıç düğümü, dal adı verilen kısımlar aracılığıyla alt düğümlere bağlanır. Dallanmalar, yaprak düğümlere kadar devam eder. Alt düğümlere bölme işlemi gerçekleştirirken, rastgele bir öznelik alt kümesi seçilmektedir (Breiman et al., 1984; Breiman, 2001; Pal and Mather, 2005; Mather and Koch, 2011).



Şekil 2.10 RF sınıflandırması (Özdarıcı Ok vd., 2011)

RF algoritması birden fazla DT yapısının kullanılması ve her bir DT tahmininden yola çıkılarak sonuç tahminlerin yapılması esasına dayanmaktadır (Pal and Mather 2005). RF sınıflandırıcı, her sınıflandırıcı, giriş vektörlerinin eğitim setinden bağımsız olarak örneklenen rastgele bir vektör kullanılarak üretilir. Her ağacın, belirli bir girdi vektörünün yerleştirileceği en popüler sınıf için bir birim oyu attığı ağaç sınıflandırıcılarının bir kombinasyonundan oluşur. (Breiman, 2001).

RF oluşturma tekniği genellikle torbalama ve rastgele alt uzay yöntemlerinin bir kombinasyonudur. Torbalama (Breiman, 1996), sınıflandırma doğruluğunu arttırmak ve aşırı uydurmayı önlemek için kullanılan bir yöntemdir. N boyutunda bir eğitim seti verildiğinde, torbalama, orijinal eğitim setinden değiştirilerek rastgele numuneler çizerek her biri n boyutunda (burada $n < N$) bir dizi yeni eğitim seti oluşturur. Ayrıca verilerin M nitelikleri (örn. Spektral bantlar) içerdiğini varsayalım. Ağacın her bir düğümü için, m ($m < M$) öz nitelikleri, o düğümdeki en iyi bölünmeyi hesaplamak için temel sağlamak üzere rastgele seçilir. Daha sonra her örnek, ormandaki tüm ağaç belirleyicilerinden en popüler oyları alan sınıfa atanır (Mather and Koch, 2011).

2.4.3. RF Algoritmasının İşlem Adımları

RF algoritması, veri setindeki en iyi bölme noktasını belirleyerek her düğümde en uygun dalı seçer. Bu işlem, rasgele seçilmiş veri setleri üzerinde gerçekleştirilir, her biri orijinal verinin yer değiştirilmesiyle oluşturulan bu veri setleri üzerinde rasgele ağaçlar türetir. Her bir ağaç, farklı bir veri seti üzerinde eğitilir. Algoritma eğitim aşamasında, tam veri setinden rasgele seçilen 2/3'lük bir kısmı kullanılır ve bu alt kümeler üzerinde ağaçlar oluşturulur. Eğitim setinin geri kalan 1/3'lük kısmı ise algoritmanın test edilmesi için kullanılır. Bu test verisi "Out Of Bag" (OOB) verisi olarak adlandırılır. Bu yaklaşım, ağaçların oluşturulmasında kullanılan veri setiyle eğitim ve test aşamalarında farklı veri setlerinin kullanılmasına dayanır (Archer, 2008; Belgiu and Dragut, 2016).

RF algoritması eğitim veri setine ait birden fazla karar ağacı içerir. RF sınıflandırıcısı bir ağaç oluşturmak için belirli parametreler gerektirir, bunlar arasında her bir düğümde kullanılan değişkenlerin sayısı ve oluşturulacak ağaçların sayısı yer alır. Bu parametreler, modelin çeşitliliğini ve genelleme yeteneğini arttırarak doğruluğunu iyileştirir. RF, bu parametreleri kullanarak bir dizi karar ağacı oluşturur ve sonuçları bir araya getirerek güçlü ve genelleştirilebilir bir sınıflandırma modeli elde eder. Kullanıcı tarafından en iyi bölünmeyi belirlemek için başlangıç

düğümde kullanılan değişkenlerin sayısı rastgele seçilir, sonraki değişken sayısı genelleştirilmiş hatalara göre artırılır ya da azaltılır. Bu esneklik, modelin karmaşıklığını ve genelleme yeteneğini kontrol etmeye yardımcı olur. Değişkenlerin sayısının azaltılmasıyla korelasyon ve güç azalırken, m'nin artırılmasıyla korelasyon ve güç artar. RF algoritması, bu adaptasyon sayesinde veri setine uygun ve dengeli bir sınıflandırma modeli oluşturabilir (Pal and Mather, 2005; Saini and Ghosh, 2017).

RF algoritmasında, öncelikle orijinal veri seti üzerinde önyükleme (bootstrap) yöntemleri kullanılarak eğitim ve test veri setleri oluşturulur. Eğitim veri setinden rastgele seçim yapılarak en büyük genişliğe sahip bir karar ağacı oluşturulur. Eğitim ve test veri setlerindeki sınıfların homojenliğini belirlemek için Gini indeksi, entropi, yanlış sınıflandırma hatası ve kazanç oranı gibi kriterler kullanılır. Gini indeksi, homojenlik durumunu belirlemek için sıkça kullanılır ve değeri düştükçe sınıfların homojenliği artar. Bu indeks, Eşitlik 2.19'de belirtilen formülle hesaplanır (Breiman, 2001; Mather and Tso, 2016).

$$Gini = 1 - \sum_{i=1}^N \left(\frac{f(C_i, T)}{|T|} \right)^2 \quad (2.19)$$

Burada C_i , rastgele seçilen örneğin sınıfını; T , eğitim veri setini; f , seçilen örneğin C_i sınıfına ait olma olasılığını ifade etmektedir. Bir alt düğümün Gini indeksi, üst düğümün Gini indeksinden daha düşük olduğunda, o dalın başarılı olduğunu gösterir. Gini indeksi hesaplandıktan sonra, her düğüme bir sınıf atanır ve ağaç, veri seti ile test edilir. Test sonrası oluşan sınıflar kaydedilir ve bu işlemler modeldeki tüm ağaçlar için tekrarlanır. Test verisi ile yapılan değerlendirme sonucunda hangi kategorilerde kaç defa sınıflandırıldığı tespit edilir. Bu sonuçlar, eğitim veri setindeki sınıflandırmalar ile karşılaştırılır. Bu süreç, modelin doğruluğunu ve performansını değerlendirmek ve sınıflandırma algoritmasının etkinliğini ölçmek için kullanılır.

2.4.4. RF Algoritmasının Avantaj ve Kısıtları

RF algoritmasının avantajları;

- RF algoritması hem regresyon hem de sınıflandırma problemlerinde kullanılabilir böylece geniş bir uygulama alanına sahiptir.
- DT uygulamalarında karşılaşılan aşırı uyum problemi RF ile azaltılmaktadır. Aşırı uyum probleminin azalmasının nedeni eğitim aşamasında farklı veri setlerinin oluşturulmasıdır.

- Her bir karar ağacının tahmin üzerindeki doğruluğunu ölçmeye imkan vermektedir.
- Her düğümde farklı eğitim verileri ve rastgele tahminciler ile dallanma gerçekleştirildiği için oluşan ağaçlar birbirinden bağımsızdır.
- Kategorik ve sürekli değerlerle iyi çalışır
- Kural tabanlı bir yaklaşım kullandığından verilerin normalleştirilmesi gerekli değildir.

RF algoritmasının kısıtları;

- Çok sayıda ağaç oluşturduğu için hesaplama gücü fazlalığı ve kaynak gerektirir.
- Her bir değişkenin önemini belirleyemez. Başka bir deyişle karar ağaçları oluşturmak için uygun değişkenlerin seçimi zordur.
- Çok sayıda ağaç oluşturduğu için çok fazla hesaplama gücü gerektirir.
- Sınıfı belirlemek için birçok karar ağacını birleştirdiğinden dolayı eğitim aşamasında çok zaman gerektirir.

2.5. Optimizasyon

Optimizasyon, bir sonlandırma koşuluna ulaşılan kadar her örneklemede verilere ekleme yaparak bir örnekleme politikasının tekrar tekrar kullanılmasıdır. Optimizasyon teknikleri yüzyılı aşkın bir süredir mevcuttur. Başlangıçta, diferansiyel hesap, birçok pratik durumda ve teorik problemde ortaya çıkan fonksiyonların maksimum veya minimumlarını bulmak için uygulanmıştır. İlk olarak, George B. Dantzig, 1947 yılında doğrusal programlama problemlerinin çözümü için simpleks algoritmasını geliştirmiş ve doğrusal programlama, optimizasyonda kullanılan temel tekniklerden biri olmuştur. Pratik hesaplama algoritmalarının sistematik gelişimi ise 1950'li yıllarda gerçekleşmiş ve bilgisayar teknolojileri ile birlikte 1980'li yıllarda gelişimi hızlanmıştır (Liberti and Maculan, 2006; Sarker and Newton, 2007; Boyd and Parrilo, 2009; Bergstra et al., 2011)

Son yarım yüzyılda, geliştirilen birçok klasik optimizasyon tekniğine ek olarak, benzetimli tavlama, tabu arama, genetik algoritmalar, sinirsel hesaplama, bulanık mantık ve karınca kolonisi optimizasyonu gibi modern sezgisel teknikler de geliştirilmiştir. Bu modern teknikler, uygulayıcılara daha karmaşık durumları ele almak için yeni yöntemler sağlamaktadır. Optimizasyon sürecinin amacı, karar verme ve modelleme süreçlerinin gerçekçi ve pratik sonuçlarının belirlenmesine

yardımcı olmaktadır. Optimizasyon sürecini adımlara ayırırsak; temel problemin tanımlanması ve formüle edilmesi, matematiksel modelin kurulması, en uygun çözümün seçilmesi, çözümün analiz edilmesi ve tercih edilen optimizasyon modellemesinin uygulanması olarak sıralanabilir. Günümüzde, uydu görüntülerinin sınıflandırılmasında kullanılan yöntemlerin içerdiği fonksiyon parametrelerinin belirlenmesi önem arz ettiği ve en iyi parametrelerin nerede aranacağı konusunda daha iyi seçimler yapmak için optimizasyon yöntemleri kullanmak etkili bir çözüm olmaktadır. Bazı basit matematiksel ifadelerde, birinci türevi sıfıra eşitleyerek denklemi çözmek yeterli olabilir. Ancak, problem türleri, türev alınabilirlik durumu ve kısıtlamalar göz önüne alındığında, optimizasyon için geliştirilmiş birçok yöntem mevcuttur (Liberti and Maculan, 2006, Zhigljavsky and Zilinskas, 2008).

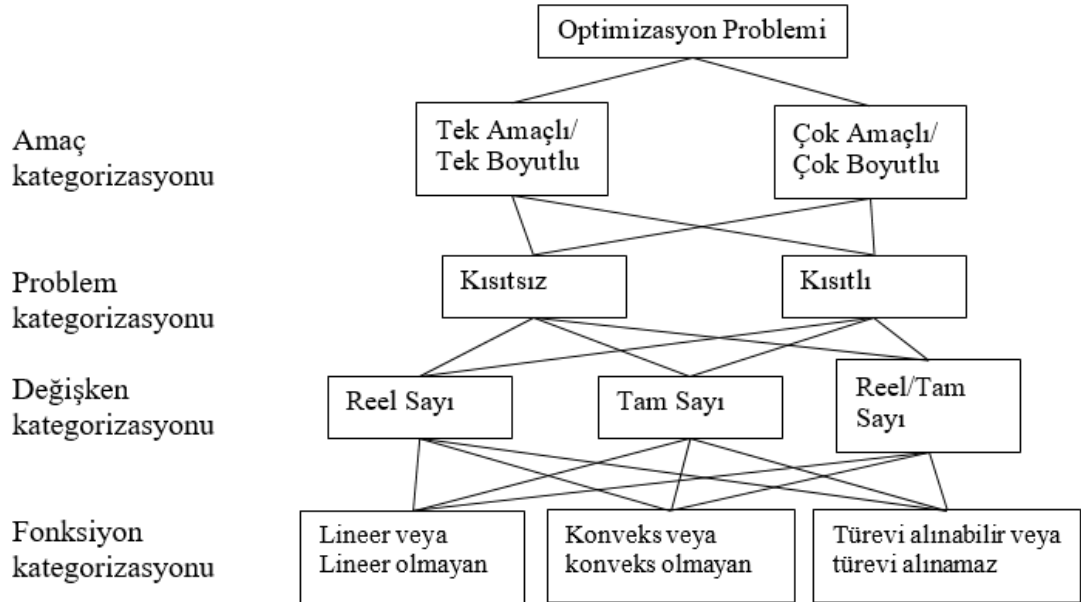
2.5.1. Optimizasyon Problemleri

On sekizinci yüzyılın ünlü İsviçreli matematikçi ve fizikçisi Leonhard Euler (1707-1783), "...Evrende bir maksimum ya da minimum kuralının ortaya çıkmadığı hiçbir şey yoktur" demiştir. Bu söz, optimizasyonun temel amacı olan minimum veya maksimum değerlerin aranmasını anlatır. Optimizasyon ile birbiriyle ilişkili bir dizi değişken için değer seçimini içeren karmaşık bir karar problemine, performansı ve kararın kalitesini ölçmek için tasarlanmış tek bir çözüme odaklanarak yaklaşılr. Bu çözüm, karar değişkenlerinin seçimini sınırlayabilecek kısıtlamalara tabi olarak maksimize veya minimize edilmesidir. Sonuç olarak, optimizasyon, felsefi olarak doğru çözümü veren bir ilke değil, bir kavramsallaştırma ve analiz aracı olarak görülmelidir (Liberti and Maculan, 2006; Sarker and Newton, 2007).

Belirli kısıtlamalara tabi olarak bir dizi değişkenin matematiksel bir fonksiyonunu minimize veya maksimize etmeye çalışan problemler, optimizasyon problemleri olarak adlandırılabilirler benzersiz bir problem sınıfı oluşturur. Optimizasyon, bilgisayar bilimleri ve makine öğrenimi, mühendislik, operasyonel araştırma, ekonomi ve ulaşım gibi birçok disiplinde kullanılmaktadır. Tüm olası alternatif çözümlerin yalnızca bir alt kümesini analiz ederek en uygun çözümleri bulan optimizasyon algoritmalarını uygulamak için, fonksiyon özelliklerine ilişkin bazı bilgilere ihtiyaç duyulur ve buna göre optimizasyon kategorilendirilir.

Şekil 2.11’de gösterildiği gibi ilk olarak amaca göre kategorizasyonu tekli veya çoklu olması ve amaç türünün maksimizasyon veya minimizasyon olarak belirlenmesi gerekmektedir. Birden fazla amaç olması durumunda, amaçlar genellikle birbiriyle çelişir. Eğer çelişmiyorsa, çoklu hedefler tek hedefli bir

probleme dönüştürülebilir. Probleme göre kategorizasyonu problemin kısıtlamalar içerip içermediğini gösterir. Genel kanı gerçek dünyada kısıtlamasız optimizasyon problemi olmadığı yönündedir, çünkü gerçek dünyada ya kısıtlama fonksiyonları ya da değişken sınırları (üst veya alt) veya her ikisi de bulunmaktadır. Kısıtsız problemlerin incelenmesi çok önemlidir çünkü birçok optimizasyon algoritması kısıtlı problemleri kısıtsız probleme dönüştürerek çözmektedir. Buna ek olarak, bazı kısıtlamasız optimizasyon teknikleri, kısıtlı problemler için çözüm prosedürleri sağlamak için genişletilebilir. Değişkenlerin kategorize edilmesi reel sayı, tam sayı veya karışık (reel/tam sayı) olabilir. Fonksiyon kategorizasyonu temel olarak fonksiyonların çözüm yaklaşımı açısından çok önemli olan matematiksel özellikleriyle ilgilidir. Amaç veya kısıtlama fonksiyonları doğrusal, doğrusal olmayan veya her ikisi birden olabilir. Konvekslik, klasik optimizasyonda önemli bir özellik olarak kabul edilir, çünkü birçok optimizasyon tekniği fonksiyonun konveks olduğu varsayımına dayalı olarak geliştirilmiştir. Son olarak fonksiyon türevlenebilir veya türevlenemez olabilir, optimizasyonda çözüm yaklaşımları iki ana gruba ayrılabilir: (i) türevli olanlar ve (ii) türevsiz olanlar. Türev tabanlı teknikler kullanılırken fonksiyonun türevlenebilirliği gereklidir. Örnek olarak, reel sayı değişkenlerine sahip tek amaçlı, kısıtlı bir problemin fonksiyon özellikleri doğrusal olmayan, dışbükey ve diferansiyel olabilir (Sarker and Newton, 2007).



Şekil 2.11 Optimizasyon problemleri kategorizasyonu (Sarker and Newton, 2007)

Verilen bu genel kategorizeye ek olarak, optimizasyon problemi tek modlu ve çok modlu özelliklerini de dikkate alır. Sadece bir tepe noktası (optimum çözüm)

olan bir fonksiyon tek modlu bir fonksiyon, birden fazla tepe noktası (yerel veya global optimum) olan bir fonksiyon çok modlu bir fonksiyon olarak kabul edilir. Fonksiyon statik ve dinamik olabilir. Çözümde karşılanması gereken kısıtlamalar sert kısıtlamalar olarak bilinirken, belirli bir ceza ile veya belirli koşullar altında ihlal edilebilen kısıtlamalar yumuşak kısıtlar olarak adlandırılır (Bazaraa et al., 1990; Hillier and Lieberman, 2005; Sarker and Newton, 2007)

Bir optimizasyon probleminin çözülmesi için aşağıdaki unsurların belirlenmesi gerekmektedir;

1. Problemin modellenmesi

- Optimizasyon değişkenlerinin tanımlanması
- Amaç fonksiyonunun formülize edilmesi
- Problem kısıtlı mı yoksa kısıtsız mı olduğunu belirlemek, kısıtların belirlenmesi

2. Problemin özelliklerinin belirlenmesi

- Optimizasyon değişkeninin boyutunu belirlemek
- Amaç doğrusal mı yoksa doğrusal değil mi? Kısıtlar (varsa) doğrusal mı yoksa doğrusal değil mi belirlemek
- Problem dışbükey mi yoksa dışbükey değil mi belirlemek

Bu iki temel unsurun sırayla uygulanması zorunlu değildir ve tamamlanması ile uygun optimizasyon yöntemi seçilir. Örneğin, doğrusal bir optimizasyon problemi en iyi şekilde doğrusal optimizasyon problemlerine uyarlanmış bir çözücü mesela simpleks yöntemin kullanılabilir. Fakat literatürdeki genel kanı her tür optimizasyon problemi için tek bir en iyi çözücü olmadığı yönündedir (Luenberger and Ye, 2008; Boyd and Parrilo, 2009; Herzog, 2022).

Optimizasyon problemlerinin matematiksel olarak genel gösterimi Eşitlik 2.20'de gösterilmektedir.

$$\text{Minimum veya maximum } f_0(x) \quad (2.20)$$

$$f_i(x) \leq b_i, \quad i=1, \dots, m$$

$x = (x_1, \dots, x_n)$: optimizasyon değişkenleri

$f_0 : R^n \rightarrow R$: amaç fonksiyon

$f_i : R^n \rightarrow R, i = 1, \dots, m$: kısıt

Bu problemin optimal çözüm sonucu x , kısıtlamaları sağlayan tüm vektörler arasında en küçük veya en büyük $f_0(x)$ değerine sahiptir olarak ifade edilir.

En yakın komşu sınıflandırmasında k değerine, destek vektör makineleri algoritmasında hangi kernel fonksiyonunun kullanılacağına, ağ modellerinde seyreltmeye, nöron sayısına, katman sayısına tasarımcı karar verdiği gibi MARS algoritmasında da tasarımcının karar vermesi gereken bazı parametreler bulunmaktadır. Bu parametrelerin başlangıçta hangi değerleri alacağı açık ve kesin değildir; çünkü bunlar veri setine, problemin yapısına göre değişiklik göstermektedir. Bu parametreler, hiper parametre (hyperparameter) olarak adlandırılır. Doğrusal olmayan problemlerin çözümünde en uygun hiper parametre değerleri bulmak için optimizasyon yöntemleri kullanılmakta olup, bu algoritmalar arasında performans ve hız bakımından farklılıklar bulunmaktadır (Bergstra et al., 2011; Feurer and Hutter, 2019; Savaş, 2019; Yang and Shami, 2020).

2.5.2. Parçacık Sürü Optimizasyonu (Particle Swarm Optimization – PSO)

PSO, 1995 yılında James Kennedy ve Russell Eberhart tarafından geliştirilen doğadan ilham alan en popüler metasezgisel optimizasyon algoritmalarından biridir. Son zamanlarda PSO, bilim ve mühendislik alanındaki çeşitli optimizasyon problemlerinin çözümünde umut verici bir algoritma olarak ortaya çıkmıştır. PSO, model setlerinin en iyi performansı göstermesi için farklı tahmin sonuçlarının bir araya getirilmesine yönelik mühendislik zorluklarına çözüm olarak geliştirilmiştir. PSO tekniğinin yakınsama hızı çok hızlıdır ve bu teknik aynı zamanda çok boyutlu problemleri de çözebilmektedir (Kennedy and Eberhart, 1995; Kennedy and Eberhart, 2007).

PSO, kuş ve balık gibi sürü halinde hareket eden hayvanların, yiyecek bulma gibi temel ihtiyaçlarını karşılamak için hareket ederken, sürüdeki diğer bireyleri etkilediği ve sürünün amacına daha kolay ulaştığı gözlemlenerek geliştirilen bir optimizasyon tekniğidir (Kennedy and Eberhart, 1995). PSO, her biri probleme potansiyel bir çözüm sunan bir parçacık sürüsünü içerir. PSO algoritmasında, her bir parçacık çözümü arayan bir unsur olarak kabul edilir ve parçacıkların bulunduğu veri kümesi sürü olarak adlandırılır. PSO'nun her iterasyonunda, parçacıklar hem kendi önceki en iyi konumlarının hafızasından hem de sürüdeki diğer parçacıklar tarafından elde edilen en iyi konumdan etkilenecek arama uzayında hareket ederler. Her bir parçacığın çözüm aradığı süre boyunca elde ettiği en iyi konum "pbest" olarak adlandırılır. Tüm sürüde, arama süresi boyunca çözüme en çok yaklaşan parçacığın o andaki durumu ise "gbest" olarak adlandırılır (Kennedy and Eberhart, 1995; Jiang et al., 2007; Kennedy and Eberhart, 2007; Inti and Tandon, 2017; Gad,

2022).

PSO, çeşitli türlerin arama alanındaki ihtiyaçlarını karşılamak için sergiledikleri sosyal ve işbirlikçi davranışlardan esinlenmiştir. Algoritma, kişisel deneyim (pbest), genel deneyim (gbest) ve parçacıkların arama uzayında bir sonraki konumlarına (değerlerine) karar vermek için mevcut hareketleri tarafından yönlendirilir. Deneyimler, iki faktör (c1 ve c2) ve [0, 1] arasında üretilen iki rastgele sayı ile hızlandırılırken, mevcut hareket (wmin, wmax) arasında değişen bir w ile çarpılır. PSO optimizasyonu Eşitlik 2.21 ve Eşitlik 2.22 kullanılarak iteratif formüller ile elde edilir. Eşitlik 5.1'de $p_{best\ i,j}(k)$ i'nci bireyinin kişisel en iyi j. parçacığın temsil ederken, $g_{best\ i,j}(k)$, k iterasyonuna kadar popülasyonun en iyi bireyinin j. bileşenini temsil eder (Kennedy and Eberhart, 1995; Kennedy and Eberhart, 2007; Gad, 2022).

$$V_{i,j}(k+1) = \omega * V_{i,j}(k) + c_1 * rand(.) * (p_{best\ i,j}(k) - X_{i,j}(k)) + c_2 * rand(.) * (g_{best\ i,j}(k) - X_{i,j}(k)) \quad (2.21)$$

$$X_{i,j}(k+1) = X_{i,j}(k) + V_{i,j}(k+1) \quad (2.22)$$

i = N büyüklüğünde ve j = D boyutundaki başlangıç sürüsünden i'nci parçacık $X_i = (X_{i,1}, X_{i,2}, \dots, X_{i,D})^T$ ile gösterilir ve burada "T" transpoze operatörünü ifade eder. Popülasyondaki her bir parçacığın hızı $V_i = (V_{i,1}, V_{i,2}, \dots, V_{i,D})^T$ olarak ifade edilir. rand(.) 0 ile 1 arasındaki rastgele sayıyı temsil eder. Bireysel ve global uç değerler sırasıyla $p_{best\ i,j}$ ve $g_{best\ i,j}$ terimleriyle ifade edilir. Öğrenme (hızlanma) faktörleri, yani c_1 ve c_2 , 2 ile 2.05 arasında değişmektedir. Bu katsayılar, parçacıkların hem kendi en iyi pozisyonlarına (bilişsel bileşen – c_1) hem de sürünün en iyi pozisyonuna (sosyal bileşen – c_2) yönelme eğilimlerini kontrol eder (Alam, 2016). Ağırlık faktörü ω , parçacıkların hızını azaltmaya ve sürüyü kontrol etmeye yardımcı olur ve Eşitlik 2.23'te ifade edilmektedir (Eberhart and Kennedy, 1995; Kennedy and Eberhart, 2007; Çatal Reis ve Bayram, 2016; Hasanoğlu, 2019).

$$\omega = \omega_{min} + \frac{(T_{max}-T)(\omega_{max}-\omega_{min})}{T_{max}} \quad (2.23)$$

Eşitlik 2.23'te T ve T_{max} mevcut ve maksimum iterasyon sayılarını göstermektedir (Çatal Reis ve Bayram, 2016; Hasanoğlu, 2019).

PSO algoritmasının adımları Şekil 2.12' de gösterilmektedir (Hasanoğlu, 2019).

Adım 1: Tahmin değerlerinin ağırlıklarının başlatılması.

Adım 2: $T = 1$ 'i ayarlayın

Adım 3: PSO algoritmasının parametre değerlerini ayarlayın ve ayrıca popülasyon boyutunu belirtin.

Adım 4: Konum (X_i) ve hız (V_i) vektörlerinden bahsederek parçacıklar üretin.

Adım 5: Üretilen her parçacığın uygunluğunu hesaplayın.

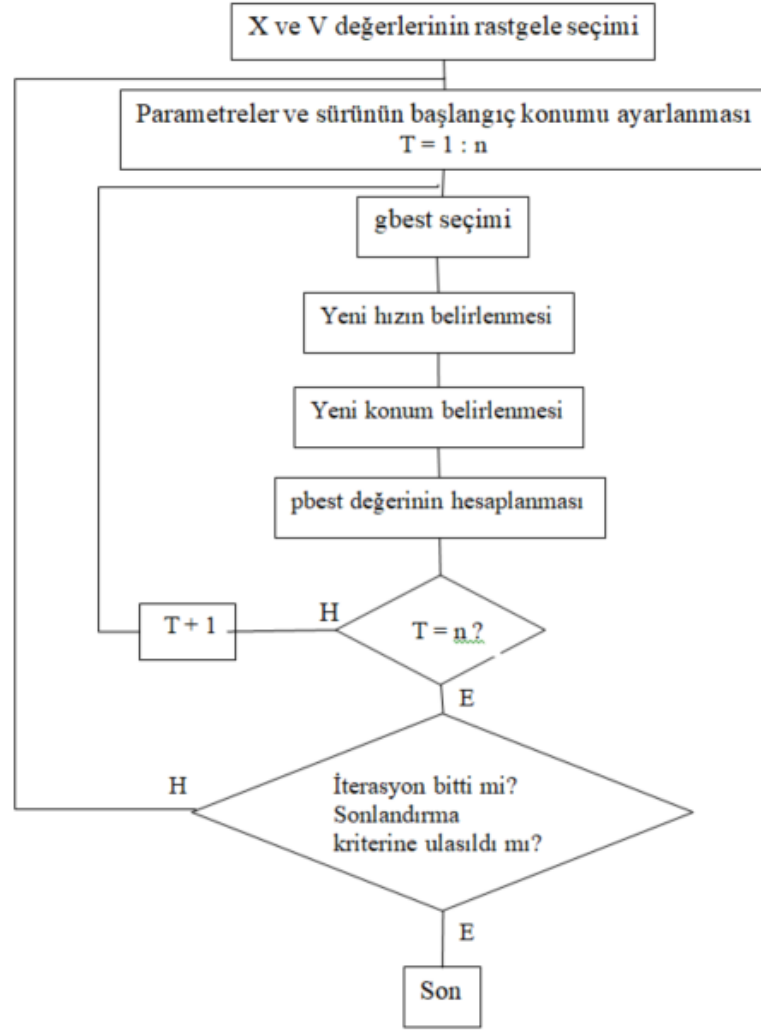
Adım 6: Uygunluk değerinin önceki uç değerden daha iyi olması durumunda bireysel uç değeri ($p_{best\ i, j}$) güncelleyin.

Adım 7: Adım 6'da bireysel uç değer güncelleniyorsa, global uç değeri ($g_{best\ i, j}$) buna göre güncelleyin. Bu arada, parçacığın hızını ve konumunu da hesaplayın ve güncelleyin.

Adım 8: $T = T_{max}$ 'a kadar iterasyon sayısını ($T = T + 1$) artırarak bu adımları tekrarlayın.

Adım 9: Optimizasyonun sonu.

Optimizasyon işlemi tamamlandıktan sonra, minimum MSE'ye sahip tahmin kombinasyonunun nihai ağırlıkları ve o zamana kadar nihai tahmin elde edilecektir.



Şekil 2.12 PSO işlem adımları (Hasanoğlu, 2019)

PSO algoritmasının avantajları (Gad, 2022);

- Eğitim aşaması hızlıdır.
- Diğer optimizasyon yöntemlerine göre daha az parametre ayarı içermesi
- Optimizasyon problemlerindeki kısıtların kolay uygulanabilmesi

PSO algoritmasının kısıtları (Gad, 2022);

- Her iterasyonda hız güncellenmesi nedeniyle daha fazla belleğe ihtiyaç duyar.
- Ayrık optimizasyon problemlerinde performansı diğer birçok optimizasyona göre düşüktür.

2.5.3. Bayes Optimizasyonu (Bayesian Optimization – BO)

Malzeme bilimi, robotik, sentetik gen tasarımı ve tahmin sistemleri gibi alanlardaki problemlere başarıyla uygulanmış olan BO, makine öğrenimi algoritmaları için parametrelerin ayarlanmasında kullanılmaya başlanmıştır (Bergstra

et al., 2011, Joy et al., 2019). BO iki ana bileşeni bulunmaktadır; birincisi bilinmeyen fonksiyonu modellemek, diğeri ise bu modele dayalı olarak uygulamanın gerçekleştirileceği bir sonraki noktayı seçmektir (Garnett et al., 2010; Gonzalez et al. 2015).

BO, bir amaç fonksiyonunun maksimum ve minimum değerlerini bulmak için Bayes teoremini kullanan bir yaklaşımdır. Bir olayın/olgunun koşullu olasılığını hesaplamak için kullanılan BO, gözlemlenen olayların diğeri olayların olasılığına bağlı olduğunu belirten Bayes olasılık teoremine dayanır. Bayes teoremi, 18. yüzyılda İngiliz matematikçi Thomas Bayes tarafından geliştirilen ve olasılık teorisinde önemli bir yer tutan bir teoremdir. Bayes teoremi, iki olay olan A ve B'nin meydana gelmesindeki koşullu olasılıklar arasındaki ilişkiyi açıklar. Genellikle $P(A|B)$ olarak yazılan koşullu olasılık, B olayının gerçekleştiği bilgisine dayanarak A olayının gerçekleşme olasılığıdır (Eşitlik 2.24) (Mockus et al., 1978; Jones, 2001; Ekici, 2009; Yu and Zhu, 2020).

$$P(A|B) = P(B|A) P(A) \quad (2.24)$$

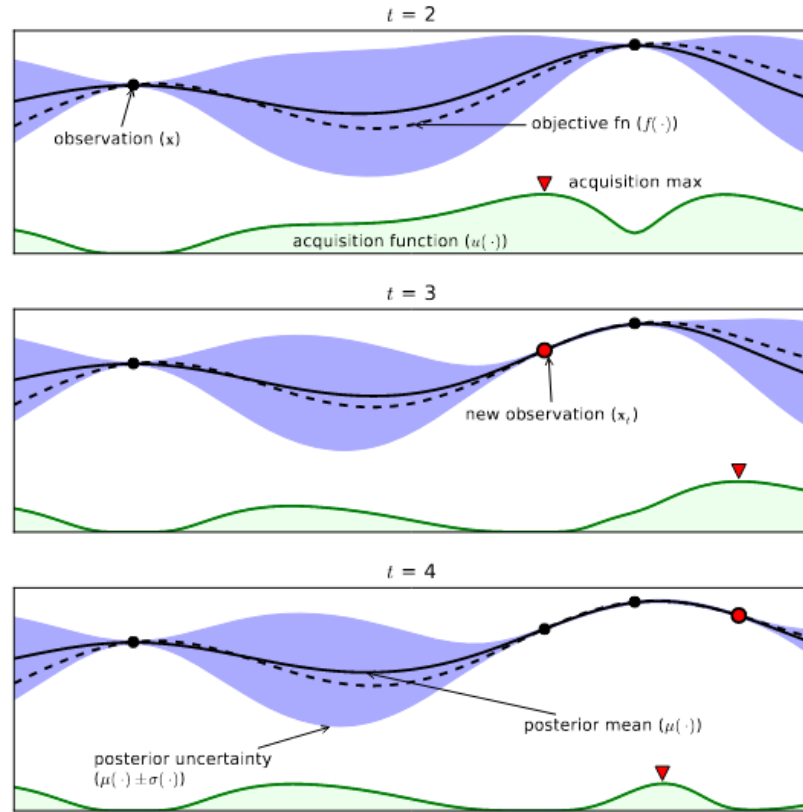
BO, Bayes teoremini kullanarak amaç fonksiyonunun sonsal dağılımını tahmin etmektedir ve bu dağılımı dikkate alarak bir sonraki parametre değerlerini belirler. BO, diğeri optimizasyon yaklaşımlarından farklı kılan, tüm girdi uzayını araştırmaması, bunun yerine bir optimum sağlama olasılığı daha yüksek olan kısıtlı bir alt kümeye odaklanmasıdır (Mockus et al., 1978; Jones, 2001).

BO algoritmasındaki iki önemli seçim bulunmaktadır; optimize edilen fonksiyon hakkındaki varsayımları ifade eden fonksiyonlar arasında öncelik seçilmesi, değerlendirilecek bir sonraki noktayı belirleyen edinim fonksiyonu seçilmesi. İlk kısım için esnekliği ve izlenebilirliği nedeniyle Gauss süreci önceliğini tercih edilmektedir. Edinim fonksiyonları ise Gauss süreci yanı sıra önceki gözlemlere de bağlı olarak değişmektedir (Mockus et al., 1978; Brochu et al., 2010).

Gauss süreci $\{x_n \in X\}_{n=1}^N$ herhangi bir sonlu N nokta kümesinin R^N üzerinde çok değişkenli bir Gauss dağılımı oluşturur. Bu noktalardan n'incisi $f(x_n)$ fonksiyon değeri olarak alınır ve bu değerler arasında bir ortak ilişki (kovaryans) belirler. Gauss sürecinin temel bileşeninden biri olan kovaryans fonksiyonu iki nokta arasındaki benzerliği ölçer ve diğeri bir bileşen ise fonksiyonun şeklinin nasıl olduğunu belirleyen ortalama fonksiyondur. Bu öncül işlemler sonucunda edinim fonksiyonu oluşturulur. Edinim fonksiyonu, amaç fonksiyonunun bir sonraki aşamada nerede örnekleneceğini belirler. Beklenen İyileştirme (Expected Improvement), Bilgi

Gradyanı (Knowledge Gradient), Entropi Arama ve Tahmini Entropi Arama (Entropy Search and Predictive Entropy Search) yaygın olarak kullanılan edinim fonksiyonlarıdır. Edinim fonksiyonuyla amaç fonksiyonuna yaklaşan bir vekil fonksiyonu oluşturulur. Yeni veri noktalarıyla yani iterasyonla, vekil fonksiyon zaman içinde daha doğru hale gelir ve optimizasyon sürecini optimum çözüme doğru yönlendirir (Mockus et al., 1978; Brochu et al., 2010; Frazier, 2018).

Optimizasyon iki nokta ile başlar. Her iterasyonda, hedef fonksiyonundan bir sonraki örnekleme nerede yapılacağını belirlemek için edinim fonksiyonunun maksimum olduğu nokta seçilir. Daha sonra, edinim fonksiyonu argmax değerinde örneklenir, Gauss süreci güncellenir ve süreç tekrarlanır. Şekil 2.13 1 boyutlu tasarım probleminde BO algoritması kullanımına bir örnektir. Örnekte amaç fonksiyonunun örneklenmiş değerlerinin 4 iterasyonu üzerinden, amaç fonksiyonunun Gauss yaklaşımını göstermektedir (Mockus et al., 1978; Jones, 2001, Sasena, 2002; Brochu et al., 2010).



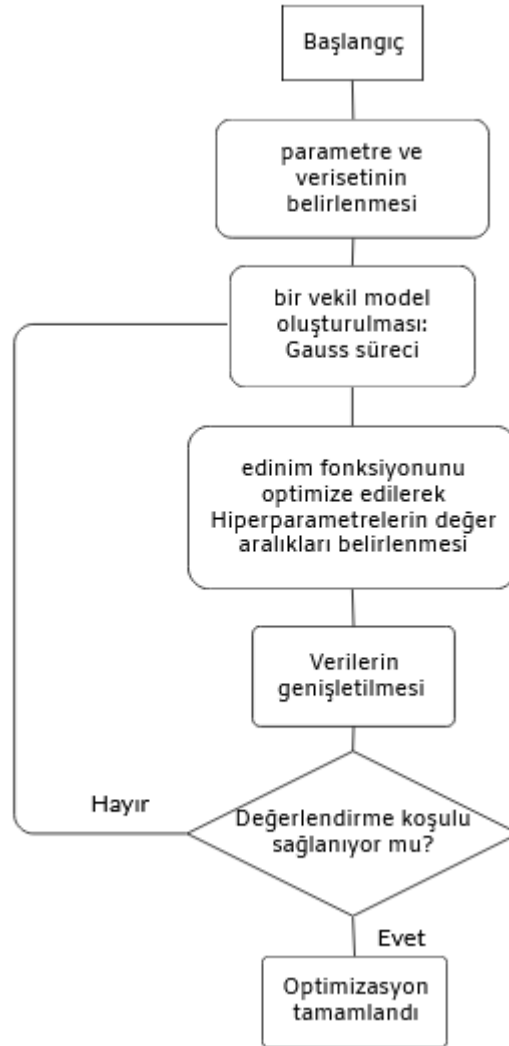
Şekil 2.13 1 boyutlu tasarım probleminde BO algoritmasının kullanımı (Brochu et al., 2010)

BO algoritması gerçekleştirilirken yapılması gereken iki önemli seçim vardır. İlk olarak, optimize edilen fonksiyonla ilgili varsayımları ifade edecek fonksiyonlar üzerinde bir öncelik yani Gauss süreci önceliğini seçilir. İkinci olarak,

değerlendirilecek bir sonraki noktanın belirlenmesine olanak tanıyan, modelin son halinden bir fayda fonksiyonu oluşturmak için kullanılan bir edinim fonksiyonu seçilmesidir (Brouch et al., 2010; Hoffman et al., 2011; Gonzalez, 2017).

BO algoritması işlem adımları Şekil 2.14'te gösterilmektedir (Brouch et al., 2010; Hoffman et al., 2011);

- Parametre ve iterasyon sayısı belirlenmesi $t = 1, 2, \dots$
- Gauss süreci ile edinim fonksiyonunu optimize ederek x_t 'nin bulunması: $x_t = \arg\min_x u(x|D_{1:t-1})$.
- Amaç fonksiyonunu örneklenmesi: $y_t = f(x_t) + \varepsilon_t$.
- Verilerin genişletilmesi $D_{1:t} = \{D_{1:t-1}, (x_t, y_t)\}$ ve Gauss sürecinin güncellenmesi.
- Optimizasyon sonu



Şekil 2.14 BO işlem adımları

BO algoritması avantajları (Brouch et al., 2010; Frazier, 2018);

- Diğer birçok algoritmanın temelini oluşturma
- Ön bilgilerin ve öznel bilginin tahminlere dahil etme
- Sınırlı veri kullanılabilirliği olan durumları etkin bir şekilde analiz etme, verilerin az olduğu durumlarda küçük veri setlerine dayalı çıkarımlar sağlama
- Klasik yöntemlere kıyasla istatistiksel analiz için daha esnek ve sezgisel bir çerçeve sunma
- Olasılık dağılımlarındaki değişikliklerin izlenmesini sağlamak ve olasılıksal modeller üzerinden istatistiksel ve nedensel çıkarımları kolaylaştırmak

BO algoritması kısıtları (Brouch et al., 2010; Frazier, 2018);

- Özellikle büyük ölçekli problemler veya yüksek boyutlu arama uzayları için hesaplama açısından yoğun olabilir
- Problemden paralellik durumu olduğunda uygulaması zordur yani BO sıralı bir süreçtir, amaç fonksiyonunda paralel hesaplama gerektiren durumlarda dezavantajlıdır.
- Kavramsal olarak karmaşıktır.

2.6. Sınıflandırma Doğruluğunun Değerlendirilmesi

Uydu görüntülerinin sınıflandırılmasında başlıca işlem adımları; sınıflandırmada kullanılacak görüntülerin ve bantların belirlenmesi, görüntü ön işleme, sınıfların belirlenmesi, eğitim örneklemelerinin seçilmesi, özellik çıkarımı, uygun sınıflandırma yaklaşımlarının seçilmesi, sınıflandırma sonrası işlem ve doğruluk değerlendirmesidir (Stehman, 1999; Grullan et al., 2009; Stehman and Wickham, 2011; Wickham et al., 2017). Sınıflandırma süreci dikkate alındığında çalışma alanının özellikleri, analistin deneyimleri, uzaktan algılama verilerinin seçimi, sınıflandırma tekniği gibi faktörlerin sınıflandırma doğruluğunu etkilediği anlaşılmaktadır (Stehman, 1999; Stehman and Czaplewski, 2003; Lu and Weng, 2007).

Bir haritanın doğruluğunun bilimsel olarak güvenilir bir şekilde değerlendirilmesi kritik önem taşımaktadır. Tematik harita doğruluğunun değerlendirilmesi, çıktıların güvenilirliği ve kalitesinin sayısal olarak analiz edilmesine olanak tanıyan önemli bir aşamadır (Congalton and Green, 2009). Doğruluk değerlendirmesinde sınıflandırılmış haritalar ile referans veriler arasındaki

istatistiksel belirlenmektedir. Doğruluk değerlendirme süreci üç temel adımdan oluşur: i) referans veriyi seçmek için kullanılan örnekleme stratejisi ii) örneklem verisini kullanarak arazi örtüsü/kullanımı haritaları oluşturan hedef tasarım ve iii) hata ve belirsizliklerin tahmin edilmesi (Stehman, 1999; Stehman and Czaplewski, 2003).

Görüntü sınıflandırma analizlerinde genellikle hata matrisi (confusion matrix), ve hata matrisinden hesaplanan toplam doğruluk, kappa istatistiği ve F1 skor gibi ölçütler kullanılarak yapılır. Hata matrisi, sınıflandırma sonuçları ile referans veriler arasındaki uyumu gösteren bir tablodur ve her bir sınıf için doğru ve yanlış sınıflandırılan piksel sayısını içerir. Toplam doğruluk, doğru sınıflandırılan örneklerin toplam örnek sayısına oranıdır ve sınıflandırmanın genel başarısını ifade eder. Kappa istatistiği ise, sınıflandırmanın rastgele tahminlere göre ne kadar iyi olduğunu gösteren bir ölçüttür ve 0 ile 1 arasında bir değer alır. Bu iki ölçüt birlikte kullanılarak, görüntü sınıflandırma analizlerinde elde edilen sonuçların doğruluğu ve güvenilirliği kapsamlı bir şekilde değerlendirilebilir (Foody, 2009). Hata matrisinden hesaplanan kesinlik (precision) ve duyarlılık (recall) ölçütleri kullanılarak hesaplanan F1 skoru doğruluk değerlendirme ölçütü ise makine öğrenmesi ve istatistiksel modelleme alanlarında yaygın olarak kullanılmaktadır (Lipton et al., 2014).

2.6.1. Toplam Doğruluk

Bir hata matrisi oluşturmak, sınıflandırma doğruluğunun değerlendirmesinde en yaygın kullanılan yöntemlerden biridir. Tablo 2.2’de gösterilen hata matrisi, sütun ve satır sayısı sınıf sayısına eşit olan bir diziden oluşur ve referans örneklerini sınıf bazında tahmin edilen sınıflandırma verileriyle karşılaştırmak için kullanılır. Hata matrisi bir $n \times n$ dizisi içerir, burada "n" sınıf sayısını temsil eder. Hata matrisinde diyagonal hücreler doğru sınıflandırılmış piksel sayısını verirken, diyagonal olmayan hücreler sınıflandırma hatalarını, yani referans görüntü ile sınıflandırılmış görüntü arasındaki uyumsuzlukları gösterir. Hata matrisi kullanılarak üretici ve kullanıcı doğruluğu ölçütleri ayrıca ihmal hatası ve dahil etme hatası elde edilmektedir (Congalton and Green, 1999; Foody, 2009).

Tablo 2.2 Hata matrisi

		Referans Veri		Toplam	Kullanıcı Doğruluğu / Kesinlik (Precision)
		A	B		
Sınıflandırılmış Veri	A	x_{AA}	x_{AB}	x_{i+}	x_{AA} / x_{i+}
	B	x_{BA}	x_{BB}	x_{i+}	x_{BB} / x_{i+}
Toplam		x_{+i}	x_{+i}	N	
Üretici Doğruluğu / Duyarlılık (Recall)		x_{AA} / x_{+i}	x_{BB} / x_{+i}		

Tablo 2.2’de eşitliği verilen üretici doğruluğu, bir spektral sınıf için referans kabul edilen piksellerden kaç tanesinin doğru şekilde sınıflandırıldığını gösterir. Her sınıf için doğru sınıflandırılmış piksellerin sayısının, o sınıf için kullanılan referans piksellerinin toplam sayısına bölünmesiyle hesaplanan üretici doğruluğu 1’den farkı alınarak ihmal hatası elde edilir. Tablo 2.2’de eşitliği verilen kullanıcı doğruluğu sınıflandırma haritasındaki bir pikselin sınıfının yeryüzünde de o pikseli temsil edip etmediğini ölçmektedir. Kullanıcı doğruluğu 1’den farkı alınarak dahil etme hatasını elde edilir (Lillesand and Kiefer 1994; Assal et al., 2015).

Hata matrisi kullanılarak hesaplanan doğruluk ölçütlerinden toplam doğruluk, tüm referans alanlarının ne kadarının doğru sınıflandırıldığı bilgisini verir. Hata matrisindeki doğru sınıflandırılmış piksel sayısı toplamının referans piksellerinin toplamına bölünmesiyle elde edilen toplam doğruluk değeri yüzde olarak ifade edilir ve Tablo 2.2’deki hata matrisi örnek alınarak formüleştirildiğinde Eşitlik 2.25’te gösterildiği gibi hesaplanmaktadır (Lillesand and Kiefer 1994; Assal et al., 2015).

$$\text{toplam doğruluk} = \frac{x_{AA} + x_{BB}}{\text{toplam referans veri sayısı}} \quad (2.25)$$

2.6.2. Kappa İstatistiği

Cohen’in Kappa istatistiği, tesadüfi bir sınıflandırıcı ile referans veri arasındaki şans uyumu ile otomatik bir sınıflandırıcı ile referans veri arasındaki gerçek uyum arasındaki farkı ifade etmektedir. Bir sınıflandırmanın doğruluğunu değerlendirmek için kullanılan istatistiksel bir değer olan Kappa, sınıflandırmanın sadece rastgele değer atamaya kıyasla ne kadar iyi performans gösterdiğini değerlendirir. Yani, sınıflandırmanın rastgele olandan daha iyi sonuç verip vermediği hakkında bilgi verir. Eşitlik 2.26’da gösterildiği gibi hesaplanan kappa değeri 0 ile 1 arasında değişebilir; 0 değeri, sınıflandırmanın rastgele bir sınıflandırmadan daha iyi

olmadığını gösterir. 1'e yakın bir kappa değeri ise sınıflandırmanın rastgele sınıflandırmadan önemli ölçüde daha iyi olduğunu gösterir. Kappa katsayısı, şans uyum olasılığını hesaba katan bir ölçüttür (Lillesand and Kiefer 1994; Assal et al., 2015).

$$k = \frac{\text{Gözlemlenen Doğruluk} - \text{Şans Uyumu}}{1 - \text{Şans Uyumu}} \quad (2.26)$$

Eşitlik 2.27 kappa istatistiği hesabının formülizasyonunu göstermektedir.

$$k = \frac{N * \sum_{i=1}^r x_{ii} - \sum_{i=1}^r x_{i+} * x_{+i}}{N^2 - \sum_{i=1}^r x_{i+} * x_{+i}} \quad (2.27)$$

Burada

r = hata matrisindeki satır sayısı

x_{ii} = i satır ve i sütunundaki veri sayısı (hata matrisinin ana çaprazındaki değerler)

x_{i+} = i satırındaki gözlemlerin toplamı

x_{+i} = i sütunundaki gözlemlerin toplamı

N = toplam referans veri sayısını ifade etmektedir.

2.6.3. F1 Skor

F1 skor ölçütünün hesaplanması için duyarlılık ve kesinlik ölçütlerinin hesaplanması gerekmektedir. Tablo 2.2'de gösterildiği gibi üretici doğruluğu, duyarlılık; kullanıcı doğruluğu ise kesinlik ile aynı şekilde hesaplanmaktadır. F skor ise Eşitlik 2.7'de gösterildiği gibi hesaplanır. Duyarlılık, kesinlik ve F1 skoru ölçütleri her bir sınıf için ayrı ayrı hesaplanır. Bir sınıflandırma sonucunda hem doğru sonuçların bulunması hem de yanlış sonuçların azaltılması önemli ise F1 skora bakılmaktadır. Bu ölçütlerinin temel avantajı, sınıflandırıcının performansını tek tek sınıflar bazında ele alarak sınıf dengesizliği sorunlarıyla ilgili olarak doğruluk gibi çok sınıflı ölçütlerin eksikliklerini hesaba katma ve aynı zamanda tek tek sınıf performanslarını göstermesidir. Bu ölçütlerin kısıtı ise tek bir sınıfı kapsamaması, tüm sınıflandırma performansını ifade etmemesidir. F ölçütü, duyarlılık ve kesinlik ölçütlerinin ağırlıklı harmonik ortalamasıdır. Herhangi bir $\alpha \in \mathbb{R}$, $\alpha > 0$ için, F ölçüsünün genel bir formülasyonu Eşitlik 2.28'de gösterilmektedir (Gamon et al., 2005; Sokolova and Lapalme, 2009; Japkowicz and Shah, 2011).

$$F \propto = \frac{(1+\alpha) * (\text{Kesinlik} * \text{Duyarlılık})}{(\alpha * \text{Kesinlik}) + \text{Duyarlılık}} \quad (2.28)$$

Eşitlik 2.28'deki fomülde $\alpha=1$ alınarak ölçüt, F1 skor olarak adlandırılır ve Eşitlik 2.29 elde edilir.

$$F1 = \frac{2*(Kesinlik*Duyarluluk)}{Kesinlik+Duyarluluk} \quad (2.29)$$

F1 ölçütünü, çoklu sınıflandırmalarda hesaplamak için 3 yöntem bulunmaktadır; makro, mikro ve ağırlıklı ortalama. Makro ortalama tüm sınıfların eşit oranda etkisini sonuç değere yansıtmaktadır. Mikro ortalama daha büyük veri sayısına sahip sınıfların etkisinin daha fazla olmasını sağlar. Ağırlıklı ortalama ise sınıf ağırlıklarının farklı olarak ele alınarak hesaplanır. Bu çalışmada sınıfların etkisi eşit olması nedeniyle makro ortalama kullanılarak sınıflandırma sonucu F1 skoru hesaplanmıştır. (Sokolova and Lapalme, 2009; Japkowicz and Shah, 2011; Lipton et al., 2014).

3. METODOLOJİ

Uydu görüntülerinin sınıflandırma performansını arttırmak için MARS algoritmasına dayalı MARS-RF, MARS-PSO ve MARS-BO hibrit modelleri geliştirilmiştir. MARS ve RF tekil algoritmaları ve MARS-RF, MARS-PSO ve MARS-BO hibrit modellerin oluşturulması ve uygulanması MATLAB ve RStudio yazılımları ortamında gerçekleştirilmiştir. Modellerin uygulanması ve testi için iki farklı çalışma alanında (Samsun-Bafra ve Samsun-Atakum) arazi örtüsü/kullanımı sınıfları tespit edilmiş ve iki farklı uydu görüntüsü (Landsat-8 OLI ve Sentinel-2 MSI) ve çalışmada kullanılacak bantlar belirlenmiştir.

3.1. MARS Algoritmasına Dayalı Hibrit Sınıflandırma Modellerinin Geliştirilmesi

Hibrit bir model oluşturmak için öncelikle problem ve veri seti incelenir. Problemi etkili bir şekilde ele almak için birleştirilebilecek farklı modellerin veya algoritmaların güçlü ve zayıf yönleri karşılaştırılır. Problem için en uygun olan algoritma seçilir. Bu algoritmaların güçlü ve zayıf yönleri, birbirlerini nasıl tamamlayabilecekleri göz önünde bulundurularak entegre edilir (Zhou et al., 2002; Liu and Ge, 2018; Ye and Ding, 2018).

Günümüzde, optimize edilmiş parametrelerin varsayılan ayarlara göre daha iyi olduğu yaygın olarak kabul edilmektedir. Literatür, hibrit sistemlerin ancak tekil sınıflandırıcılar doğru ve çeşitliyse, yani düşük hata oranları sergiliyorsa ve farklı hatalar yapıyorsa veya hataları bağımsızsa etkili olduğunu açıkça göstermiştir (Thornton et al., 2013; Mantovani et al., 2016; Sanders and Giraud-Carrier, 2017).

Bu çalışmada hibrit sınıflandırmada kullanılacak tekniklerin seçiminde bu hususlara dikkat edilmiştir. Bu bağlamda öncelikle MARS algoritması avantaj ve kısıtları incelenmiş, kısıtları elimine edebilecek veya en aza indirgeyebilecek teknikler araştırılmıştır. Araştırılan tekniklerin hibrit kullanım esnekliğine sahip olup olmadığı incelendiğinde RF algoritması hibrit kullanımda etkinliği ve ayrıca görüntü sınıflandırma çalışmalarındaki performansı ile öne çıkmıştır (Zhou et al., 2002; Liu and Ge, 2018; Ye and Ding, 2018). Bu nedenle çalışma kapsamındaki ilk hibrit model olarak MARS-RF oluşturulmuştur.

Ayrıca MARS algoritması incelendiğinde en uygun temel fonksiyon sayısı ve ceza parametresi belirleme kısıtının performansa etkisini optimizasyon yöntemleri ile elimine etme veya en aza indirgemenin yapılabileceği öngörülmüştür. Optimizasyon,

bir dizi işlem uygulanıp, amaç fonksiyonunu maksimum veya minimum yapan değerlerin elde edilmesidir. Günümüzde, optimize edilmiş hiper parametrelerin varsayılan ayarlara göre daha iyi olduğu yaygın olarak kabul edilmektedir (Thornton et al., 2013; Mantovani et al., 2016; Sanders and Giraud-Carrier, 2017). Bu çalışma kapsamında PSO ve BO, MARS algoritmasının hiper parametrelerinin optimize edilmesinde kullanılarak sırasıyla MARS-PSO ve MARS-BO hibrit modelleri geliştirilmiştir.

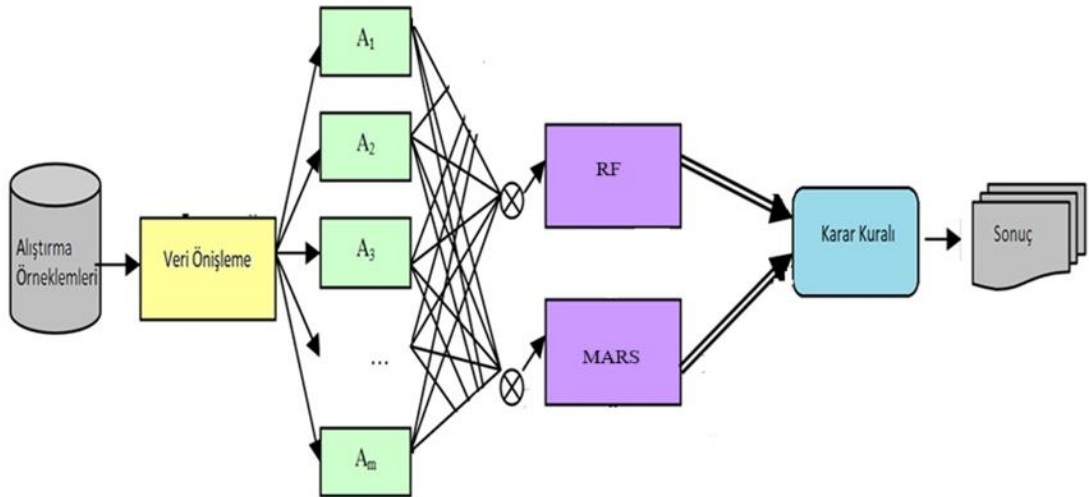
3.1.1. MARS-RF Hibrit Sınıflandırma Modeli

MARS ile RF algoritmaları uydu görüntüsü sınıflandırma analizlerinde ayrı ayrı uygulandığında yüksek doğruluklu sonuçlar vermektedirler. Ayrıca MARS ile RF algoritmalarının avantaj ve kısıtları incelendiğinde birbirlerinin kısıtlarını elimine edebileceği öngörülmüştür. MARS algoritmasının kısıtı olan aşırı uyum problemi, RF algoritmasında bulunmamaktadır. Ayrıca RF algoritmasının kısıtı olan uygun değişkenlerin seçimi, MARS algoritmasının değişkenler arasındaki etkileşimi hesaplayarak uygun değişkenleri seçmesiyle giderilmektedir. Bu avantaj ve kısıtlar göz önüne alınarak MARS ve RF sınıflandırma sonuçları birleştirildiğinde kısıtların etkisini en aza indirme veya elimine etmesi amacıyla hibrit model geliştirilmiştir.

MARS-RF hibrit modeli, hibrit sınıflandırma tekniklerinden paralel kombinasyon ile oluşturulmuş (Şekil 3.1), modelin performansı toplam doğruluk, kappa istatistiği ve F1 skor ile değerlendirilmiştir. Şekil 3.1’de gösterilen işlem adımlarında alıştırma örneklemeleri ENVI yazılımı kullanılarak seçilmiş ve RStudio yazılımına aktarmak amacıyla düzenlenmiştir. RStudio yazılımında MARS algoritması “earth” paket programı kullanılarak uygulanmıştır (çapraz geçerlilik, aşırı uyum cezası, modelin derecesi, maksimum fonksiyon sayısı gibi parametreler değiştirilebilmektedir). MARS algoritmasında genelleştirilmiş çapraz geçerlilik değeri en küçük olan model en ideal model olarak seçilmektedir. Bu doğrultuda RStudio programı içerisinde en küçük değere sahip GCV hesaplatılması için döngü oluşturulmuş, MARS algoritması bu döngüye bağlı çalıştırılmıştır. MARS algoritması ile görüntü sınıflandırma işlemi gerçekleştirmek amacıyla öncelikle RStudio yazılımı kullanılarak temel fonksiyonlar elde edilmiştir. Bu temel fonksiyonlar kullanılarak tüm görüntü sınıflandırılmıştır. Yine aynı alıştırma örneklemeleri kullanılarak paralel olarak RF sınıflandırıcısı RStudio yazılımında “randomForest” kodlarıyla çalıştırılmıştır. Bu kodlar içerisinde oluşturulacak ağaç sayısı ve her dallanmada aday olarak rastgele örneklenen değişken sayısını kullanıcı

belirleyebilmektedir.

RStudio yazılımında ASCII formatında kaydedilen sınıflandırma sonuç görüntüleri, hibrit model oluşturmak amacıyla MATLAB yazılımında açılmıştır. MATLAB yazılımında iki sınıflandırılmış görüntü incelenerek her iki yöntemde de aynı sınıfa ve farklı sınıflara atanan pikseller kodlama yapılarak belirlenmiştir. Paralel kombinasyon yöntemini uygulamak için karar kuralı olarak her iki algoritmada farklı sınıfa atanan pikseller için F1 skor doğruluğu yüksek olan modelin sonucunun seçilmesi olarak belirlenmiştir. Bu doğrultuda kodlamada MARS ve RF algoritmalarının her sınıf için F1 skor doğruluk değerleri kullanılarak “if” döngüsü kurulmuştur. Böylece her pikselin sınıf seçimi yapılmıştır. Fakat farklı sınıflara atanan piksellerin her iki sınıflandırıcıda da F1 skor doğruluğunun eşit olması durumu göz önüne alınarak kodlamaya kappa değeri büyük olan algoritmanın sonucu seçilmesi yönünde ekleme yapılmıştır.



Şekil 3.1 MARS-RF işlem adımları

3.1.2. MARS-PSO Hibrit Sınıflandırma Modeli

MARS-PSO hibrit modeli, en iyi MARS modeli için parametreleri optimize etmek üzere PSO tekniği kullanılarak geliştirilmiştir. PSO tekniği sadece MARS modelinin parametrelerini optimize etmekle kalmaz, aynı zamanda modelin doğrusal olmayan ve değişkenler arasındaki karmaşık etkileşimleri yakalama yeteneğini geliştirir (Bui et al., 2017; Nguyen et al., 2022; Garai et al., 2024).

PSO, tahmin modelinin doğruluğunu arttırmak için her iterasyonda parçacıkların konumlarını yinelemeli olarak ayarlayarak optimum çözüme doğru gelişir. Bu parçacıklar MARS algoritmasındaki maksimum temel fonksiyon sayısı, ceza parametresi hiper parametreleridir ve en uygun değerlerinin bulunması

amaçlanmaktadır. MARS-PSO hibrit sınıflandırma modelinin teorik temelleri literatürdeki benzer çalışmalarla desteklenmektedir (Benemaran and Esmaeili-Falak, 2020; Nguyen et al., 2022; Hanteh et al., 2023; Garai et al., 2024). Şekil 3.2’de hiper parametrelerin optimize edilme süreci akış şeması olarak gösterilmiştir.



Şekil 3.2 MARS-PSO işlem adımları

MARS modelinin tahmin kabiliyeti maksimum temel fonksiyon sayısı ve ceza parametresinden etkilenmektedir. Bu çalışmada, MARS'ın sınıflandırma sonucunu optimize etmek amacıyla MATLAB yazılımı kullanılarak temel fonksiyon sayısı ve ceza parametresinin uygun değerlerini aramak için PSO optimizasyonu kullanılmıştır.

Maksimum iterasyon 100, maksimum parçacık sayısı 50, temel fonksiyon sayısı ve ceza parametresi başlangıç değeri alt ve sınır aralığında Eşitlik 3.1’e göre rastgele belirlenir (Bui et al., 2019).

$$X_{i,0} = LB + R_{(0,1)} * (UB - LB) \quad (3.1)$$

$R(0,1)$, 0 ile 1 arasında eşit dağılımlı rastgele bir sayıyı ifade etmektedir. LB ve UB ise parametreler için alt ve üst sınırların iki vektörüdür. Temel fonksiyon sayısı

ve ceza parametresinin alt sınırları sırasıyla 3 ve 2; üst sınırları ise sırasıyla 50 ve 5 olarak belirlenmiştir.

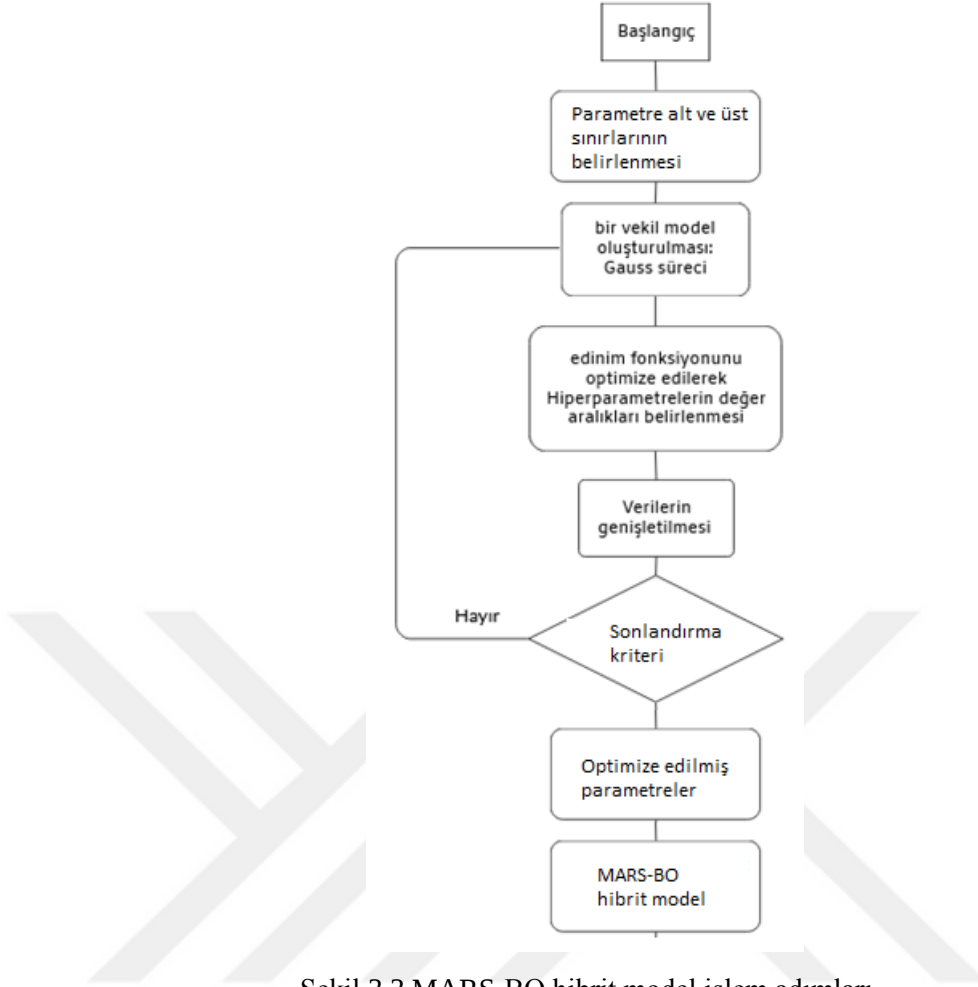
Optimizasyon süreci boyunca, PSO popülasyonu MARS'ın parametreleri olan çok sayıda değişkeni keşfetmiştir. Her iterasyonda, her bir parametre seti kullanılarak GCV hesaplanmıştır. GCV fonksiyonunun değerini minimum yapan parametreler seçilerek MARS algoritması gerçekleştirilmiştir.

3.1.3. MARS-BO Hibrit Sınıflandırma Modeli

MARS-BO hibrit modeli, en iyi MARS modeli için parametreleri optimize etmek üzere BO tekniği kullanılarak geliştirilmiştir. MARS algoritmasındaki maksimum temel fonksiyon sayısı, ceza parametresi hiper parametreleridir ve en uygun değerlerinin bulunması hedeflenmektedir. MARS-BO modelinde BO kullanılarak amaç fonksiyonunun yani GCV fonksiyonunun minimumunu bulmak amaçlanmaktadır.

BO, ilk olarak birkaç noktanın değerlendirilmesiyle başlar ve ilk vekil model oluşturulur. Bu model ve edinim fonksiyonuna dayanarak, vekil modeli ve optimum çözüm arayışını iyileştirmek için sonraki nokta seçilir. Bu iteratif süreç, maksimum değerlendirme sayısına ulaşılan veya tatmin edici bir çözüme yakınsama gibi bir durdurma kriteri karşılanıncaya kadar devam eder.

MATLAB yazılımı kullanılarak gerçekleştirilen MARS-BO hibrit modelin işlem adımları Şekil 3.3'te gösterilmektedir. MARS-BO hibrit modelinde, BO'da hiper parametrelerin alt ve üst sınırları kullanıcı tarafından belirlenir. Bu çalışmada temel fonksiyon sayısı ve ceza parametresinin alt sınırları sırasıyla 3 ve 2 olarak, üst sınırları sırasıyla 50 ve 5 olarak belirlenmiştir. BO'da parametrelerin başlangıç değerleri üst ve alt sınırlarına göre rastgele seçilmiştir. Edinim fonksiyonu olarak ise beklenen iyileştirme seçilmiştir ve edinim fonksiyonuyla amaç fonksiyonuna yaklaşan bir vekil fonksiyonu oluşturulmuştur. İterasyonla, vekil fonksiyon zaman içinde daha doğru hale gelir ve optimizasyon sürecini optimum çözüme doğru yönlendirir.

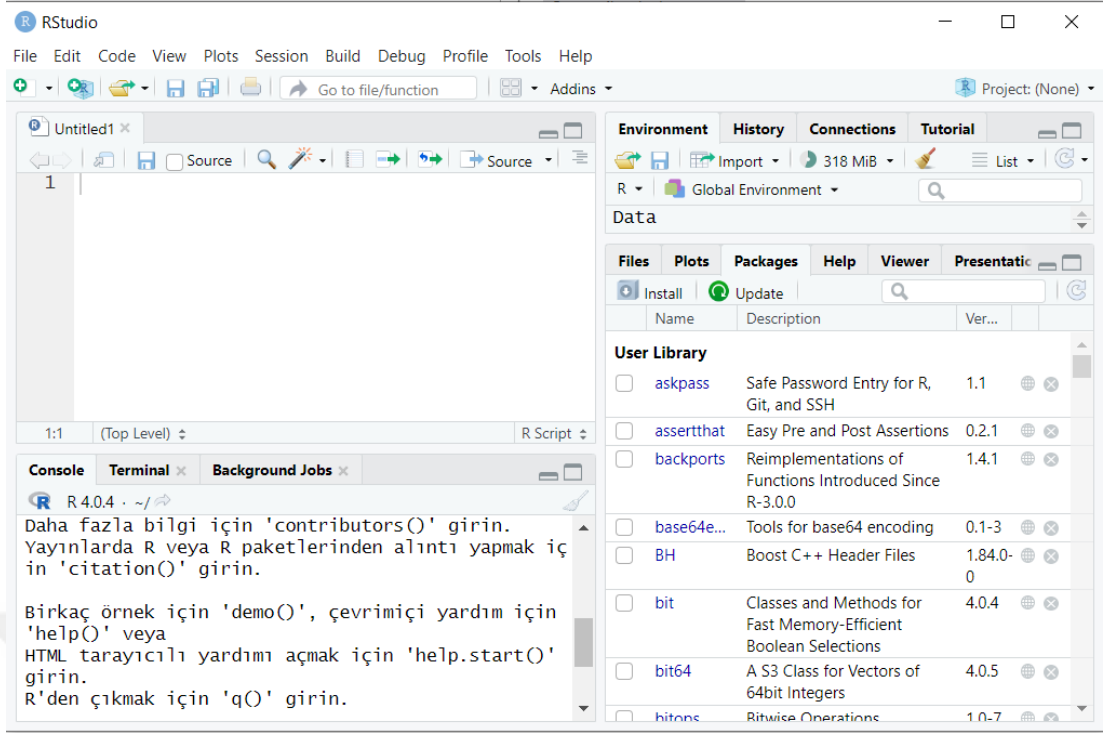


Şekil 3.3 MARS-BO hibrit model işlem adımları

3.2. Kodlama

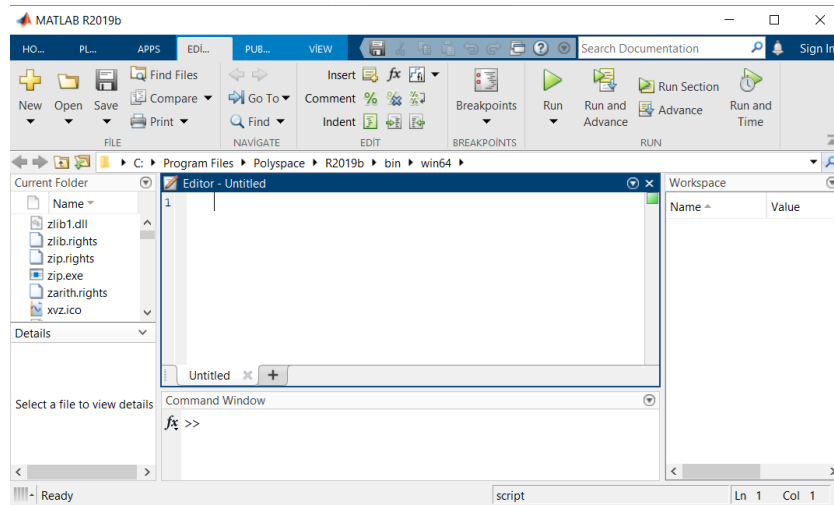
Bu çalışma kapsamında yazılan kodlar MATLAB ve RStudio yazılımı üzerinden gerçekleştirilmiştir. Her iki yazılımda da hem hazır komutlar kullanılıp hem de hazır komutlarda değişiklik ve ilave kodlama yapılmıştır.

R, istatistiksel hesaplama ve grafikler için bir programlama dili ve yazılım ortamıdır. Yeni Zelanda Auckland Üniversitesinden Ross Ihaka ve Robert Gentleman tarafından geliştirilmiştir. R programının farklı arayüzleri bulunmakta olup, bu çalışmada RStudio arayüzü (Şekil 3.4) kullanılmıştır. RStudio, kullanıcıların veri analizi, istatistiksel modelleme, veri görselleştirme ve raporlama gibi çeşitli işlemleri gerçekleştirmelerine olanak tanır ve geniş paket kütüphanesine sahiptir. Ayrıca RStudio, açık kaynaklı bir yazılımdır ve bu özelliği, kullanıcıların yazılımın işleyişini tam olarak anlamalarını ve gerektiğinde değiştirmelerini mümkün kılar (Anonim, 2021). RStudio yazılımı, RF ve MARS algoritmalarının kodlama yönünden incelenmesi ve kodlamayı değiştirebilme esnekliğine sahip olması açısından avantaj sağlamıştır.



Şekil 3.4 RStudio arayüzü

MATLAB (Şekil 3.5), matematiksel hesaplamalar, veri analizi, modelleme, simülasyon ve grafiksel gösterimler için kullanılan yüksek seviyeli bir programlama dilidir. MathWorks tarafından geliştirilmiş olan MATLAB, mühendislik, bilim, ekonomi ve daha birçok alanda geniş bir kullanım alanına sahiptir. MATLAB'ın kullanıcı dostu arayüzü, kapsamlı dökümantasyonu, araştırmacıların kod yazmayı ve geliştirmeyi hızlı ve verimli bir şekilde öğrenmelerine olanak sağlar. Ayrıca MATLAB geniş bir fonksiyon ve kütüphaneye sahiptir (MATLAB Help Center, 2024).



Şekil 3.5 MATLAB yazılımı

RStudio yazılımında uydu görüntülerinin açılması ve analize hazırlanması amacıyla “raster” paket kodlaması her iki algoritma için kullanılmıştır. RF algoritması için RStudio yazılımında “randomForest” ve “e1071” paket kodlamaları üzerinde değişiklikler yapılarak toplam 71 satır kodlama ile sınıflandırma işlemi gerçekleştirilmiştir. Çalışmada öncelikle “raster”, “e1071” ve “randomForest” paketleri kütüphaneden yüklenmiş ve alıştırma örnekleri açılmıştır. Ardından RF algoritması parametreleri ayarlanarak sınıflandırma gerçekleştirilmiş ve sonuç görüntü, harita oluşturmak amacıyla ASCII formatında kaydedilmiştir.

```
library(raster)
library(e1071)
library(randomForest)

samples15 <- readOGR(dirname("D:/train/samples"),
                      tools::file_path_sans_ext(basename("D:/train/samples")))
.
.
.
result_RF = predict(RO,
                    ro,
                    filename = "classification_RF.tif",
                    overwrite = TRUE)

plot(result_RF, add=TRUE)
varImp(ro, scale=FALSE)
ra = aggregate(result_RF, fact=2)
writeRaster(ra, "D:/train/arcgis/RF.asc", format="ascii")
```

MARS algoritması için ise “earth” paket kodlaması kullanılarak toplam 43 satır kodlama ile sınıflandırma gerçekleştirilmiştir. Öncelikle “raster” ve “earth” paketleri kütüphaneden yüklenmiş ve alıştırma örnekleri açılmıştır. Ardından MARS algoritması parametreleri ayarlanarak “for” döngüsü oluşturulup sınıflandırma gerçekleştirilmiş ve sonuç görüntü, harita oluşturmak amacıyla ASCII formatında kaydedilmiştir.

```
library(raster)
library(earth)
samples15 <- readOGR(dirname("D:/train/samples"),
```

```

tools::file_path_sans_ext(basename("D:/train/samples")))
.
.
.
summary(mars, digits=3, style="max")
evimp(mars)

result <- predict(RGB,
                  mars,
                  filename = "classification_mars.tif",
                  overwrite = TRUE)

writeRaster(ra, "D:/train/arcgis/MARS.asc", format="ascii")

```

MARS-RF hibrit model RStudio ve MATLAB yazılımları kullanılarak gerçekleştirilmiştir. Yukarıda anlatıldığı gibi MARS ve RF algoritmaları RStudio yazılımı kullanılarak sınıflandırma gerçekleştirilmiştir. RStudio yazılımında ASCII formatında elde edilen görüntüler MATLAB yazılımında açılmış ve iki algoritma sonucundan paralel kombinasyon yöntemine göre hibrit model oluşturulmuştur. Hibrit model oluşturmak için öncelikle MATLAB yazılımında açılan sonuç görüntüler incelenmiştir. Bu inceleme paralel kombinasyon ve belirlenen karar kuralınca; her iki algoritmada da aynı sınıfa atanan pikseller ve iki algoritmada farklı sınıflara ayrılan piksellerin belirlenmesini içermektedir. Bu inceleme MATLAB yazılımında hazırlanan kod ile yapılmıştır. Belirlenen piksellere karar kuralına göre piksellerin sınıf ataması olması amacıyla “if” döngüsü kodlaması yapılmıştır ve MARS-RF hibrit model sonucu elde edilmiştir. Karar kuralının MATLAB yazılımında kodu aşağıdaki gibidir;

```

% Karar Kuralı
combined_preds = zeros(size(y_test));
for i = 1:length(y_test)
    if mars_preds(i) == rf_preds(i)
        combined_preds(i) = mars_preds(i);
    else
        mars_class = mars_preds(i);
        rf_class = rf_preds(i);
        if mars_pa(mars_class) > rf_pa(rf_class)

```

```

        combined_preds(i) = mars_class;
    else
        combined_preds(i) = rf_class;
    end
end
end
end

```

MARS-PSO hibrit modeli MARS algoritması parametrelerine optimizasyon uygulanması ile elde edilen modeldir. MARS algoritması maksimum temel fonksiyon sayısı ve ceza parametresi parametrelerinin en uygun değerlerinin belirlenmesi amacıyla PSO algoritmasının kullanılarak optimize edilmesi MATLAB yazılımında 109 satır kodlama ile gerçekleştirilmiştir.

% MARS algoritmasının hiperparametrelerinin belirleneceği aralıklar belirlenir

```
lower_bound = [3 2]; % alt sınır
```

```
upper_bound = [50 5]; % üst sınır
```

% PSO algoritmasının parametreleri belirlenir

```
num_particles = 50; % parçacık sayısı
```

```
max_iterations = 100; % maksimum iterasyon sayısı
```

```
.
```

```
.
```

```
.
```

% En iyi hiperparametreler bulunur

```
best_params = bestPoint(results);
```

% MARS modeli oluşturulur

```
model = fitMars(X, Y, ...
```

MARS-BO hibrit modeli MARS algoritması parametrelerine optimizasyon uygulanması ile elde edilen modeldir. MARS algoritması maksimum temel fonksiyon sayısı ve ceza parametresi parametrelerinin en uygun değerlerinin belirlenmesi amacıyla BO algoritmasının kullanılarak optimize edilmesi MATLAB yazılımında kodlama ile gerçekleştirilmiştir.

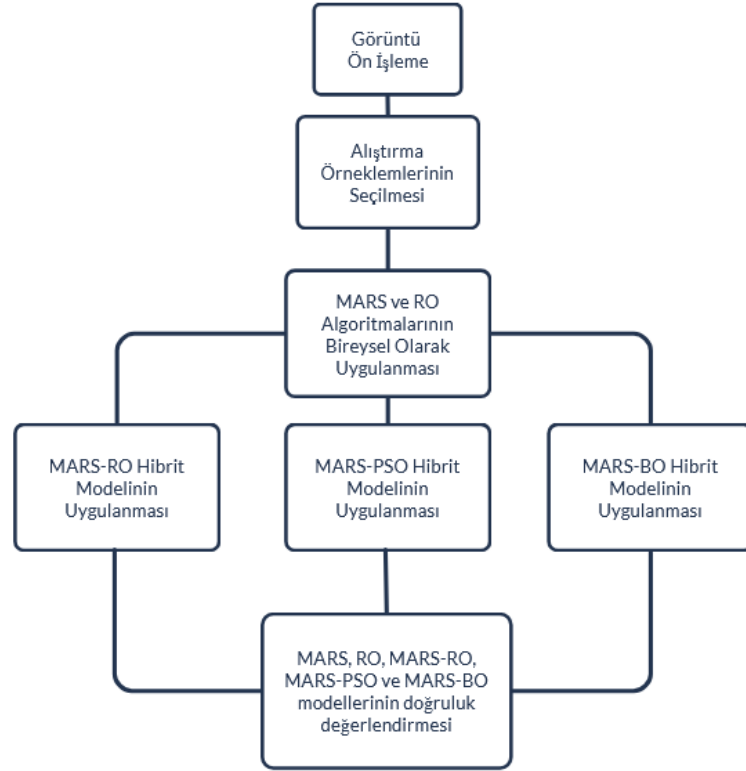
```

% BO için aralıklar belirlenir
hyperparams.min_nodes = 2;
hyperparams.max_nodes = 20;
hyperparams.min_threshold = 0.01;
hyperparams.max_threshold = 0.5;
hyperparams.min_degree = 1;
hyperparams.max_degree = 3;
% MARS algoritmasının hiperparametrelerinin belirleneceği aralıklar
belirlenir
lower_bound = [3 2]; % alt sınır
upper_bound = [50 5]; % üst sınır
.
.
.
% En iyi hiperparametreler bulunur
best_params = bestPoint(results);
.
% MARS modeli oluşturulur
model = fitMars(X, Y, ...

```

3.3. Uygulama ve Modellerin Testi

MARS algoritmasına dayalı hibrit sınıflandırma modellerinin performans değerlendirmesi için iki farklı çalışma alanı ve iki farklı uydu görüntüsü kullanılmıştır. Başka bir ifade ile önerilen yöntemlerin performansı, iki ayrı bölgede farklı çözünürlüklere sahip iki farklı uydu görüntüsü kullanılarak test edilmiştir. Bu kapsamda gerçekleştirilen ana işlem aşamaları Şekil 3.6'da gösterilmiştir. Görüntülerde ilk önce ön işleme gerçekleştirilmiştir. Alıştırma örneklerinin seçiminin ardından MARS ve RF sınıflandırmaları ile MARS-RF, MARS-PSO ve MARS-BO hibrit sınıflandırmaları uygulanmış ve sınıflandırma sonuçlarının karşılaştırılması doğruluk analizi gerçekleştirilmiştir.



Şekil 3.6 Uygulama ve test için iş akış şeması

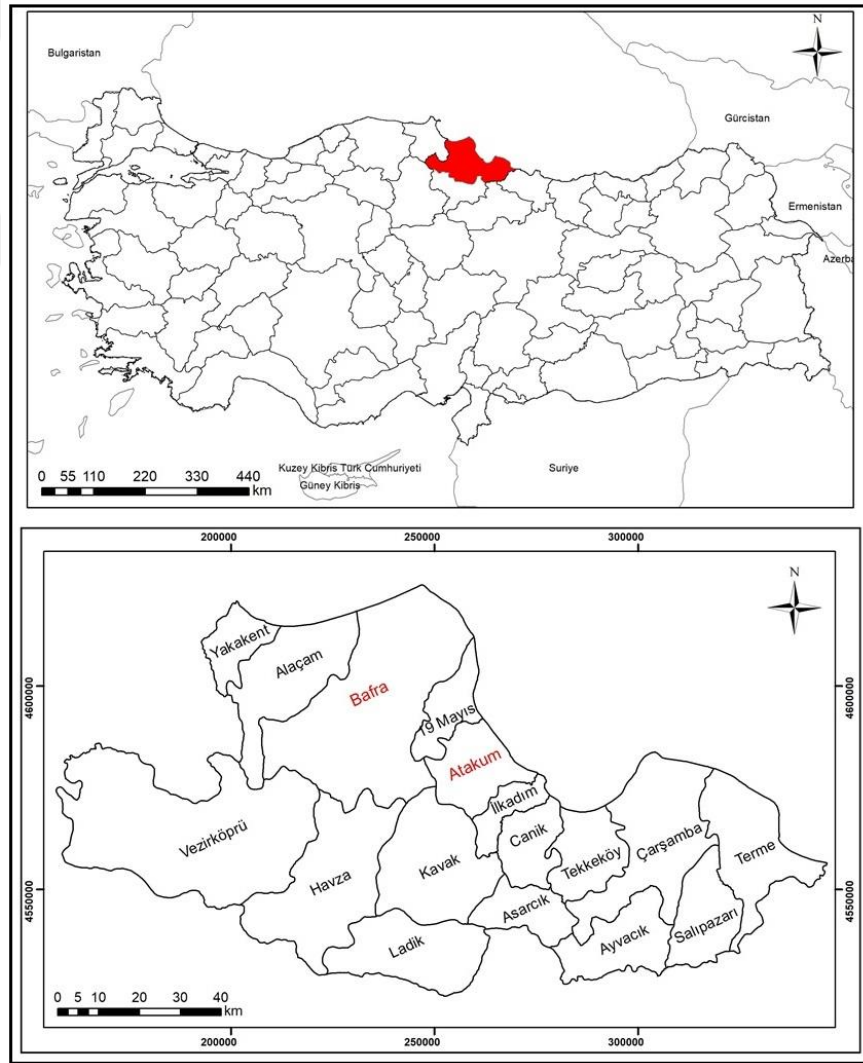
3.3.1. Test Alanlarının Belirlenmesi

Görüntü sınıflandırma yöntemlerinin performanslarını değerlendirmek için, çeşitli arazi örtüsü/kullanımı içeren farklı çalışma alanları gereklidir. Bu yaklaşım, sınıflandırma yöntemlerinin çeşitli koşullardaki performanslarını ve yeteneklerini anlamak için önemlidir. Farklı çalışma alanlarından elde edilen veri setleri, görüntü sınıflandırma algoritmalarının çeşitli coğrafi ve topolojik özelliklere nasıl tepki verdiğini değerlendirmek için kullanılabilir. Aynı zamanda, bu veri setleri, yöntemlerin genellikle farklı iklim koşulları, bitki örtüsü tipleri veya yerel özelliklere uyum sağlama yeteneklerini değerlendirmek için de kullanılabilir. Böylece sınıflandırma yöntemlerinin performansları, farklı arazi örtüsü/kullanımı için karşılaştırılabilir (Gregorio and Jansen, 1998; Li et al., 2014; Mather and Tso, 2016).

Test alanı 1 olarak seçilen Bafra ilçesi, Samsun ilinin kuzeybatısında, Karadeniz'e kıyısı olan bir yerleşim yeridir ve yaklaşık olarak $41^{\circ} 35' - 41^{\circ} 45'$ kuzey enlemleri ve $35^{\circ} 54' - 36^{\circ} 15'$ doğu boylamları arasında yer alır (Şekil 3.6). İlçenin toplam yüzölçümü Harita Genel Müdürlüğü (HGM) verilerine göre 1503 km^2 olup, Türkiye İstatistik Kurumu (TÜİK) 2023 yılı verilerine göre 143109 nüfusa sahiptir. Bafra ilçesi, Karadeniz kıyısından iç kesimlere doğru uzanan geniş bir alanı kaplamaktadır. İlçe tarımsal üretim açısından Samsun'un en önemli bölgelerinden

biridir. İlçede ayrıca ormanlık alanlar ve kırsal yerleşimler bulunmaktadır. Bu çeşitlilik ile Bafra arazi kullanımı yönünden farklılık göstermekte ve çeşitli ekosistemlere ev sahipliği yapmaktadır (Hekimoğlu vd., 2007).

Test alanı 2 olarak seçilen Atakum ilçesi, Samsun ilinin batısında, Karadeniz kıyısında yer almaktadır (Şekil 3.7). İlçe, yaklaşık $41^{\circ} 20'$ - $41^{\circ} 25'$ kuzey enlemleri ve $36^{\circ} 18'$ - $36^{\circ} 24'$ doğu boylamları arasında konumlanmıştır. Bu coğrafi konumla, Atakum hem deniz hem de dağ topografyasına sahiptir. Atakum'un toplam yüzölçümü HGM verilerine göre 351 km^2 olup, TÜİK 2023 yılı verilerine göre 245328 nüfusa sahiptir. Arazi kullanımı ve örtüsü sınıflandırması açısından Atakum'da kentsel alanlar, tarım arazileri ve ormanlık bölgeler öne çıkmaktadır. Bu çeşitlilik, arazi kullanımı ve örtüsü sınıflandırma çalışmaları için zengin veri sağlamaktadır (Hekimoğlu vd., 2007).



Şekil 3.7 Test alanları: Bafra ve Atakum ilçeleri (Samsun)

3.3.2. Uydu Görüntüleri ve Bantların Seçimi

Tez kapsamında farklı arazi örtüsü/kullanımı türlerinin ayırt edilmesinde Landsat-8 OLI ve Sentinel-2A MSI uydu görüntüleri kullanılmıştır. Bu çalışmada, Landsat-8 OLI uydu görüntüsünün 6 bandı (Bant 2, Bant 3, Bant 4, Bant 5, Bant 6 ve Bant 7) ve Sentinel-2A MSI uydu görüntüsünün 10 bandı (Bant 2, Bant 3, Bant 4, Bant 5, Bant 6, Bant 7, Bant 8, Bant 8A, Bant 11 ve Bant 12) kullanılmıştır. Bu bantların seçiminde, çalışma alanlarının özellikleri ve belirlenen sınıflar değerlendirmeye alınmıştır. Uydu sistemleri ile ilgili teknik bilgiler aşağıdaki bölümlerde detaylı olarak verilmiştir.

3.3.2.1. Sentinel-2A/B

Sentinel-2 uydu sistemi Sentinel-2A ve Sentinel-2B olmak üzere iki uydudan oluşmaktadır. Sentinel-2A, Haziran 2015, Sentinel-2B ise Temmuz 2016 tarihinde fırlatılmış olup, Avrupa Uzay Ajansı'nın (ESA) yeryüzünden 786 km yükseklikteki yörüngesinde, güneşle senkronize bir yer gözlem uydusudur. Sentinel-2A ve Sentinel-2B takım uyduları gibi çalıştığı ve aynı çözünürlüklere sahip oldukları için birlikte kullanılabilirler. Sentinel-2, 10 m (Bant 2, Bant 3, Bant 4, Bant 8), 20 m (Bant 5, Bant 6, Bant 7, Bant 8A, Bant 11, Bant 12) ve 60 m (Bant 1, Bant 9, Bant10) çözünürlüğe sahiptir. Bu çalışmada kullanılan Sentinel-2A MSI görüntüleri ESA web sitesinden indirilmiştir. Sentinel-2A MSI uydusuna ait teknik özellikler Tablo 3.1'de verilmiştir (Sentinel-2A User Handbook, 2015).

Tablo 3.1 Sentinel-2A MSI uydu görüntüsünün teknik özellikleri

Bantlar	Merkez dalgaboyu (nm)	Piksel boyutu (m)
Bant 1- Coastal	443	60
Bant 2- Blue	490	10
Bant 3- Green	560	10
Bant 4- Red	665	10
Bant 5- Vegetation Red Edge	705	20
Bant 6- Vegetation Red Edge	740	20
Bant 7- Vegetation Red Edge	783	20
Bant 8- NIR	842	10
Bant 8A- Narrow NIR	865	20
Bant 9- Water Vapour	945	60
Bant 10- SWIR Cirrus	1375	60
Bant 11- SWIR1	1610	20
Bant 12- SWIR2	2190	20

3.3.2.2. Landsat-8 OLI

11 Şubat 2013 tarihinde yeryüzünden 705 km yükseklikteki yörüngesine oturtulan Landsat-8 OLI, güneşle senkronize bir yer gözlem uydusudur. Uydu üzerinde iki algılayıcı bulunmaktadır: OLI (Operational Land Imager) ve TIRS (Thermal Infrared Sensor). Landsat-8 OLI, 15 m PAN, 30 m MS (RGB, NIR, SWIR) ve 100 m termal bant çözünürlüğüne sahiptir. Bu çalışmada kullanılan Landsat-8 OLI görüntüleri, Amerika Birleşik Devletleri Jeoloji Araştırmaları Merkezi'nin (USGS) web sitesinden indirilmiştir. Tablo 3.2'de Landsat-8 OLI uydusuna ait teknik özellikler gösterilmektedir (Landsat-8 OLI Data Users Handbook, 2018).

Tablo 3.2 Landsat-8 OLI uydu görüntüsünün teknik özellikleri

Bantlar	Merkez dalgaboyu (nm)	Piksel boyutu (m)
Bant 1 - Coastal aerosol	443	30
Bant 2 - Blue	483	30
Bant 3 - Green	560	30
Bant 4 - Red	660	30
Bant 5 - Near Infrared (NIR)	865	30
Bant 6 - SWIR 1	1650	30
Bant 7 - SWIR 2	2220	30
Bant 8 - Panchromatic	640	15
Bant 9 - Cirrus	1375	30

3.3.3. Sınıflandırmada Kullanılacak Arazi Örtüsü/Kullanımı Sınıflarının Belirlenmesi

Dünya genelinde, özellikle gelişmekte olan ülkelerde, tarımsal üretime olan yoğun bağımlılıkları ve artan nüfus nedeniyle arazi örtüsü/kullanımında hızlı değişimler gözlenmektedir. Bu değişimlerin etkin planlama ve yönetimi için güncel ve doğru arazi örtüsü/kullanımı bilgisinin tespit edilmesi gerekmektedir. Bu durum, kısa süreler içerisinde daha geniş alanlardan farklı arazi örtüsü/kullanımı verilerinin elde edilmesi için modern metodolojilerin geliştirilmesini gerektirmektedir (Anderson et al., 1976; Büttner 2011; Nedd et al., 2021).

Bu çalışmada belirlenecek sınıf sayısı ve sınıflar için literatürde kabul görmüş üç adet arazi örtüsü/kullanımı sınıflandırma sistemi incelenmiştir; USGS Arazi örtüsü/kullanımı sınıflandırma sistemi (Anderson et al., 1976), Gıda ve Tarım Örgütü'nün (Food and Agriculture Organization – FAO) Arazi örtüsü sınıflandırma sistemi ve CORINE (Coordination of Information on the Environment – Çevresel

Bilginin Koordinasyonu) (Nedd et al., 2021). Çalışma alanlarının içerdikleri sınıf özellikleri dikkate alınarak CORINE sisteminin kullanılmasına karar verilmiştir.

CORINE Avrupa Çevre Ajansının oluşturduğu arazi örtüsü/kullanımı sınıflandırmasına göre uydu görüntüleri üzerinden bilgisayar destekli görsel yorumlama metodu ile üretilen arazi örtüsü/kullanımı verisidir. Sınıflandırma sistemi, Avrupa genelinde arazi örtüsü/kullanımı verilerinin standart bir metodoloji ile toplanmasını, düzenlenmesini ve analiz edilmesini sağlayan bir sistemdir. Literatür incelendiğinde CORINE sınıflandırma sistemi, çevre yönetimi, planlama, araştırma ve politika geliştirme gibi birçok alanda yaygın olarak kullanılmaktadır (Grullon et al., 2009; Nedd et al., 2021).

Bu çalışmada CORINE arazi örtüsü/kullanımı sınıflandırma sisteminin belirlediği kriterler ve sınıflandırma sistemi incelenerek ve çalışma alanları göz önünde bulundurularak; Bafra ilçesi için “Yapay alanlar”, “Ekilebilir Alanlar”, “Karışık Tarım Alanları”, “Orman ve Yarı Doğal Alanlar”, “Sulak Alanlar”, “Suyla Kaplı Alanlar” sınıfları; Atakum ilçesi için “Yapay Alanlar”, “Tarımsal Alanlar”, “Orman ve Yarı Doğal Alanlar”, “Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar”, “Suyla Kaplı Alanlar” sınıfları seçilmiştir. Arazi örtüsü/kullanımı sınıfları, çalışılan alan ve işlevsellik gibi çeşitli gereksinimlere esnek bir şekilde uyarlanmalıdır. Bu nedenle farklı çalışma alanları için farklı sınıflar ve sınıf seviyeleri seçilmiştir.

Bu çalışmada, MARS ve RF tekil algoritmaları ve MARS algoritmasına dayalı üç farklı hibrit model (MARS-RF, MARS-PSO ve MARS-BO) uygulanmıştır. Gerçekleştirilen tüm sınıflandırma analizleri için iki farklı uydu sistemi kullanılmıştır;

- 23.07.2023 tarihli Landsat-8 OLI uydu görüntüsünün 6 bandı (Bant 2, Bant 3, Bant 4, Bant 5, Bant 6 ve Bant 7);
- 27.07.2023 tarihli Sentinel-2A MSI uydu görüntüsünün 10 bandı (Bant 2, Bant 3, Bant 4, Bant 5, Bant 6, Bant 7, Bant 8, Bant 8A, Bant 11 ve Bant 12) kullanılmıştır.

3.3.4. Görüntü Ön İşleme

Uydu görüntüleri çevresel etkiler nedeniyle sistematik ve sistematik olmayan hatalar içermektedir. Bu nedenle, uydu görüntülerine uygulamalarda kullanılmadan önce düzeltmelerin uygulanması gerekmektedir.

Çalışma kapsamında kullanılan Landsat-8 OLI Collection 2 Level-2 görüntülerinin radyometrik kalibrasyon ve atmosferik düzeltme işlemi USGS

tarafından uygulanmıştır ve analize hazır verilerdir. Atmosferik düzeltme, arazi yüzey yansıma kodu (Land Surface Reflectance Code – LaSRC) algoritması kullanılarak gerçekleştirilmiştir (USGS, 2023).

Çalışma kapsamında Sentinel-2A Seviye-1C uydu görüntüsüne uygulanan atmosferik düzeltme işlemcisi Sen2Cor, ESA tarafından geliştirilmiştir. Sen2Cor, temel amacı Sentinel-2 Seviye-1C Atmosfer Üstü (TOA) ürünlerini atmosferin etkilerinden düzelterek Atmosfer Altı (BOA) yüzey yansıma ürünü sunmaktır (Main-Knorn et al., 2017). Görüntü ön işlemesi için Sentinel Application Platform (SNAP) yazılımında bulunan Sen2Cor aracı kullanılmıştır. Sentinel-2A MSI uydu görüntüsü çalışma alanlarından Bafra ilçesi için iki çerçevede yer almaktadır bu nedenle ArcGIS yazılımında görüntü mozaikleme işlemi gerçekleştirilmiştir.

Sentinel-2A MSI uydu görüntülerinin tüm bantlarının piksel boyutu yönünden uyumlu olmasını sağlamak için 20 m piksel boyutlu bantlar en yakın komşu enterpolasyonu ile 10 m piksel boyutuna yeniden örneklenmiştir. Son olarak uydu görüntüleri, çalışma alanları olan Bafra ve Atakum ilçe idari sınırlarına göre kesilmiştir. Suyla kaplı alanların da çalışmaya eklenmesi amacıyla sınırlar deniz yönünde 2 km genişletilmiştir.

3.3.5. Alıştırma Örneklemelerinin Seçilmesi

Bir sınıflandırma yönteminin performansı, alıştırma örneklem veri setinin boyutu, çalışma alanının karmaşıklığı, görüntülerin çözünürlüğü ve kullanılan sınıflandırma algoritması dahil olmak üzere çeşitli faktörlere bağlıdır. Alıştırma örneklemelerinin sayısı, kullanılan veri ve algoritma türüne göre değişir. Bu sayı, alıştırma örneklemelerinin sınıf üyeliği hakkında yeterli bilgi edinmek için yeterince büyük olmalı ancak eğitim verilerinin aşırı eğitilmesine neden olacak kadar büyük olmamalıdır. Alıştırma örneklemelerinin seçimi sınıflandırma performansını doğrudan etkileyen önemli faktörlerdendir ve alıştırma örneklemelerinin seçiminde yardımcı verilerin kullanılması gerekmektedir. Makine öğrenmesi alanında yapılan önceki araştırmalar, alıştırma örneklem veri setinin yalnızca eğitim hızını etkilemekle kalmayıp aynı zamanda sınıflandırıcıların performansı üzerinde de etkili olduğunu göstermektedir (Rendell and Cho, 1990; Shavlik, 1991; Brodley, 1995).

Bu tez kapsamında, ENVI yazılımında oluşturulan standart sahte renkli görüntü, kontrolsüz sınıflandırma yöntemi olan ISODATA sınıflandırma sonuç görüntüsü, NDVI indeks sonuç görüntüsü ve Google Earth Pro görüntüleri kullanılarak alıştırma örneklemeleri seçilmiştir. Bu yardımcı veri kümelerinin dahil

edilmesiyle, alıştırma örneklemelerinin seçiminde daha doğru ve güvenilir sonuçlar elde edilir. Bu veri kaynakları, uydu görüntülerinden elde edilen spektral bilgileri tamamlayan yardımcı verilerdir. Bu yardımcı veriler kullanılarak uydu görüntüsünden alıştırma örneklemeleri verisi için ENVI yazılımında ROI'ler (region of interest) oluşturulup, ASCII formatında kaydedilmiştir. ASCII formatında kaydedilen spektral yansıtım değerleri sınıf bilgisi de eklenerek düzenlenmiştir. Eğitim ve model performanslarını değerlendirmek amacıyla; Bafra ilçesi için toplam 2248 ve Atakum ilçesi için toplam 1738 piksel veri örneği kullanılmıştır. Veri setleri eğitim seti (%70) ve test seti (%30) olarak ikiye ayrılmıştır. Tablo 3.3'te her sınıf için eğitim ve test veri seti piksel sayıları gösterilmektedir.

Tablo 3.3 Eğitim ve test veri sayıları

Çalışma Alanı	Sınıf	Eğitim Veri Sayısı	Test Veri Sayısı	Toplam
Bafra	Suyla Kaplı Alanlar	242	104	2248
	Sulak Alanlar	252	108	
	Orman ve Yarı Doğal Alanlar	264	113	
	Karışık Tarımsal Alanlar	334	107	
	Ekilebilir Alanlar	249	105	
	Yapay Alanlar	259	111	
Atakum	Suyla Kaplı Alanlar	235	101	1738
	Orman ve Yarı Doğal Alanlar	247	106	
	Tarımsal Alanlar	247	106	
	Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	245	105	
	Yapay Alanlar	242	104	

3.3.6. Görüntü Sınıflandırma

Bu çalışma kapsamında MARS ve RF algoritmaları temel sınıflandırıcı olarak kullanılmıştır. Temel sınıflandırıcıların algoritmaları RStudio yazılımında gerçekleştirilmiştir. Hibrit modeller ise, RStudio yazılımından elde edilen sonuç görüntülerin MATLAB yazılımında hazırlanan kodlama ile işlenmesiyle oluşturulmuştur.

RF algoritması ile görüntü sınıflandırma işlemi RStudio yazılımında “randomForest” paket programıyla gerçekleştirilmiştir. RF algoritmasında kullanıcının belirlemesi gereken her bir düğümde kullanılan değişkenlerin sayısı (m) ve oluşturulacak ağaçların sayısı gibi parametreler bulunmaktadır. Bu nedenle bu çalışmada, parametreleri ayarlamak, doğruluk açısından en iyi sonucu veren

parametrelerin optimum büyüklüklerini elde etmek için ağaç sayısı (Ntree - 'n_estimators') ve değişken sayısı (Mtry - 'max_features') değer aralıkları dikkate alınmıştır. Burada Ntree ('n_estimators') =20 ve Mtry ('max_features') = auto, yani değişken sayısının kareköküdür. Ardından, Ntree değeri 50 ve 100'e çıkarılmış ve Mtry'yi varsayılan değerlerde (auto) tutarak bunlar test edilmiştir. En iyi Ntree değeri belirlendikten sonra, bu değer sabit tutulmuş ve Mtry = 5, 7, 10 gibi bir dizi büyüklüğü göz önünde bulundurarak Mtry'nin en iyi değerini bulmak için algoritma tekrar çalıştırılmıştır. RF algoritmasında bu iki parametrenin en iyi iki değerine karar vermek için toplam dokuz test gerçekleştirilmiştir.

MARS sınıflandırıcısı ise RStudio yazılımında “earth” paketinde çalıştırılmıştır. Açık kaynak kodlu bir yazılım olan RStudio ile çapraz geçerlilik, ceza parametresi, modelin derecesi, maksimum fonksiyon sayısı gibi parametreler, kodlamada düzenlemeler yaparak değiştirilebilmektedir. MARS algoritmasında genelleştirilmiş çapraz geçerlilik değeri en küçük olan model en ideal model olarak seçilmektedir. Bu doğrultuda RStudio programı içerisinde en küçük değere sahip genelleştirilmiş çapraz geçerlilik hesaplatılması için döngü oluşturulmuş, MARS algoritması bu döngüye bağlı çalıştırılmıştır. Ceza parametre değeri (penalty) yaklaşık 2 ila 5 aralığında değerler önermektedir. İleri adımda belirlenen maksimum fonksiyon sayısı (nk) 20 olarak belirlenmiştir. Maksimum etkileşim derecesi ise (degree) varsayılan ayarlarda 1 olarak tanımlanmıştır yani etkileşim terimleri modele dahil etmez, bu çalışmada (degree) 1 ve 2 olarak belirlenmiştir. Çapraz doğrulama katlarının sayısı (nfold) varsayılan değeri 0'dır yani çapraz doğrulama yoktur. 1'den büyükse, kodlamada (earth) önce tüm verilerle her zamanki gibi standart bir model oluşturur, sonra her seferinde dışarıda bırakılmış veriler üzerinde en küçük kareler yöntemine göre modeller oluşturur. Bu parametreler denenerek en küçük GCV değerinin bulunmasıyla ideal MARS modeli yani temel fonksiyonlar elde edilmiş olur. Bu temel fonksiyonlar kullanılarak tüm görüntü sınıflandırılmıştır.

3.3.7. Sınıflandırma Doğruluk Analizi

Uydu görüntülerinin sınıflandırma doğruluk analizi, sınıflandırmanın ne kadar başarılı olduğunu belirlemek için kullanılır. Sınıflandırma yöntemlerini karşılaştırmada hata matrisi kullanılarak kullanıcı doğruluğu/kesinlik, üretici doğruluğu/duyarlılık, toplam doğruluk, kappa istatistiği ve F1 skor gibi doğruluk ölçütleri kullanılır.

Bu çalışmada, kullanıcı doğruluğu/kesinlik, üretici doğruluğu/duyarlılık ve F1

skor ölçütleri sınıf bazında karşılaştırmada kullanılmıştır. Sınıflandırma sonucunun genel karşılaştırması ise toplam doğruluk, kappa istatistiği ve makro ortalama alınarak hesaplanan F1 skor değerleri ile yapılmıştır. Bu doğruluk ölçütlerinin kullanılması, modelin performansının genel ve sınıf bazında kapsamlı bir şekilde değerlendirme yapılmasını sağlar. Bu sayede, sınıflandırıcıların nasıl performans gösterdiği ve hangi yöntemlerin daha güvenilir sonuçlar verdiği hakkında daha detaylı bilgi elde edilebilir.



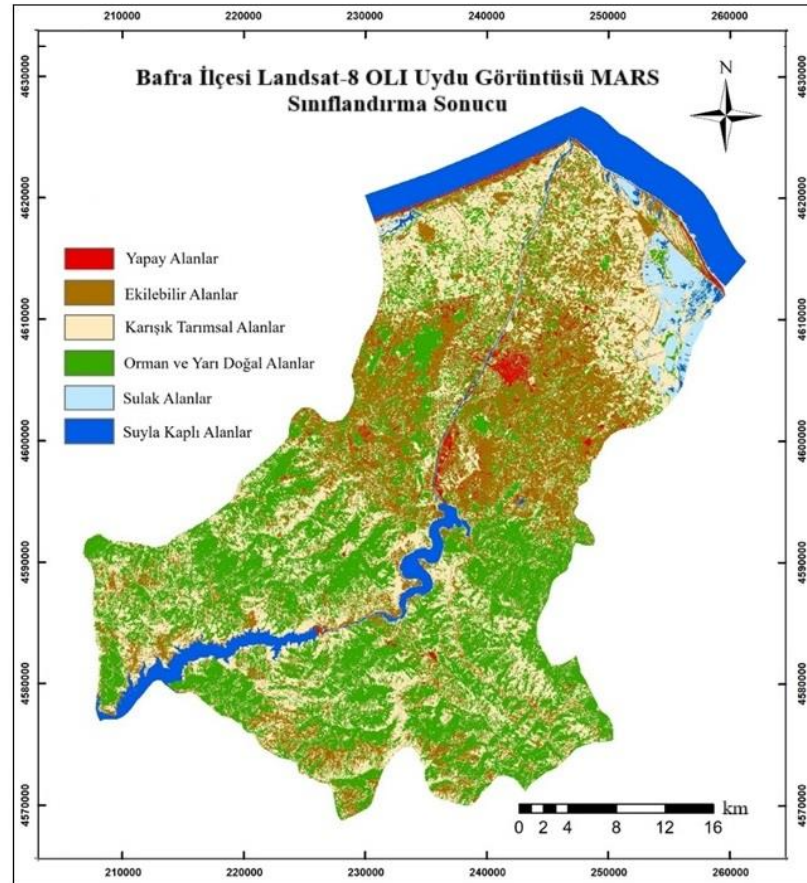
4. BULGULAR VE TARTIŞMA

MARS, RF, MARS-RF, MARS-PSO ve MARS-BO'nun sınıflandırma doğruluğunu belirlemek için, sınıf bazında kullanıcı doğruluğu/kesinlik, üretici doğruluğu/duyarlılık ve F1 skor ölçütleri ve genel sınıflandırma değerlendirmesi için toplam doğruluk, kappa değeri ve makro ortalama ile hesaplanan F1 skor ölçütleri kullanılmış ve tüm modeller için elde edilen sonuçlar karşılaştırılmıştır.

4.1. Test Alanı 1: Bafra İlçesi Sınıflandırma Sonuçları

4.1.1. Test Alanı 1: Bafra İlçesi MARS Sınıflandırma Sonuçları

Test alanı 1 olarak seçilen Bafra ilçesi MARS algoritma sınıflandırması iki uydu görüntüsü için ayrı ayrı gerçekleştirilmiştir. Şekil 4.1 Landsat-8 OLI uydu görüntüsüne ait MARS sınıflandırma sonucunu göstermektedir. Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS sınıflandırma sonucu hata matrisi Tablo 4.1'de verilmiştir.

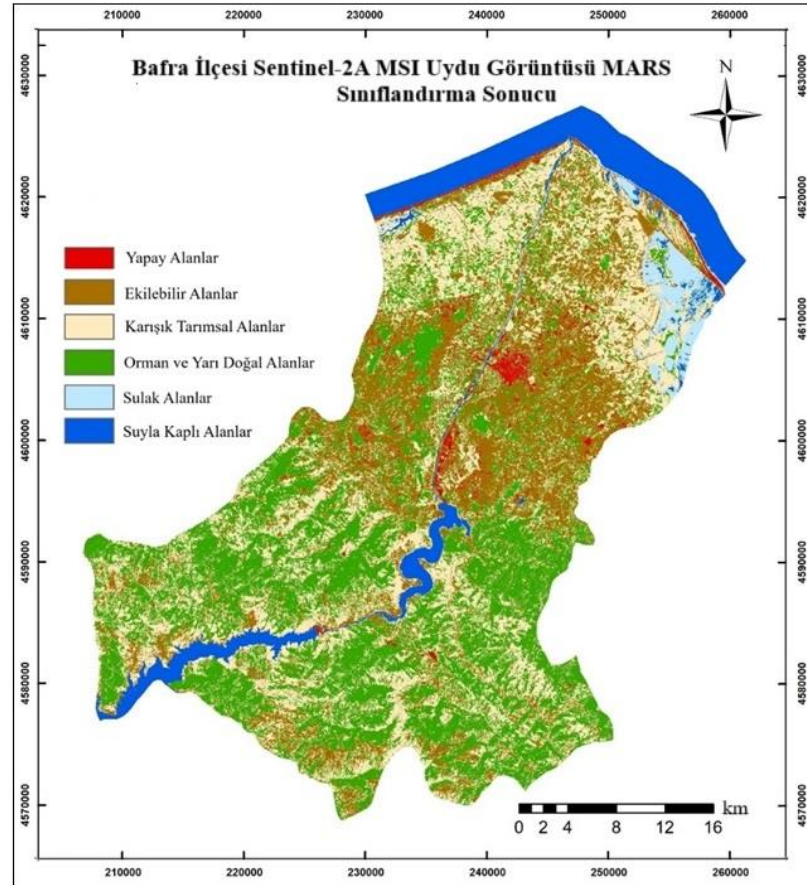


Şekil 4.1 Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS sınıflandırma sonucu

Tablo 4.1 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS sınıflandırma sonucu hata matrisi

	Referans Verisi						
	Yapay Alanlar	Ekilebilir Alanlar	Karışık Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Sulak Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	82	10	14	10	0	0	116
Ekilebilir Alanlar	17	75	4	0	0	0	96
Karışık Tarımsal Alanlar	3	16	80	13	0	0	112
Orman ve Yarı Doğal Alanlar	9	4	9	85	8	0	115
Sulak Alanlar	0	0	0	5	89	4	98
Suyla Kaplı Alanlar	0	0	0	0	11	100	111
Toplam	111	105	107	113	108	104	648

Şekil 4.2 Sentinel-2A MSI uydu görüntüsüne ait MARS sınıflandırma sonucunu göstermektedir. Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS sınıflandırma sonucu hata matrisi Tablo 4.2’de verilmiştir.



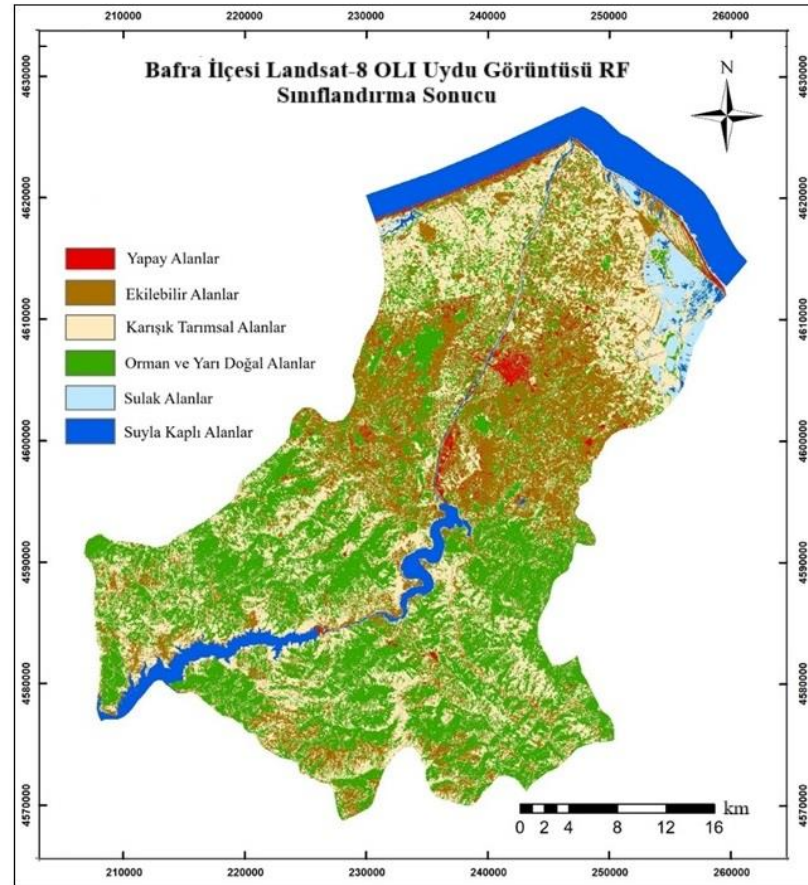
Şekil 4.2 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS sınıflandırma sonucu

Tablo 4.2 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS sınıflandırma sonucu hata matrisi

	Referans Verisi						
	Yapay Alanlar	Ekilebilir Alanlar	Karışık Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Sulak Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	90	6	13	9	0	0	118
Ekilebilir Alanlar	12	74	2	0	0	0	88
Karışık Tarımsal Alanlar	0	20	87	11	0	0	118
Orman ve Yarı Doğal Alanlar	9	5	5	87	8	0	114
Sulak Alanlar	0	0	0	6	83	5	94
Suyla Kaplı Alanlar	0	0	0	0	17	99	116
Toplam	111	105	107	113	108	104	648

4.1.2. Test Alanı 1: Bafra İlçesi RF Sınıflandırma Sonuçları

Şekil 4.3 Landsat-8 OLI uydu görüntüsüne ait RF sınıflandırma sonucunu göstermektedir. Bafra ilçesi Landsat-8 OLI uydu görüntüsü RF sınıflandırma sonucu hata matrisi Tablo 4.3'te verilmiştir.

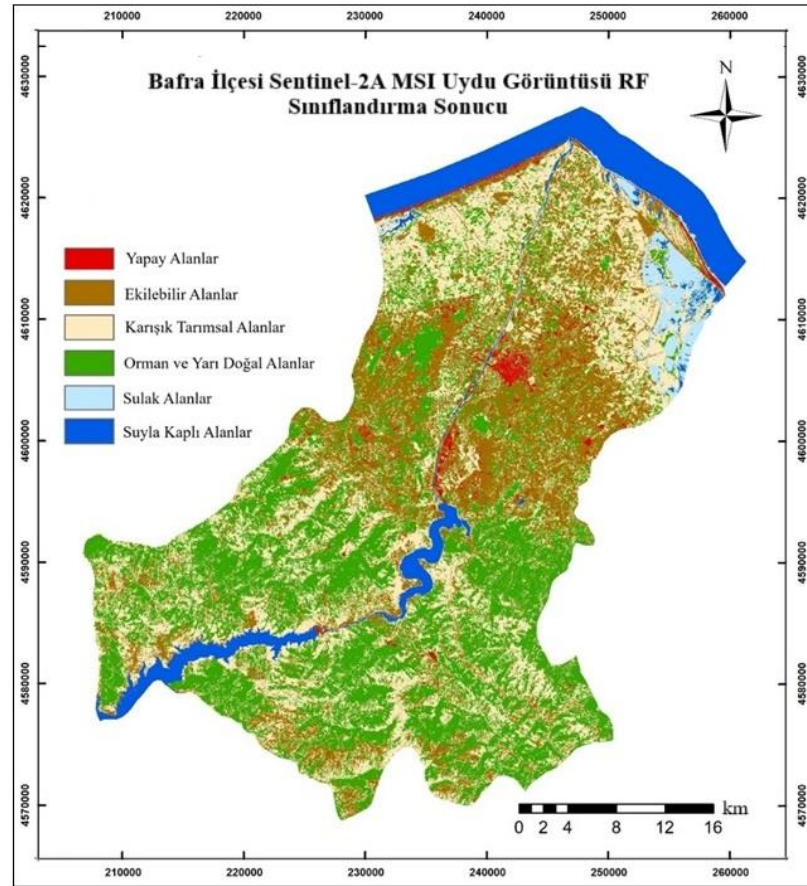


Şekil 4.3 Bafra ilçesi Landsat-8 OLI uydu görüntüsü RF sınıflandırma sonucu

Tablo 4.3 Bafra ilçesi Landsat-8 OLI uydu görüntüsü RF sınıflandırma sonucu hata matrisi

	Referans Verisi						
	Yapay Alanlar	Ekilebilir Alanlar	Karışık Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Sulak Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	80	9	15	10	0	0	114
Ekilebilir Alanlar	18	76	3	0	0	0	97
Karışık Tarımsal Alanlar	4	15	81	14	0	0	114
Orman ve Yarı Doğal Alanlar	9	5	8	83	9	0	114
Sulak Alanlar	0	0	0	6	88	5	99
Suyla Kaplı Alanlar	0	0	0	0	11	99	110
Toplam	111	105	107	113	108	104	648

Şekil 4.4 Sentinel-2A MSI uydu görüntüsüne ait RF sınıflandırma sonucunu göstermektedir. Bafra ilçesi Sentinel-2A MSI uydu görüntüsü RF sınıflandırma sonucu hata matrisi Tablo 4.4’te verilmiştir.



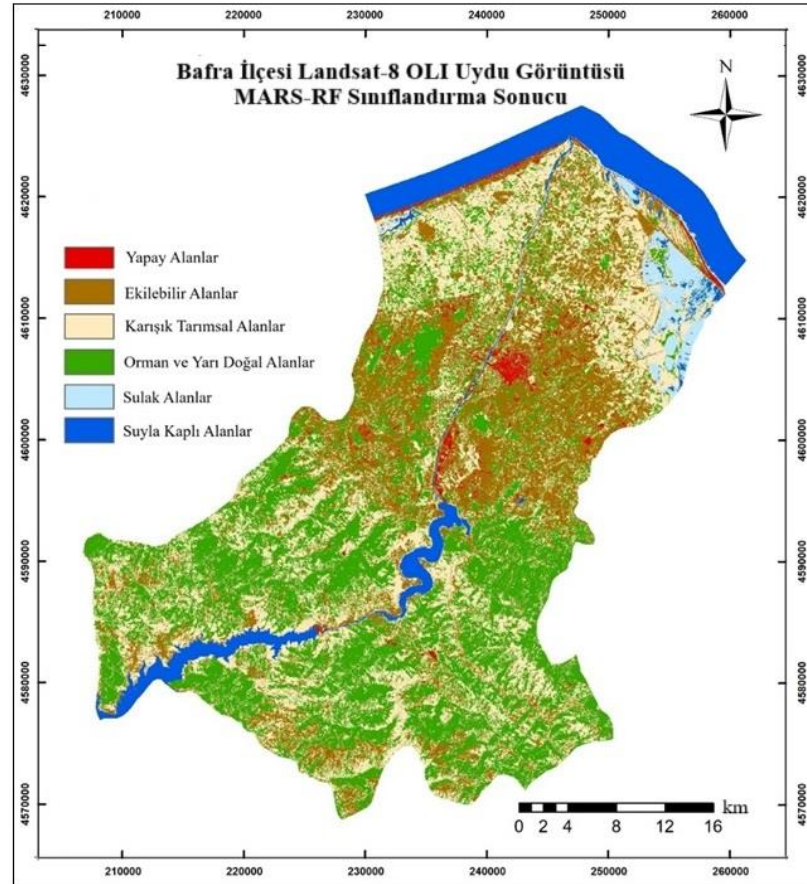
Şekil 4.4 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü RF sınıflandırma sonucu

Tablo 4.4 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü RF sınıflandırma sonucu hata matrisi

	Referans Verisi						
	Yapay Alanlar	Ekilebilir Alanlar	Karışık Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Sulak Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	89	6	14	10	0	0	119
Ekilebilir Alanlar	13	77	1	0	0	0	91
Karışık Tarımsal Alanlar	0	18	87	12	0	0	117
Orman ve Yarı Doğal Alanlar	9	4	5	84	8	0	110
Sulak Alanlar	0	0	0	7	81	5	93
Suyla Kaplı Alanlar	0	0	0	0	19	99	118
Toplam	111	105	107	113	108	104	648

4.1.3. Test Alanı 1: Bafra İlçesi MARS-RF Hibrit Sınıflandırma Sonuçları

Şekil 4.5 Landsat-8 OLI uydu görüntüsüne ait MARS-RF sınıflandırma sonucunu göstermektedir. Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS-RF sınıflandırma sonucu hata matrisi Tablo 4.5'te verilmiştir.

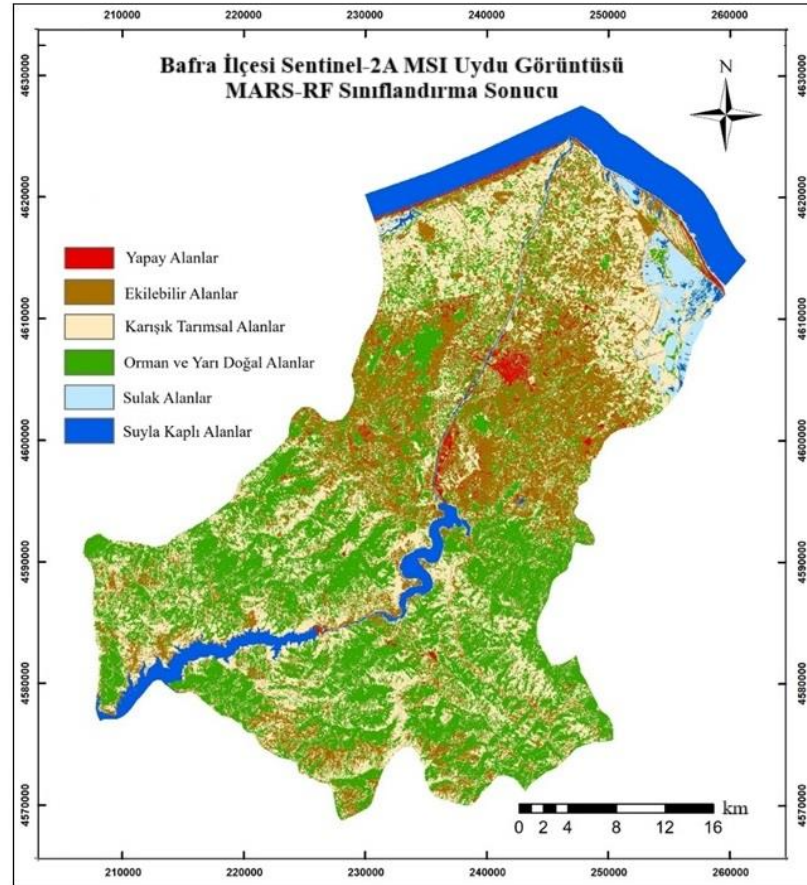


Şekil 4.5 Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS-RF sınıflandırma sonucu

Tablo 4.5 Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS-RF sınıflandırma sonucu hata matrisi

	Referans Verisi						
	Yapay Alanlar	Ekilebilir Alanlar	Karışık Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Sulak Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	82	9	15	10	0	0	116
Ekilebilir Alanlar	17	76	3	0	0	0	96
Karışık Tarımsal Alanlar	3	15	81	13	0	0	112
Orman ve Yarı Doğal Alanlar	9	5	8	85	8	0	115
Sulak Alanlar	0	0	0	5	89	4	98
Suyla Kaplı Alanlar	0	0	0	0	11	100	111
Toplam	111	105	107	113	108	104	648

Şekil 4.6 Sentinel-2A MSI uydu görüntüsüne ait MARS-RF sınıflandırma sonucunu göstermektedir. Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS-RF sınıflandırma sonucu hata matrisi Tablo 4.6’da verilmiştir.



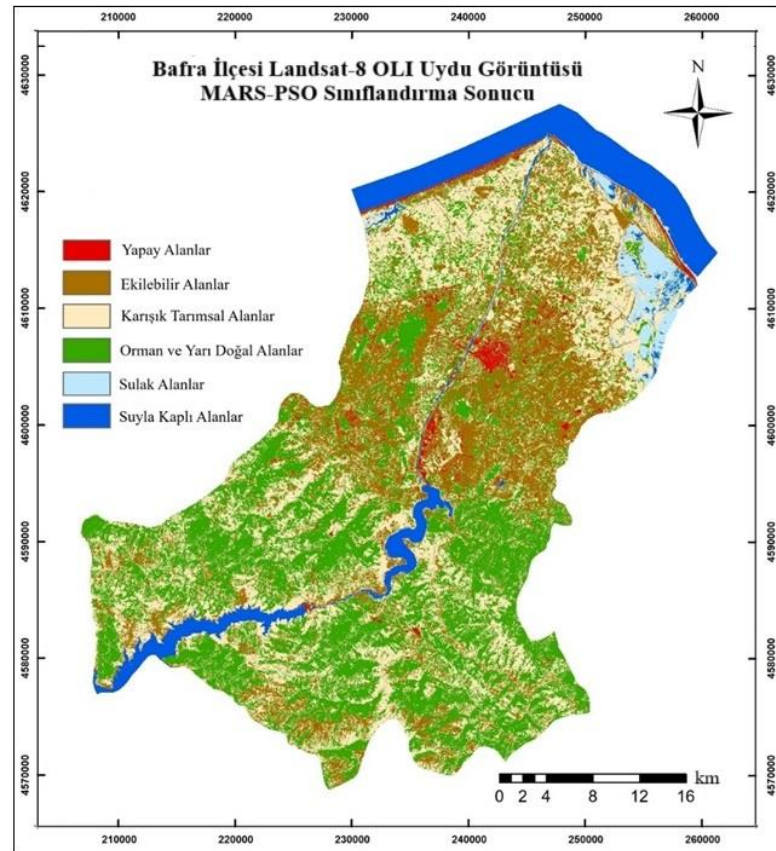
Şekil 4.6 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS-RF sınıflandırma sonucu

Tablo 4.6 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS-RF sınıflandırma sonucu hata matrisi

	Referans Verisi						
	Yapay Alanlar	Ekilebilir Alanlar	Karışık Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Sulak Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	90	6	14	9	0	0	119
Ekilebilir Alanlar	12	77	1	0	0	0	90
Karışık Tarımsal Alanlar	0	18	87	11	0	0	116
Orman ve Yarı Doğal Alanlar	9	4	5	87	8	0	113
Sulak Alanlar	0	0	0	6	83	5	94
Suyla Kaplı Alanlar	0	0	0	0	17	99	116
Toplam	111	105	107	113	108	104	648

4.1.4. Test Alanı 1: Bafra İlçesi MARS-PSO Hibrit Sınıflandırma Sonuçları

Şekil 4.7 Landsat-8 OLI uydu görüntüsüne ait MARS-PSO sınıflandırma sonucunu göstermektedir. Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS-PSO sınıflandırma sonucu hata matrisi Tablo 4.7’de verilmiştir.

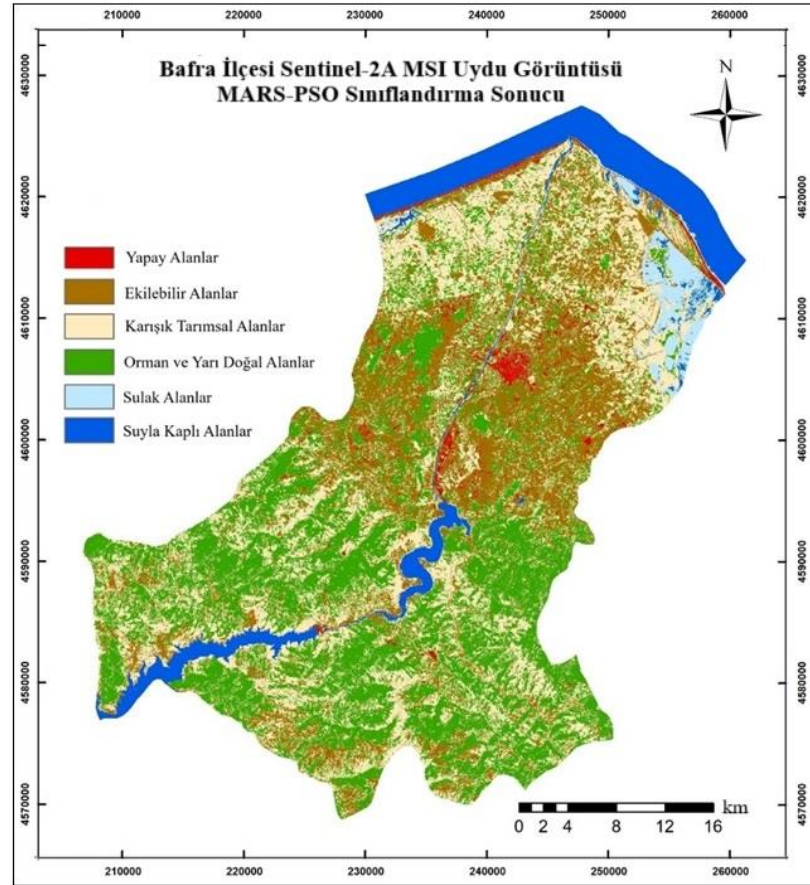


Şekil 4.7 Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS-PSO sınıflandırma sonucu

Tablo 4.7 Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS-PSO sınıflandırma sonucu hata matrisi

	Referans Verisi						
	Yapay Alanlar	Ekilebilir Alanlar	Karışık Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Sulak Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	83	8	13	10	0	0	114
Ekilebilir Alanlar	15	77	3	0	0	0	95
Karışık Tarımsal Alanlar	3	14	83	12	0	0	112
Orman ve Yarı Doğal Alanlar	10	6	8	86	8	0	118
Sulak Alanlar	0	0	0	5	90	4	99
Suyla Kaplı Alanlar	0	0	0	0	10	100	110
Toplam	111	105	107	113	108	104	648

Şekil 4.8 Sentinel-2A MSI uydu görüntüsüne ait MARS-PSO sınıflandırma sonucunu göstermektedir. Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS-PSO sınıflandırma sonucu hata matrisi Tablo 4.8’de verilmiştir.



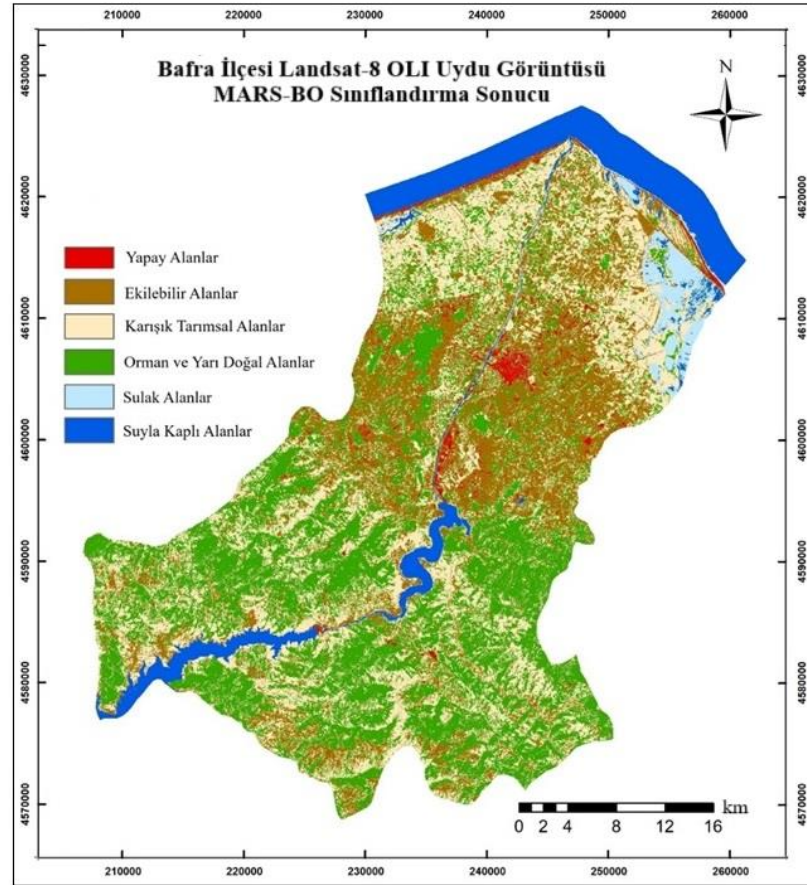
Şekil 4.8 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS-PSO sınıflandırma sonucu

Tablo 4.8 Bafra ilçesi Sentinel-2A uydu görüntüsü MARS-PSO sınıflandırma sonucu hata matrisi

	Referans Verisi						
	Yapay Alanlar	Ekilebilir Alanlar	Karışık Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Sulak Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	91	5	12	7	0	0	115
Ekilebilir Alanlar	12	79	3	0	0	0	94
Karışık Tarımsal Alanlar	0	13	85	10	0	0	108
Orman ve Yarı Doğal Alanlar	8	8	7	91	9	0	123
Sulak Alanlar	0	0	0	5	84	3	92
Suyla Kaplı Alanlar	0	0	0	0	15	101	116
Toplam	111	105	107	113	108	104	648

4.1.5. Test Alanı 1: Bafra İlçesi MARS-BO Hibrit Sınıflandırma Sonuçları

Şekil 4.9 Landsat-8 OLI uydu görüntüsüne ait MARS-BO sınıflandırma sonucunu göstermektedir. Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS-BO sınıflandırma sonucu hata matrisi Tablo 4.9’da verilmiştir.

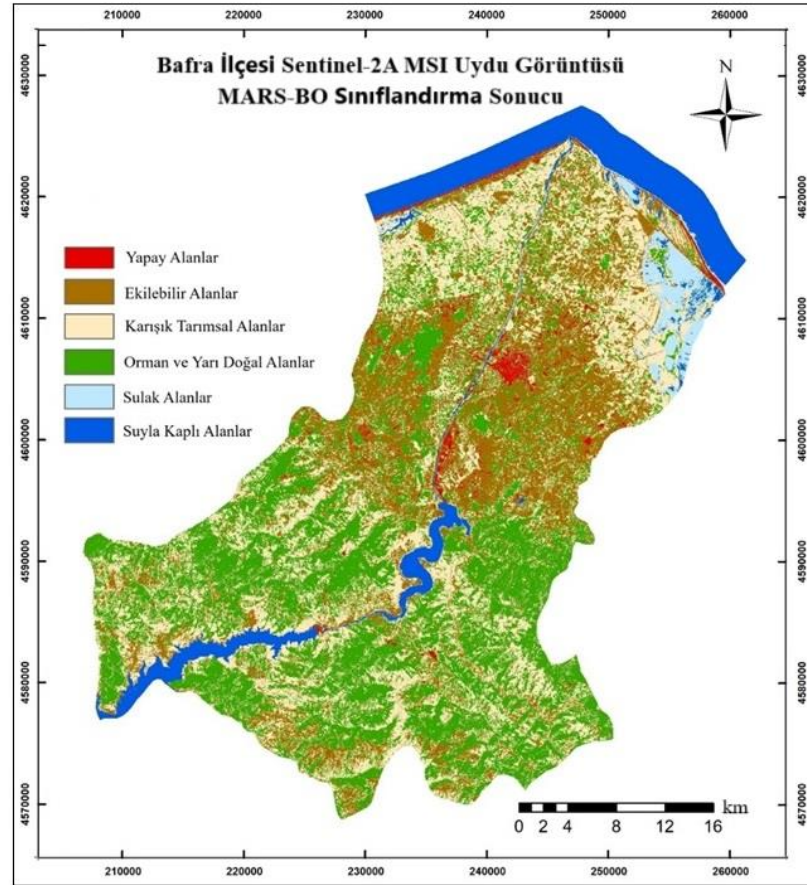


Şekil 4.9 Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS-BO sınıflandırma sonucu

Tablo 4.9 Bafra ilçesi Landsat-8 OLI uydu görüntüsü MARS-BO sınıflandırma sonucu hata matrisi

	Referans Verisi						
	Yapay Alanlar	Ekilebilir Alanlar	Karışık Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Sulak Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	85	8	13	10	0	0	116
Ekilebilir Alanlar	14	79	3	0	0	0	96
Karışık Tarımsal Alanlar	3	13	83	11	0	0	110
Orman ve Yarı Doğal Alanlar	9	5	8	87	7	0	116
Sulak Alanlar	0	0	0	5	91	4	100
Suyla Kaplı Alanlar	0	0	0	0	10	100	110
Toplam	111	105	107	113	108	104	648

Şekil 4.10 Sentinel-2A MSI uydu görüntüsüne ait MARS-BO sınıflandırma sonucunu göstermektedir. Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS-BO sınıflandırma sonucu hata matrisi Tablo 4.10’da verilmiştir.



Şekil 4.10 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS-BO sınıflandırma sonucu

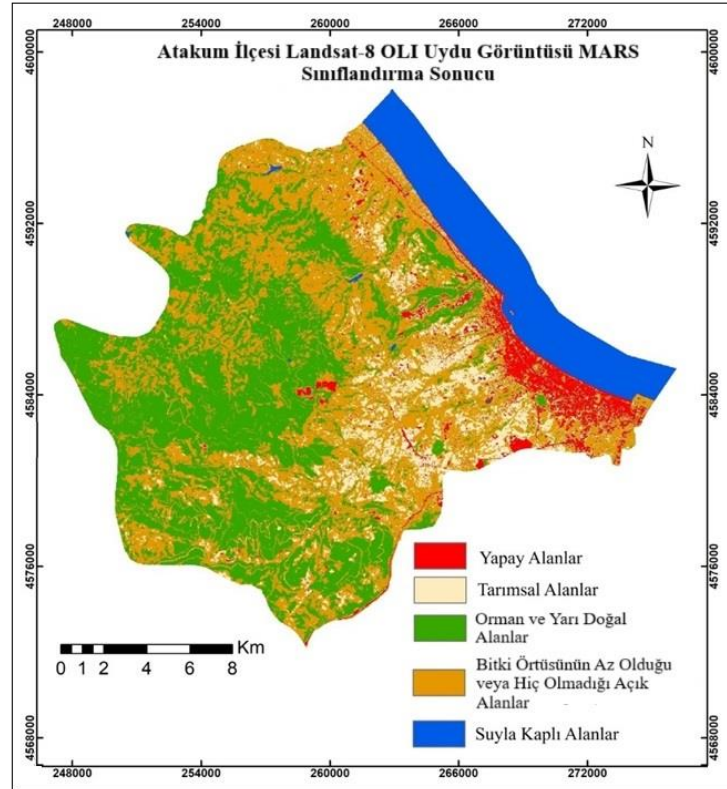
Tablo 4.10 Bafra ilçesi Sentinel-2A MSI uydu görüntüsü MARS-BO sınıflandırma sonucu hata matrisi

	Referans Verisi						
	Yapay Alanlar	Ekilebilir Alanlar	Karışık Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Sulak Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	91	4	11	5	0	0	111
Ekilebilir Alanlar	12	80	4	0	0	0	96
Karışık Tarımsal Alanlar	0	12	85	9	0	0	106
Orman ve Yarı Doğal Alanlar	8	9	7	94	9	0	127
Sulak Alanlar	0	0	0	5	84	3	92
Suyla Kaplı Alanlar	0	0	0	0	15	101	116
Toplam	111	105	107	113	108	104	648

4.2. Test Alanı 2: Atakum İlçesi Sınıflandırma Sonuçları

4.2.1. Test Alanı 2:Atakum İlçesi MARS Sınıflandırma Sonuçları

Şekil 4.11 Landsat-8 OLI uydu görüntüsüne ait MARS sınıflandırma sonucunu göstermektedir. Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS sınıflandırma sonucu hata matrisi Tablo 4.11’de verilmiştir.

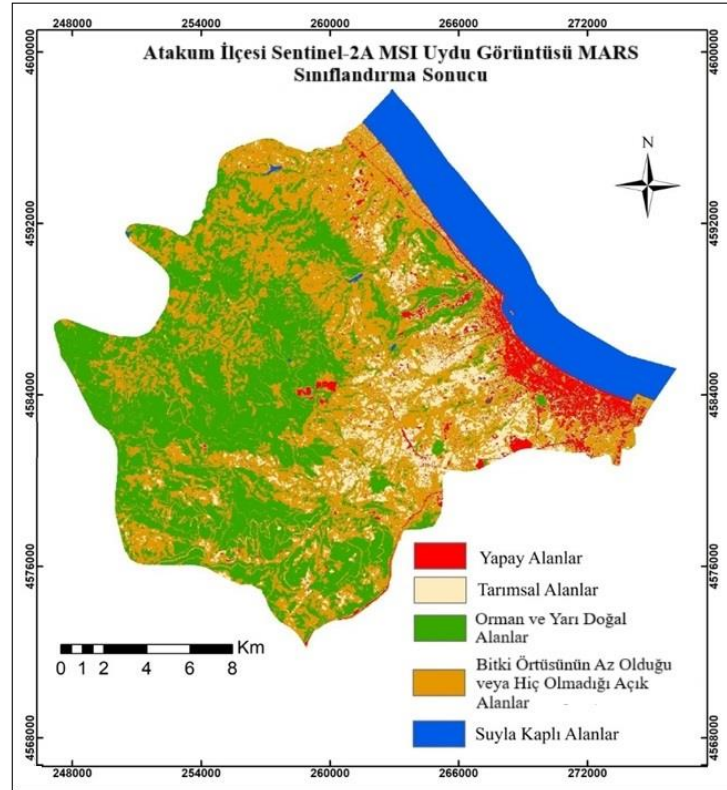


Şekil 4.11 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS sınıflandırma sonucu

Tablo 4.11 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS sınıflandırma sonucu hata matrisi

	Referans Verisi					
	Yapay Alanlar	Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	77	13	0	13	0	103
Tarımsal Alanlar	5	82	9	10	0	106
Orman ve Yarı Doğal Alanlar	4	7	84	3	0	98
Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	18	4	13	79	0	114
Suyla Kaplı Alanlar	0	0	0	0	101	101
Toplam	104	106	106	105	101	522

Şekil 4.12 Sentinel-2A MSI uydu görüntüsüne ait MARS sınıflandırma sonucunu göstermektedir. Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-BO sınıflandırma sonucu hata matrisi Tablo 4.12’de verilmiştir.



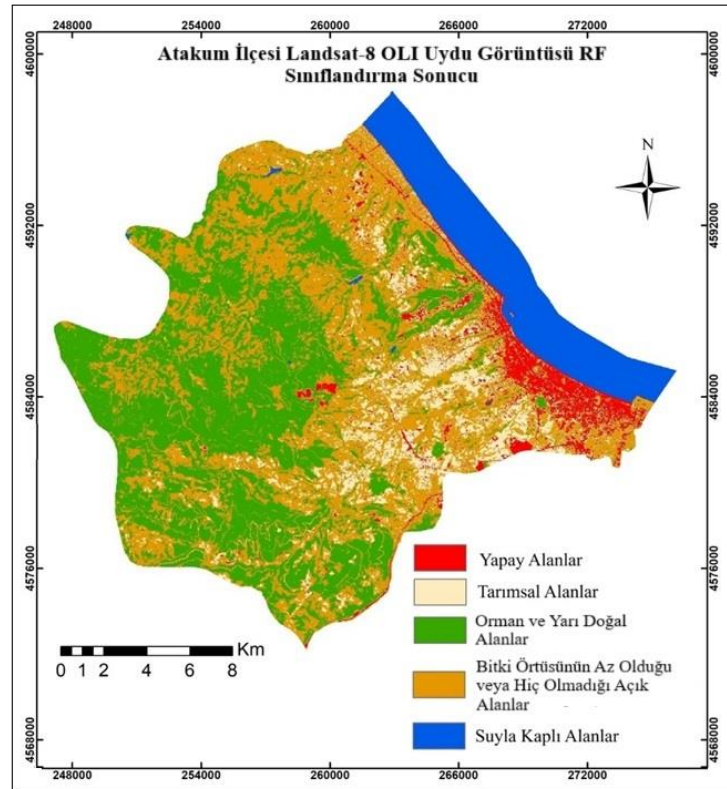
Şekil 4.12 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS sınıflandırma sonucu

Tablo 4.12 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS sınıflandırma sonucu hata matrisi

	Referans Verisi					
	Yapay Alanlar	Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	78	12	0	13	0	98
Tarımsal Alanlar	8	84	9	8	0	114
Orman ve Yarı Doğal Alanlar	4	6	85	3	0	98
Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	14	4	12	81	0	111
Suyla Kaplı Alanlar	0	0	0	0	101	101
Toplam	104	106	106	105	101	522

4.2.2. Test Alanı 2: Atakum İlçesi RF Sınıflandırma Sonuçları

Şekil 4.13 Landsat-8 OLI uydu görüntüsüne ait RF sınıflandırma sonucunu göstermektedir. Atakum ilçesi Landsat-8 OLI uydu görüntüsü RF sınıflandırma sonucu hata matrisi Tablo 4.13’te verilmiştir.

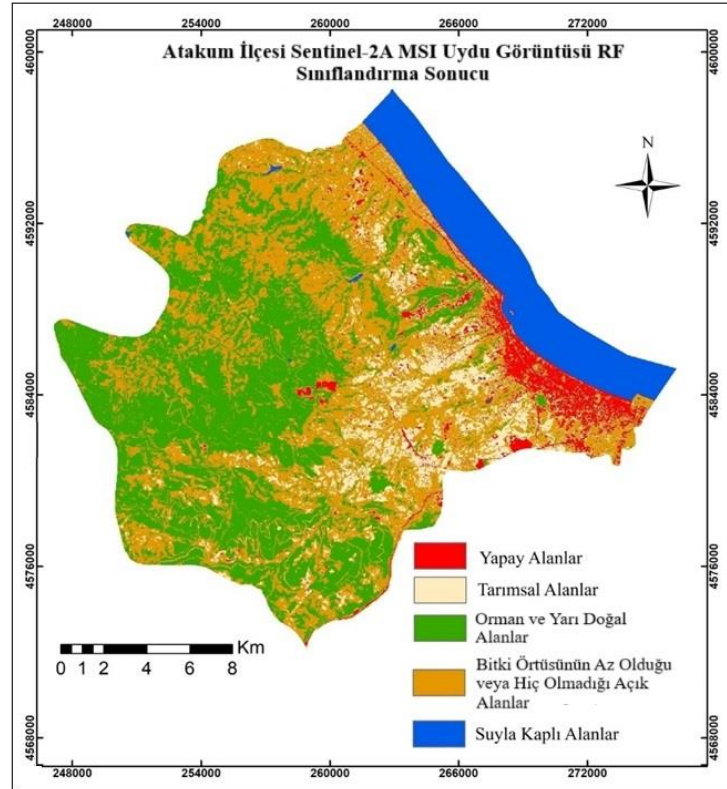


Şekil 4.13 Atakum ilçesi Landsat-8 OLI uydu görüntüsü RF sınıflandırma sonucu

Tablo 4.13 Atakum ilçesi Landsat-8 OLI uydu görüntüsü RF sınıflandırma sonucu hata matrisi

	Referans Verisi					
	Yapay Alanlar	Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	76	14	0	12	0	102
Tarımsal Alanlar	7	81	11	10	0	109
Orman ve Yarı Doğal Alanlar	4	6	82	2	0	94
Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	17	5	13	81	0	116
Suyla Kaplı Alanlar	0	0	0	0	101	101
Toplam	104	106	106	105	101	522

Şekil 4.14 Sentinel-2A MSI uydu görüntüsüne ait RF sınıflandırma sonucunu göstermektedir. Atakum ilçesi Sentinel-2A MSI uydu görüntüsü RF sınıflandırma sonucu hata matrisi Tablo 4.14’te verilmiştir.



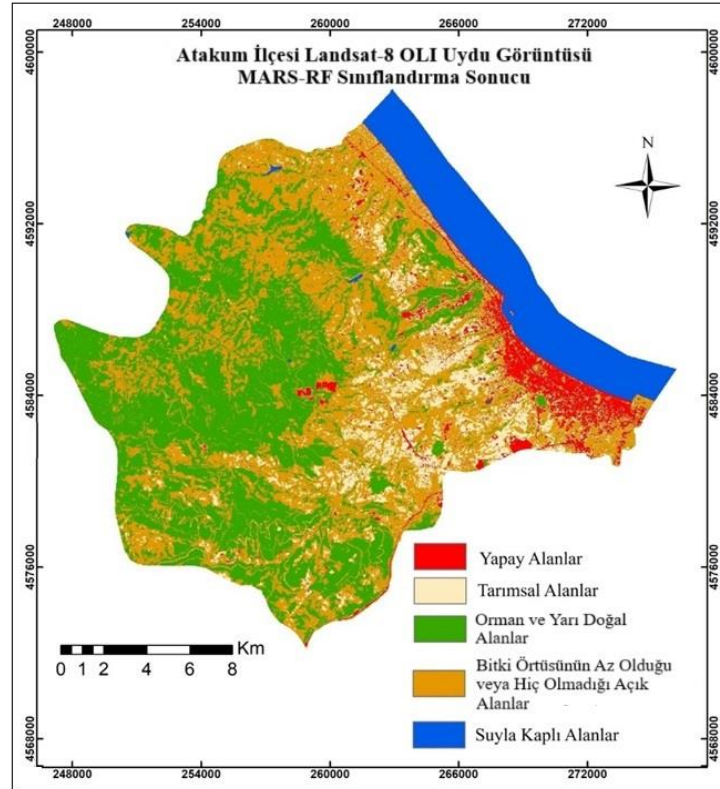
Şekil 4.14 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü RF sınıflandırma sonucu

Tablo 4.14 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü RF sınıflandırma sonucu hata matrisi

	Referans Verisi					
	Yapay Alanlar	Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	78	12	0	8	0	98
Tarımsal Alanlar	8	84	9	13	0	114
Orman ve Yarı Doğal Alanlar	4	6	85	3	0	98
Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	14	4	12	81	0	111
Suyla Kaplı Alanlar	0	0	0	0	101	101
Toplam	104	106	106	105	101	522

4.2.3. Test Alanı 2: Atakum İlçesi MARS-RF Sınıflandırma Sonuçları

Şekil 4.15 Landsat-8 OLI uydu görüntüsüne ait MARS-RF sınıflandırma sonucunu göstermektedir. Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS-RF sınıflandırma sonucu hata matrisi Tablo 4.15’te verilmiştir.

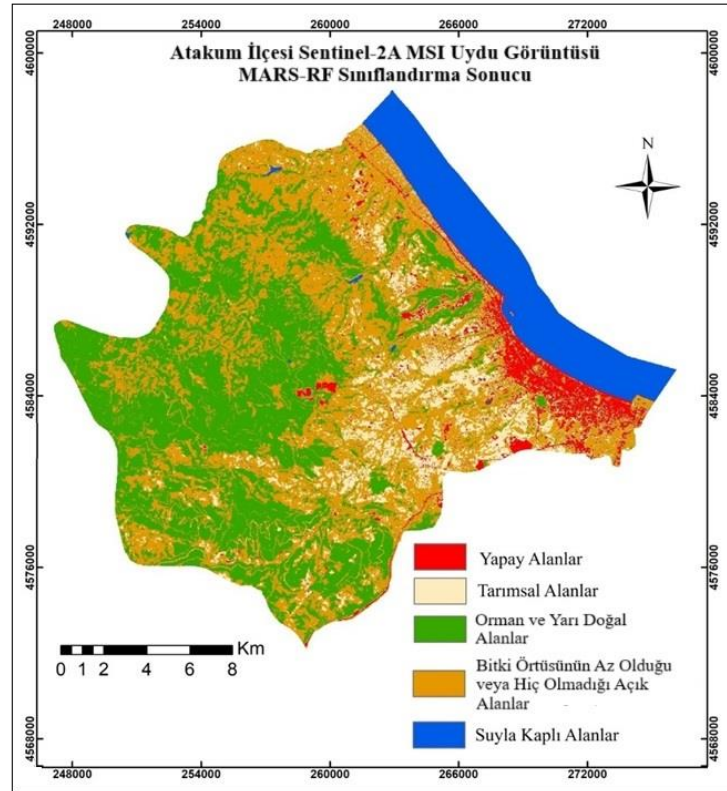


Şekil 4.15 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS-RF sınıflandırma sonucu

Tablo 4.15 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS-RF sınıflandırma sonucu hata matrisi

	Referans Verisi					
	Yapay Alanlar	Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	77	13	0	12	0	102
Tarımsal Alanlar	5	82	9	10	0	106
Orman ve Yarı Doğal Alanlar	4	7	84	2	0	97
Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	18	4	13	81	0	116
Suyla Kaplı Alanlar	0	0	0	0	101	101
Toplam	104	106	106	105	101	522

Şekil 4.16 Sentinel-2A MSI uydu görüntüsüne ait MARS-RF sınıflandırma sonucunu göstermektedir. Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-RF sınıflandırma sonucu hata matrisi Tablo 4.16’da verilmiştir.



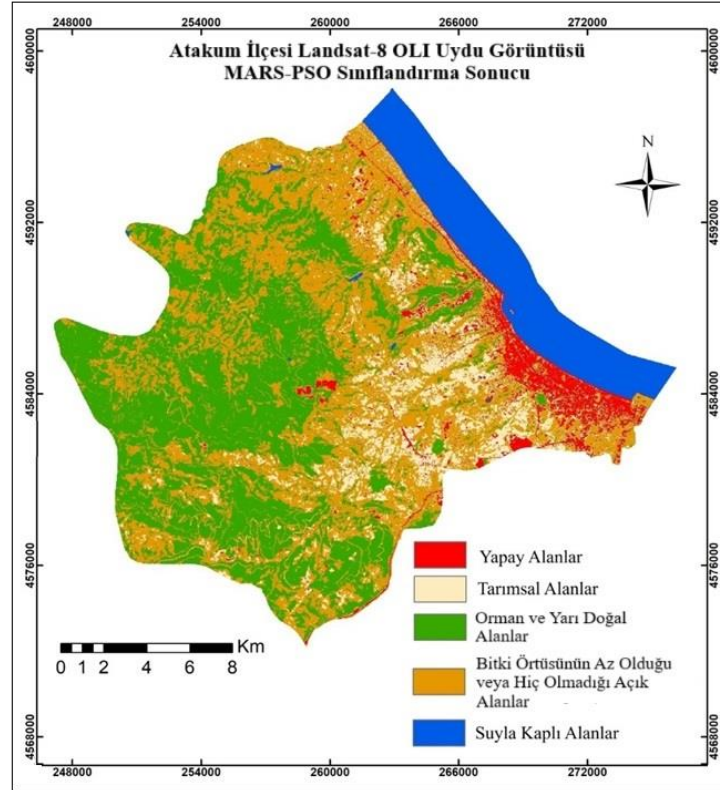
Şekil 4.16 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-RF sınıflandırma sonucu

Tablo 4.16 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-RF sınıflandırma sonucu hata matrisi

	Referans Verisi					
	Yapay Alanlar	Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	78	12	0	8	0	98
Tarımsal Alanlar	8	84	9	11	0	112
Orman ve Yarı Doğal Alanlar	4	6	85	3	0	98
Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	14	4	12	83	0	113
Suyla Kaplı Alanlar	0	0	0	0	101	101
Toplam	104	106	106	105	101	522

4.2.4. Test Alanı 2: Atakum İlçesi MARS-PSO Sınıflandırma Sonuçları

Şekil 4.17 Landsat-8 OLI uydu görüntüsüne ait MARS-RF sınıflandırma sonucunu göstermektedir. Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS-PSO sınıflandırma sonucu hata matrisi Tablo 4.17’de verilmiştir.

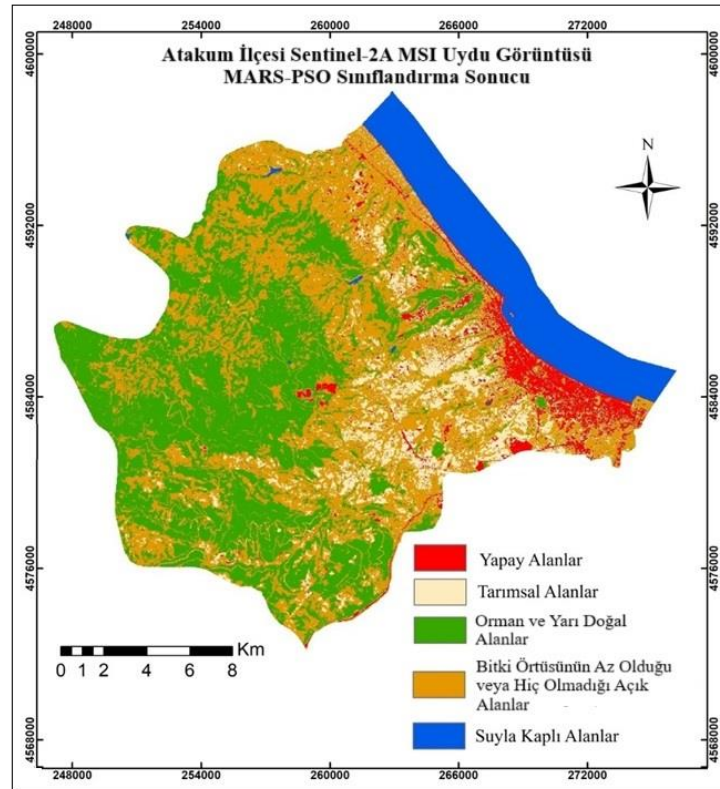


Şekil 4.17 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS-PSO sınıflandırma sonucu

Tablo 4.17 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS-PSO sınıflandırma sonucu hata matrisi

	Referans Verisi					
	Yapay Alanlar	Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	78	13	0	12	0	103
Tarımsal Alanlar	7	84	10	8	0	109
Orman ve Yarı Doğal Alanlar	3	9	86	3	0	101
Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	16	0	10	82	0	108
Suyla Kaplı Alanlar	0	0	0	0	101	101
Toplam	104	106	106	105	101	522

Şekil 4.18 Sentinel-2A MSI uydu görüntüsüne ait MARS-PSO sınıflandırma sonucunu göstermektedir. Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-PSO sınıflandırma sonucu hata matrisi Tablo 4.18’de verilmiştir.



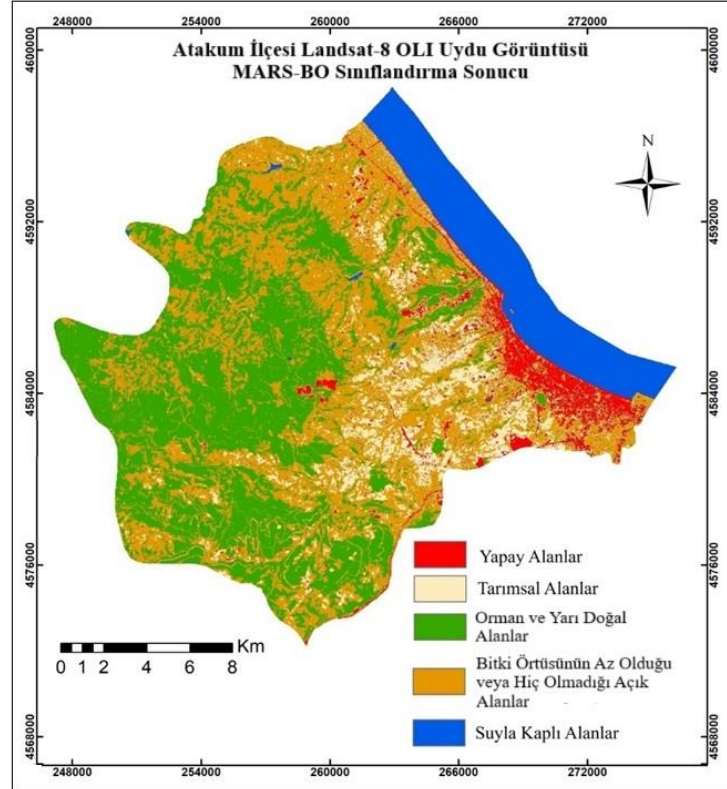
Şekil 4.18 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-PSO sınıflandırma sonucu

Tablo 4.18 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-PSO sınıflandırma sonucu hata matrisi

	Referans Verisi					
	Yapay Alanlar	Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	80	12	0	10	0	102
Tarımsal Alanlar	7	86	8	8	0	109
Orman ve Yarı Doğal Alanlar	4	7	87	3	0	101
Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	13	1	11	84	0	109
Suyla Kaplı Alanlar	0	0	0	0	101	101
Toplam	104	106	106	105	101	522

4.2.5. Test Alanı 2: Atakum İlçesi MARS-BO Sınıflandırma Sonuçları

Şekil 4.19 Landsat-8 OLI uydu görüntüsüne ait MARS-BO sınıflandırma sonucunu göstermektedir. Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS-BO sınıflandırma sonucu hata matrisi Tablo 4.19’da verilmiştir.

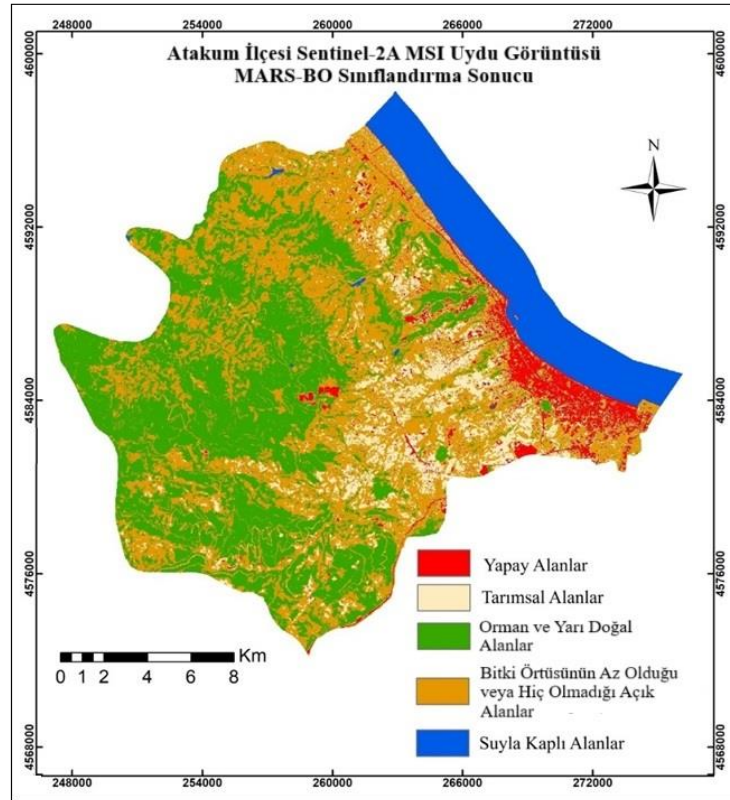


Şekil 4.19 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS-BO sınıflandırma sonucu

Tablo 4.19 Atakum ilçesi Landsat-8 OLI uydu görüntüsü MARS-BO sınıflandırma sonucu hata matrisi

	Referans Verisi					
	Yapay Alanlar	Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	78	12	0	11	0	101
Tarımsal Alanlar	7	85	11	8	0	111
Orman ve Yarı Doğal Alanlar	6	8	87	3	0	104
Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	13	1	8	83	0	105
Suyla Kaplı Alanlar	0	0	0	0	101	101
Toplam	104	106	106	105	101	522

Şekil 4.20 Sentinel-2A MSI uydu görüntüsüne ait MARS-BO sınıflandırma sonucunu göstermektedir. Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-BO sınıflandırma sonucu hata matrisi Tablo 4.20’de verilmiştir.



Şekil 4.20 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-BO sınıflandırma sonucu

Tablo 4.20 Atakum ilçesi Sentinel-2A MSI uydu görüntüsü MARS-BO sınıflandırma sonucu hata matrisi

	Referans Verisi					
	Yapay Alanlar	Tarımsal Alanlar	Orman ve Yarı Doğal Alanlar	Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	Suyla Kaplı Alanlar	Toplam
Yapay Alanlar	81	12	0	10	0	103
Tarımsal Alanlar	6	86	8	7	0	107
Orman ve Yarı Doğal Alanlar	4	8	89	3	0	104
Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	13	0	9	85	0	107
Suyla Kaplı Alanlar	0	0	0	0	101	101
Toplam	104	106	106	105	101	522

4.3. Sınıflandırma Sonuçlarının Karşılaştırılması

MARS ve RF tekil sınıflandırıcılar ve MARS-RF, MARS-PSO ve MARS-BO hibrit modellerin sınıflandırma sonuçlarının karşılaştırmalı olarak değerlendirilebilmesi için tüm yöntemlerin Landsat-8 OLI ve Sentinel-2A MSI uydu görüntüsündeki sonuçları aynı tabloda gösterilmiştir. Bu amaçla sınıf bazında değerlendirmeler için kullanıcı doğruluğu/kesinlik, üretici doğruluğu/duyarlılık ve F1 skor doğruluk ölçütlerini kapsayan ve sınıflandırmanın genel doğruluk değerlendirmesi için toplam doğruluk, kappa ve makro ortalama F1 skor değerlerini kapsayan ayrı tablolar oluşturulmuştur. Ayrıca sınıf sayıları ve sınıflar iki çalışma alanı için farklılık gösterdiği için ayrı ayrı ele alınmıştır.

Tablo 4.21’de Bafra İlçesi Landsat-8 OLI ve Sentinel-2A MSI veri setlerinin sonuçları kullanıcı doğruluğu/kesinlik, üretici doğruluğu/duyarlılık ve F1 skor olarak verilmiştir.

Tablo 4.21 Bafra İlçesi Kullanıcı Doğruluğu / Kesinlik, Üretici Doğruluğu / Duyarlılık ve F1 Skor sonuçları

		Landsat-8 OLI			Sentinel-2A MSI		
Sınıf	Yöntem	Kullanıcı Doğruluğu / Kesinlik	Üretici Doğruluğu / Duyarlılık	F1 Skor	Kullanıcı Doğruluğu / Kesinlik	Üretici Doğruluğu / Duyarlılık	F1 Skor
Yapay Alanlar	MARS	0,7069	0,7387	0,7225	0,7627	0,8108	0,7860
	RF	0,7018	0,7207	0,7111	0,7479	0,8018	0,7739
	MARS-RF	0,7069	0,7387	0,7225	0,7563	0,8108	0,7826
	MARS-PSO	0,7281	0,7477	0,7378	0,7913	0,8198	0,8053
	MARS-BO	0,7328	0,7658	0,7489	0,8198	0,8198	0,8198
Ekilebilir Alanlar	MARS	0,7813	0,7143	0,7463	0,8409	0,7048	0,7668
	RF	0,7835	0,7238	0,7525	0,8462	0,7333	0,7857
	MARS-RF	0,7917	0,7238	0,7562	0,8556	0,7333	0,7897
	MARS-PSO	0,8105	0,7333	0,7700	0,8404	0,7524	0,7940
	MARS-BO	0,8229	0,7524	0,7861	0,8333	0,7619	0,7960
Karışık Tarımsal Alanlar	MARS	0,7143	0,7477	0,7306	0,7373	0,8131	0,7733
	RF	0,7105	0,7570	0,7330	0,7436	0,8131	0,7768
	MARS-RF	0,7232	0,7570	0,7397	0,7500	0,8131	0,7803
	MARS-PSO	0,7411	0,7757	0,7580	0,7870	0,7944	0,7907
	MARS-BO	0,7545	0,7757	0,7650	0,8019	0,7944	0,7981
Orman ve Yarı Doğal Alanlar	MARS	0,7391	0,7522	0,7456	0,7632	0,7699	0,7665
	RF	0,7281	0,7345	0,7313	0,7636	0,7434	0,7534
	MARS-RF	0,7391	0,7522	0,7456	0,7699	0,7699	0,7699
	MARS-PSO	0,7288	0,7611	0,7446	0,7398	0,8053	0,7712
	MARS-BO	0,7500	0,7699	0,7598	0,7402	0,8319	0,7833
Sulak Alanlar	MARS	0,9082	0,8241	0,8641	0,8830	0,7685	0,8218
	RF	0,8889	0,8148	0,8502	0,8710	0,7500	0,8060
	MARS-RF	0,9082	0,8241	0,8641	0,8830	0,7685	0,8218
	MARS-PSO	0,9091	0,8333	0,8696	0,9130	0,7778	0,8400
	MARS-BO	0,9100	0,8426	0,8750	0,9130	0,7778	0,8400
Suyla Kaplı Alanlar	MARS	0,9009	0,9615	0,9302	0,8534	0,9519	0,9000
	RF	0,9000	0,9519	0,9252	0,8390	0,9519	0,8919
	MARS-RF	0,9009	0,9615	0,9302	0,8534	0,9519	0,9000
	MARS-PSO	0,9091	0,9615	0,9346	0,8707	0,9712	0,9182
	MARS-BO	0,9091	0,9615	0,9346	0,8707	0,9712	0,9182

Tablo 4.21 incelendiğinde Bafra İlçesi Yapay Alanlar sınıfı Landsat-8 OLI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,7069), üretici doğruluğu/duyarlılık (0,7387) ve F1 skor (0,7225) sonuçlarını vermiştir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,7018), üretici

doğruluğu/duyarlılık (0,7207) ve F1 skoru (0,7111) değerleri ile MARS algoritmasından biraz daha düşük performans göstermektedir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7069) ve üretici doğruluğu/duyarlılık (0,7387) ve F1 skor (0,7225) sonuçları tekil MARS algoritmasıyla aynı sonuçları vermiştir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7281), üretici doğruluğu/duyarlılık (0,7477) ve F1 skor (0,7378) değerleri ile MARS algoritmasından daha iyi sonuç vermektedir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7328), üretici doğruluğu/duyarlılık (0,7658) ve F1 skor (0,7489) ile en yüksek değerlere sahiptir.

Bafra İlçesi Yapay Alanlar sınıfı Sentinel-2A MSI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,7627), üretici doğruluğu/duyarlılık (0,8108) ve F1 skor (0,7860) sonuçlarını vermiştir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,7479), üretici doğruluğu/duyarlılık (0,8018) ve F1 skor (0,7739) değerleri ile MARS algoritmasından daha düşük sonuç vermiştir.

- MARS-RF hibrit modeli, üretici doğruluğu/duyarlılık (0,8108) ve F1 skoru (0,7826) ile MARS algoritması ile aynı ve yakın sonuç veriyor, ancak kullanıcı doğruluğu/kesinlik (0,7563) değeri MARS algoritmasında daha düşük bir değere sahiptir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7913), üretici doğruluğu/duyarlılık (0,8198) ve F1 skoru (0,8053) ile en iyi ikinci sonuçları vermektedir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8198), üretici doğruluğu/duyarlılık (0,8198) ve F1 skor (0,8198) ile en yüksek değerlere sahiptir.

Bafra İlçesi Ekilebilir Alanlar sınıfı Landsat-8 OLI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,7813), üretici doğruluğu/duyarlılık (0,7143) ve F1 skor (0,7463) sonucunu vermektedir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,7835), üretici doğruluğu/duyarlılık (0,7238) ve F1 skor (0,7525) ile MARS algoritmasına yakın bir performans göstermektedir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7917), üretici doğruluğu/duyarlılık (0,7238) ve F1 skor (0,7562) ile her iki tekil algoritmadan biraz daha iyi sonuç vermektedir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8105), üretici doğruluğu/duyarlılık (0,7333) ve F1 skor (0,7700) sonucunu vermektedir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8229), üretici doğruluğu/duyarlılık (0,7524) ve F1 skor (0,7861) ile en yüksek performansa sahip modeldir.

Bafra İlçesi Ekilebilir Alanlar sınıfı Sentinel-2A MSI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,8409), üretici doğruluğu/duyarlılık (0,7048) ve F1 skor (0,7668) sonucunu vermektedir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,8462), üretici doğruluğu/duyarlılık (0,7333) ve F1 skorunda (0,7857) ile MARS algoritmasından biraz daha yüksek performans göstermektedir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8556), üretici doğruluğu/duyarlılık (0,7333) ve F1 skor (0,7897) ile her iki tekil algoritmadan biraz daha iyi sonuç vermektedir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8404), üretici doğruluğu/duyarlılık (0,7524) ve F1 skor (0,7940) sonucunu vermektedir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8333), üretici doğruluğu/duyarlılık (0,7619) ve F1 skor (0,7960) ile Sentinel-2A MSI veri seti için en iyi performansa sahip modeldir.

Bafra İlçesi Karışık Tarımsal Alanlar sınıfı Landsat-8 OLI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,7143), üretici doğruluğu/duyarlılık (0,7477) ve F1 skor (0,7306) sonucunu vermektedir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,7105), üretici doğruluğu/duyarlılık (0,7570) ve F1 skor (0,7330) değerleri ile MARS algoritmasına yakın bir performans göstermektedir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7232), üretici doğruluğu/duyarlılık (0,7570) ve F1 skor (0,7397) ile tekil algoritmalarından daha iyi sonuç vermektedir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7411), üretici doğruluğu/duyarlılık (0,7757) ve F1 skor (0,7580) ile en iyi ikinci sonucu vermektedir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7545), üretici doğruluğu/duyarlılık (0,7757) ve F1 skor (0,7650) ile Landsat-8 OLI veri seti için en yüksek değerlere sahip modeldir.

Bafra İlçesi Karışık Tarımsal Alanlar sınıfı Sentinel-2A MSI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,7373), üretici doğruluğu/duyarlılık (0,8131) ve F1 skor (0,7733) sonucu vermektedir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,7436), üretici doğruluğu/duyarlılık (0,8131) ve F1 skor (0,7768) ile MARS algoritmasından biraz daha yüksek performans göstermektedir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7500), üretici doğruluğu/duyarlılık (0,8131) ve F1 skorunda (0,7803) ile tekil algoritmalara göre biraz daha iyi sonuçlar vermektedir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7870), üretici doğruluğu/duyarlılık (0,7944) ve F1 skor (0,7907) ile güçlü sonuçlar vermektedir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8019), üretici doğruluğu/duyarlılık (0,7944) ve F1 skor (0,7981) ile Sentinel-2A MSI veri seti için en yüksek değerlere sahip modeldir.

Bafra İlçesi Orman ve Yarı Doğal Alanlar sınıfı Landsat-8 OLI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,7391), üretici doğruluğu/duyarlılık (0,7522) ve F1 skor (0,7456) sonucunu vermektedir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,7281), üretici doğruluğu/duyarlılık (0,7345) ve F1 skor (0,7313) ile MARS algoritmasından daha düşük performans göstermektedir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7391), üretici doğruluğu/duyarlılık (0,7522) ve F1 skor (0,7456) ile MARS algoritmasıyla aynı sonuçları vermektedir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7288), üretici doğruluğu/duyarlılık (0,7611) ve F1 skor (0,7446) sonucunu vermektedir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7500), üretici doğruluğu/duyarlılık (0,7699) ve F1 skor (0,7598) ile Landsat-8 OLI veri seti için en yüksek değerlere sahip modeldir.

Bafra İlçesi Orman ve Yarı Doğal Alanlar sınıfı Sentinel-2A MSI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,7632), üretici doğruluğu/duyarlılık (0,7699) ve F1 skor (0,7665) sonucunu vermektedir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,7636), üretici doğruluğu/duyarlılık (0,7434) ve F1 skor (0,7534) ile MARS algoritmasına yakın sonuçlar vermektedir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7699), üretici doğruluğu/duyarlılık (0,7699) ve F1 skor (0,7699) ile MARS algoritmasından biraz daha iyi sonuçlar vermektedir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik 0,7398, üretici doğruluğu/duyarlılık 0,8053 ve F1 skor 0,7712 sonucunu vermektedir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7402), üretici doğruluğu/duyarlılık (0,8319) ve F1 skor (0,7833) ile Sentinel-2A MSI veri seti için en yüksek değerlere sahip modeldir.

Bafra İlçesi Sulak Alanlar sınıfı Landsat-8 OLI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,9082), üretici doğruluğu/duyarlılık (0,8241) ve F1 skor (0,8641) sonucunu vermektedir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,8889), üretici doğruluğu/duyarlılık (0,8148) ve F1 skor (0,8502) ile MARS algoritmasından daha düşük bir performans göstermektedir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,9082), üretici doğruluğu/duyarlılık (0,8241) ve F1 skor (0,8641) ile MARS algoritmasıyla aynı sonuçları vermektedir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,9091), üretici doğruluğu/duyarlılık (0,8333) ve F1 skor (0,8696) sonucunu vermektedir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,9100), üretici doğruluğu/duyarlılık (0,8426) ve F1 skor (0,8750) ile Landsat-8 OLI veri seti için en yüksek performansı gösteren modeldir.

Bafra İlçesi Sulak Alanlar sınıfı Sentinel-2A MSI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,8830), üretici doğruluğu/duyarlılık (0,7685) ve F1 skor (0,8218) sonucunu vermektedir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,8710), üretici doğruluğu/duyarlılık (0,7500) ve F1 skor (0,8060) ile MARS algoritmasından daha düşük sonuçlar vermiştir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8830), üretici doğruluğu/duyarlılık (0,7685) ve F1 skor (0,8218) ile MARS algoritmasıyla aynı sonuçları vermektedir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,9130), üretici doğruluğu/duyarlılık (0,7778) ve F1 skor (0,8400) ile güçlü sonuçlar vermektedir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,9130), üretici

doğruluğu/duyarlılık (0,7778) ve F1 skor (0,8400) ile Sentinel-2A MSI veri seti için en iyi sonuçları vermektedir.

Bafra İlçesi Suyula Kaplı Alanlar sınıfı Landsat-8 OLI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,9009), üretici doğruluğu/duyarlılık (0,9615) ve F1 skor (0,9302) sonucunu vermektedir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,9000), üretici doğruluğu/duyarlılık (0,9519) ve F1 skor (0,9252) ile MARS algoritmasından biraz daha düşük performans göstermiştir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,9009), üretici doğruluğu/duyarlılık (0,9615) ve F1 skor (0,9302) ile MARS algoritmasıyla aynı sonuçları vermiştir.

- MARS-PSO ve MARS-BO hibrit modelleri, kullanıcı doğruluğu/kesinlik (0,9091), üretici doğruluğu/duyarlılık (0,9615) ve F1 skor (0,9346) değerleri aynı ve en iyi performansı gösteren modellerdir.

Bafra İlçesi Suyula Kaplı Alanlar sınıfı Sentinel-2A MSI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,8534), üretici doğruluğu/duyarlılık (0,9519) ve F1 skor (0,9000) sonucunu vermektedir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,8390), üretici doğruluğu/duyarlılık (0,9519) ve F1 skor (0,8919) değerleri ile MARS algoritmasından daha düşük sonuçlar vermiştir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8534), üretici doğruluğu/duyarlılık (0,9519) ve F1 skor (0,9000) ile MARS algoritmasıyla aynı sonuçları vermektedir.

- MARS-PSO ve MARS-BO hibrit modelleri, kullanıcı doğruluğu/kesinlik (0,8707), üretici doğruluğu/duyarlılık (0,9712) ve F1 skor (0,9182) değerleri aynı ve en iyi performansı gösteren modellerdir.

Tekil algoritmaların performansları kendi içinde karşılaştırıldığında Yapay Alanlar, Sulak Alanlar ve Suyula Kaplı Alanlar için MARS algoritması en iyi doğruluk ölçütü değerlerini vermiştir. Ekilebilir Alanlar sınıfı için RF algoritması en iyi sonuçları vermektedir. Karışık Tarımsal Alanlar sınıfında ise Landsat-8 OLI veri setinde Kullanıcı Doğruluğu / Kesinlik ölçütü MARS algoritmasında daha yüksek iken Üretici Doğruluğu / Duyarlılık ve F1 Skor RF algoritmasında daha yüksektir. Sentinel-2A MSI veri setinde ise Üretici Doğruluğu / Duyarlılık hem RF hem de MARS algoritmasında eşit, Kullanıcı Doğruluğu / Kesinlik ve F1 Skor ölçütlerinde

ise RF algoritması daha yüksek sonuç vermiştir. Orman ve Yarı Doğal Alanlar sınıfında ise Sentinel-2A MSI veri setinde Kullanıcı Doğruluğu / Kesinlik RF algoritmasında, diğer tüm ölçütler ise MARS algoritmasında yüksek değerlere sahiptir.

MARS-RF hibrit modeli tekil MARS algoritmasından genel olarak daha iyi veya eşit sonuçlar vermiştir. Hem Landsat-8 OLI hem de Sentinel-2A MSI veri seti için en iyi performansı MARS-PSO ve MARS-BO algoritmaları göstermiştir. En iyi performansı gösteren MARS-BO hibrit modeli tekil MARS algoritması ile karşılaştırıldığında her iki veri seti ve üç ölçüt için ortalama yaklaşık olarak 0,02 artış göstermektedir.

Tablo 4.22 Test alanı 2 olan Atakum İlçesi sınıflandırma sonuçlarını 5 model için sınıf bazında vermektedir. Sınıflandırma sonuçları Landsat-8 OLI ve Sentinel-2A MSI veri setleri için ayrı sütunlarda gösterilmiştir.

Tablo 4.22 Atakum İlçesi Kullanıcı Doğruluğu / Kesinlik, Üretici Doğruluğu / Duyarlılık ve F1 Skor sonuçları

		Landsat-8 OLI			Sentinel-2A MSI		
		Kullanıcı Doğruluğu / Kesinlik	Üretici Doğruluğu / Duyarlılık	F1 Skor	Kullanıcı Doğruluğu / Kesinlik	Üretici Doğruluğu / Duyarlılık	F1 Skor
Yapay Alanlar	MARS	0,7476	0,7404	0,7440	0,7959	0,7500	0,7723
	RF	0,7451	0,7308	0,7379	0,7857	0,7404	0,7624
	MARS-RF	0,7549	0,7404	0,7476	0,7959	0,7500	0,7723
	MARS-PSO	0,7573	0,7500	0,7536	0,7843	0,7692	0,7767
	MARS-BO	0,7723	0,7500	0,7610	0,7864	0,7788	0,7826
Tarımsal Alanlar	MARS	0,7736	0,7736	0,7736	0,7368	0,7925	0,7636
	RF	0,7431	0,7642	0,7535	0,7477	0,7830	0,7650
	MARS-RF	0,7736	0,7736	0,7736	0,7500	0,7925	0,7706
	MARS-PSO	0,7706	0,7925	0,7814	0,7890	0,8113	0,8000
	MARS-BO	0,7658	0,8019	0,7834	0,8037	0,8113	0,8075
Orman ve Yarı Doğal Alanlar	MARS	0,8571	0,7925	0,8235	0,8673	0,8019	0,8333
	RF	0,8723	0,7736	0,8200	0,8469	0,7830	0,8137
	MARS-RF	0,8660	0,7925	0,8276	0,8673	0,8019	0,8333
	MARS-PSO	0,8515	0,8113	0,8309	0,8614	0,8208	0,8406
	MARS-BO	0,8365	0,8208	0,8286	0,8558	0,8396	0,8476
Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar	MARS	0,6930	0,7524	0,7215	0,7297	0,7714	0,7500
	RF	0,6983	0,7714	0,7330	0,7281	0,7905	0,7580
	MARS-RF	0,6983	0,7714	0,7330	0,7345	0,7905	0,7615
	MARS-PSO	0,7593	0,7810	0,7700	0,7706	0,8000	0,7850
	MARS-BO	0,7905	0,7905	0,7905	0,7944	0,8095	0,8019
Suyla Kaplı Alanlar	MARS	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
	RF	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
	MARS-RF	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
	MARS-PSO	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
	MARS-BO	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

Tablo 4.22 incelendiğinde Atakum İlçesi Yapay Alanlar sınıfı Landsat-8 OLI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,7476), üretici doğruluğu/duyarlılık (0,7404) ve F1 skor (0,7440) sonucunu vermiştir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,7451), üretici doğruluğu/duyarlılık (0,7308) ve F1 skor (0,7379) ile MARS algoritmasından biraz daha düşük performans sergilemektedir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7549), üretici doğruluğu/duyarlılık (0,7404) ve F1 skor (0,7476) ile MARS ve RF

algoritmalarından biraz daha iyi sonuç vermektedir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7573), üretici doğruluğu/duyarlılık (0,7500) ve F1 skor (0,7536) sonucunu vermiştir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7723), üretici doğruluğu/duyarlılık (0,7500) ve F1 skor (0,7610) ile en yüksek doğruluk ölçütü değerlerine sahip modeldir.

Atakum İlçesi Yapay Alanlar sınıfı Sentinel-2A MSI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,7959), üretici doğruluğu/duyarlılık (0,7500) ve F1 skor (0,7723) sonuç vermektedir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,7857), üretici doğruluğu/duyarlılık (0,7404) ve F1 skor (0,7624) ile MARS algoritmasından daha düşük sonuçlar vermektedir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7959), üretici doğruluğu/duyarlılık (0,7500) ve F1 skor (0,7723) ile MARS algoritmasıyla aynı sonuçları vermektedir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7843), üretici doğruluğu/duyarlılık (0,7692) ve F1 skor (0,7767) sonucunu vermiştir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7864), üretici doğruluğu/duyarlılık (0,7788) ve F1 skor (0,7826) ile en yüksek doğruluk ölçütü değerlerine sahip modeldir.

Atakum İlçesi Tarımsal Alanlar sınıfı Landsat-8 OLI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,7736), üretici doğruluğu/duyarlılık (0,7736) ve F1 skor (0,7736) sonucunu vermektedir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,7431), üretici doğruluğu/duyarlılık (0,7642) ve F1 skor (0,7535) ile MARS algoritmasından daha düşük performans göstermektedir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7736), üretici doğruluğu/duyarlılık (0,7736) ve F1 skor 0.7736 ile MARS algoritmasıyla aynı sonuçları vermektedir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7706), üretici doğruluğu/duyarlılık (0,7925) ve F1 skor (0,7814) ile en iyi ikinci sonucu vermektedir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7658), üretici

doğruluğu/duyarlılık (0,8019) ve F1 skor (0,7834) ile en yüksek doğruluk ölçütü değerlerine sahip modeldir.

Atakum İlçesi Tarımsal Alanlar sınıfı Sentinel-2A MSI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,7368), üretici doğruluğu/duyarlılık (0,7925) ve F1 skor (0,7636) sonucunu vermektedir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,7477), üretici doğruluğu/duyarlılık (0,7830) ve F1 skor (0,7650) ile MARS algoritmasından biraz daha iyi performans sergilemektedir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7500), üretici doğruluğu/duyarlılık (0,7925) ve F1 skor 0,7706 ile MARS algoritmasından biraz daha iyi sonuç vermektedir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7890), üretici doğruluğu/duyarlılık (0,8113) ve F1 skor (0,8000) ile güçlü sonuçlar vermektedir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8037), üretici doğruluğu/duyarlılık (0,8113) ve F1 skor (0,8075) ile en yüksek doğruluk ölçütü değerlerine sahip modeldir.

Atakum İlçesi Orman ve Yarı Doğal Alanlar sınıfı Landsat-8 OLI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,8571), üretici doğruluğu/duyarlılık (0,7925) ve F1 skor (0,8235) sonucunu vermektedir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,8723), üretici doğruluğu/duyarlılık (0,7736) ve F1 skor (0,8200) ile MARS algoritmasına yakın sonuçlar vermektedir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8660), üretici doğruluğu/duyarlılık (0,7925) ve F1 skor (0,8276) ile MARS algoritmasından daha iyi sonuç vermektedir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8515), üretici doğruluğu/duyarlılık (0,8113) ve F1 skor (0,8309) ile güçlü sonuçlar vermektedir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8365), üretici doğruluğu/duyarlılık (0,8208) ve F1 skor (0,8286) ile en yüksek doğruluk ölçütü değerlerine sahip modeldir.

Atakum İlçesi Orman ve Yarı Doğal Alanlar sınıfı Sentinel-2A MSI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,8673), üretici doğruluğu/duyarlılık (0,8019) ve F1 skor (0,8333) sonucunu vermektedir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,8469), üretici doğruluğu/duyarlılık (0,7830) ve F1 skor (0,8137) ile MARS algoritmasından daha düşük performans göstermektedir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8673), üretici doğruluğu/duyarlılık (0,8019) ve F1 skor (0,8333) ile MARS algoritmasıyla aynı sonuçları vermektedir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8614), üretici doğruluğu/duyarlılık (0,8208) ve F1 skor (0,8406) sonucunu vermiştir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,8558), üretici doğruluğu/duyarlılık (0,8396) ve F1 skor (0,8476) ile en yüksek doğruluk ölçütü değerlerine sahip modeldir.

Atakum İlçesi Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar sınıfı Landsat-8 OLI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,6930), üretici doğruluğu/duyarlılık (0,7524) ve F1 skor (0,7215) sonucunu vermiştir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,6983), üretici doğruluğu/duyarlılık (0,7714) ve F1 skor (0,7330) ile MARS algoritmasından biraz daha iyi sonuç vermiştir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,6983), üretici doğruluğu/duyarlılık (0,7714) ve F1 skor (0,7330) ile RF algoritmasıyla aynı sonuçları vermiştir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7593), üretici doğruluğu/duyarlılık (0,7810) ve F1 skor (0,7700) sonucunu vermiştir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7905), üretici doğruluğu/duyarlılık (0,7905) ve F1 skor (0,7905) ile en yüksek doğruluk ölçütü değerlerine sahip modeldir.

Atakum İlçesi Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar sınıfı Sentinel-2A MSI veri seti için;

- MARS algoritması, kullanıcı doğruluğu/kesinlik (0,7297), üretici doğruluğu/duyarlılık (0,7714) ve F1 skor (0,7500) sonucunu vermiştir.

- RF algoritması, kullanıcı doğruluğu/kesinlik (0,7281), üretici doğruluğu/duyarlılık (0,7905) ve F1 skor (0,7580) ile MARS algoritmasından daha iyi sonuçlar vermiştir.

- MARS-RF hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7345), üretici

doğruluğu/duyarlılık (0,7905) ve F1 skor (0,7615) ile tekil MARS ve RF algoritmalarından biraz daha iyi sonuç vermiştir.

- MARS-PSO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7706), , üretici doğruluğu/duyarlılık (0,8000) ve F1 skor (0,7850) ile güçlü sonuçlar vermiştir.

- MARS-BO hibrit modeli, kullanıcı doğruluğu/kesinlik (0,7944), üretici doğruluğu/duyarlılık (0,8095) ve F1 skor (0,8019) ile en yüksek doğruluk ölçütü değerlerine sahip modeldir.

Atakum İlçesi Suyla Kaplı Alanlar sınıfı Landsat-8 OLI ve Sentinel-2A MSI veri seti için;

- MARS, RF, MARS-RF, MARS-PSO ve MARS-BO modellerinde kullanıcı doğruluğu/kesinlik, üretici doğruluğu/duyarlılık ve F1 skor 1,0000 değerini vermiştir. Suyla Kaplı Alanlar sınıfının, sınıflandırma doğruluğu genellikle yüksek doğrulukta elde edilmektedir, bu çalışmada da bu sonuç elde edilmiştir.

Tekil algoritmalar kendi içinde karşılaştırıldığında Yapay Alanlar için MARS algoritması en iyi doğruluk ölçütü değerlerini vermiştir. Tarımsal Alanlar Landsat-8 OLI veri setinde MARS algoritması en yüksek sonuçları verirken, Sentinel-2A MSI veri seti için kullanıcı doğruluğu/kesinlik ve F1 skoru RF algoritmasında, üretici doğruluğu/duyarlılık ise MARS algoritmasında az farkla yüksek sonuç vermiştir. Orman ve Yarı Doğal Alanlar sınıfı Sentinel-2A MSI veri seti için MARS algoritması en yüksek doğruluk ölçüt değerlerine sahipken, Landsat-8 OLI veri seti için kullanıcı doğruluğu/kesinlik RF algoritmasında, üretici doğruluğu/duyarlılık ve F1 skor ise MARS algoritmasında daha yüksek sonuç vermiştir. Bitki Örtüsünün Az Olduğu veya Hiç Olmadığı Açık Alanlar Landsat-8 OLI veri seti için RF algoritması en yüksek değerlere sahipken, Sentinel-2A MSI veri setinde kullanıcı doğruluğu/kesinlik MARS algoritmasında, üretici doğruluğu/duyarlılık ve F1 skor değerleri RF algoritmasında daha yüksek sonuç vermiştir.

MARS-RF hibrit modeli tekil MARS algoritmasından genel olarak daha iyi veya eşit sonuçlar vermiştir. Hem Landsat-8 OLI hem de Sentinel-2A MSI veri seti için en iyi performansı MARS-PSO ve MARS-BO algoritmaları göstermiştir. En iyi performansı gösteren MARS-BO hibrit modeli tekil MARS algoritması ile karşılaştırıldığında her iki veri seti ve üç ölçüt için ortalama yaklaşık olarak 0,03 artış göstermektedir.

Tablo 4.23'te Bafra ve Atakum ilçeleri için Landsat-8 OLI ve Sentinel-2A MSI veri setlerinin sınıflandırılmasında Toplam Doğruluk, Kappa ve makro ortalama ile

hesaplanan F1 skor ölçütleri verilmiştir.

Tablo 4.23 Bafra ve Atakum ilçeleri için Landsat-8 OLI ve Sentinel-2A MSI veri setlerinin sınıflandırılmasında Toplam Doğruluk, Kappa ve makro ortalama ile hesaplanan F1 skor sonuçları

		Landsat-8 OLI			Sentinel-2A MSI		
		Toplam Doğruluk	Kappa	F1 Skor	Toplam Doğruluk	Kappa	F1 Skor
Bafra	MARS	%78,86	0,7542	0,7899	%80,25	0,7629	0,8024
	RF	%78,24	0,7388	0,7839	%79,78	0,7574	0,7979
	MARS-RF	%79,17	0,7499	0,7931	%80,71	0,7685	0,8074
	MARS-PSO	%80,09	0,7610	0,8024	%81,94	0,7833	0,8199
	MARS-BO	%81,02	0,7792	0,8116	%82,56	0,7907	0,8259
Atakum	MARS	%81,03	0,7629	0,8125	%82,18	0,7773	0,8239
	RF	%80,65	0,7581	0,8089	%81,80	0,7725	0,8198
	MARS-RF	%81,42	0,7677	0,8164	%82,57	0,7821	0,8275
	MARS-PSO	%82,57	0,7821	0,8272	%83,91	0,7988	0,8405
	MARS-BO	%83,14	0,7892	0,8327	%84,67	0,8084	0,8479

Tablo 4.23 incelendiğinde Bafra ilçesi Landsat-8 OLI veri seti için;

- MARS (%78,86) ve RF (%78,24) birbirine oldukça yakın toplam doğruluk değerine sahiptir. Ancak hibrit yöntemler daha iyi performans göstermektedir; MARS-RF (%79,17) ve MARS-PSO (%80,09) hibrit modelleri MARS ve RF algoritmalarına göre daha yüksek toplam doğruluk sağlamıştır. En yüksek toplam doğruluk ise MARS-BO (%81.02) hibrit modeliyle elde edilmiştir.

- MARS-BO hibrit modeli 0,7792 kappa değeriyle en yüksek sonucu vermiştir. MARS-PSO (0,7610) ve MARS-RF (0,7499) hibrit modelleri de RF (0,7388) ve MARS (0,7542) algoritmalarına göre daha iyi sonuçlar sunmuştur.

- MARS-BO hibrit modeli 0,8116 ile en yüksek F1 skorunu vermektedir. MARS-PSO (0,8024) ve MARS-RF (0,7931) hibrit modelleri de MARS (0,7899) ve RF (0,7839) algoritmalarına göre daha yüksek sonuçlar vermiştir.

Bafra ilçesi Sentinel-2 MSI veri seti için;

- En yüksek toplam doğruluk MARS-BO (%82,56) hibrit modeli ile elde edilmiştir. MARS-PSO (%81,94) ve MARS-RF (%80,71) hibrit modelleri de MARS (%80,25) ve RF (%79,78) algoritmalarına göre daha yüksek toplam doğruluk sağlamaktadır.

- MARS-BO (0,7907) hibrit modeli en yüksek kappa değerine sahiptir. MARS-PSO (0,7833) ve MARS-RF (0,7685) hibrit modelleri MARS (0,7629) ve RF (0,7574) yöntemlerine göre daha iyi sonuçlar vermektedir.

- MARS-BO hibrit modeli 0,8259 ile en yüksek F1 skoruna sahiptir. MARS-PSO (0,8199) ve MARS-RF (0,8074) hibrit modelleri MARS (0,8024) ve RF (0,7979) yöntemlerine göre daha iyi performans göstermektedir.

Atakum ilçesi Landsat-8 OLI veri seti için;

- MARS-BO hibrit modeli en yüksek toplam doğruluğu (%83,14) vermektedir. MARS-PSO (%82,57) ve MARS-RF (%81,42) hibrit modelleri MARS (%81,03) ve RF (%80,65) algoritmalarına göre daha iyi sonuçlar vermiştir.

- En yüksek kappa değeri yine MARS-BO (0,7892) hibrit modeli ile elde edilmiştir. MARS-PSO (0,7821) ve MARS-RF (0,7677) hibrit modelleri ile MARS (0,7629) ve RF (0,7581) yöntemlerine göre daha iyi kappa değerleri gözlemlenmiştir.

- En yüksek F1 skoru MARS-BO (0,8327) hibrit modeli ile elde edilmiştir. MARS-PSO (0,8272) ve MARS-RF (0,8164) hibrit modelleri MARS (0,8125) ve RF (0,8089) yöntemlerine göre daha iyi performans gözlemlenmiştir.

Atakum ilçesi Sentinel-2 MSI veri seti için;

- En yüksek toplam doğruluk MARS-BO (%84,67) hibrit modeli ile elde edilmiştir. MARS-PSO (%83,91) ve MARS-RF (%82,57) hibrit modelleri MARS (%82,18) ve RF (%81,80) yöntemlerine göre daha iyi sonuçlar sunmuştur.

- En yüksek kappa değeri MARS-BO (0,8084) hibrit modeli ile elde edilmiştir. MARS-PSO (0,7988) ve MARS-RF (0,7821) hibrit modelleri yine MARS (0,7773) ve RF (0,7725) yöntemlerine göre daha yüksek performans göstermektedir.

- En yüksek F1 skoru yine MARS-BO (0,8479) hibrit modelinde elde edilmiştir. MARS-PSO (0,8405) ve MARS-RF (0,8275) hibrit modelleri MARS (0,8239) ve RF (0,8198) yöntemlerine göre daha iyi F1 skor değerine sahiptir.

Bu çalışmanın temel sınıflandırıcısı olan MARS algoritması ve bu algoritmaya dayalı hibrit yaklaşımlar karşılaştırıldığında, hibrit yaklaşımların sınıflandırma doğruluğu daha yüksektir. Tablo 4.23 incelendiğinde, üç doğruluk ölçütünün de hibrit modellerin tekil sınıflandırıcılardan daha iyi performansa sahip olduklarını görülmektedir. Hibrit modeller kendi içinde kıyaslandığında ise optimizasyon algoritmalarının kullanıldığı hibrit modeller (MARS-PSO ve MARS-BO), MARS algoritmasının RF algoritması ile hibrit yaklaşımını sunan MARS-RF modelinden daha iyi performansa sahiptir. MARS-BO ve MARS-PSO modelleri arasındaki fark çok az olmakla birlikte MARS-BO modeli daha iyi bir performansa sahiptir.

4.4. MARS Algoritmasına Dayalı Hibrit Sınıflandırma Modellerinin Katkıları ve Kısıtları

Gelişen teknoloji ile arazi üzerinde yapılan maliyetli ve zaman alan çalışmalar yerine, uydu görüntülerinin sınıflandırılması daha yüksek doğrulukta, çok daha ekonomik ve hızlı şekilde gerçekleştirilebilir. Ayrıca, farklı sınıflandırıcılar aynı çalışma alanı için farklı sonuçlar verebilir fakat tek bir sınıflandırıcı tüm sınıflar için en iyi performansı gösteremez. Başka bir deyişle, her bir sınıflandırma yönteminin kendine özgü güçlü yanları ve kısıtları vardır, farklı sınıflandırıcıların güçlü yanları kullanılarak sınıflandırma doğruluğu artırılabilir. Bu amaç doğrultusunda, sınıflandırma doğruluğunun artırılması ve daha güvenilir sonuçların elde edebilmesi için yeni yaklaşımlar araştırılmaktadır. Bu çalışmalar, sınıflandırma performansını arttırmak için hibrit yaklaşımlara odaklanmıştır.

MARS algoritması tıp, biyoloji, ekonomi alanında kullanılmasına rağmen görüntü sınıflandırma alanında yeni bir yöntemdir. Yeryüzündeki doğal ve yapay nesneler, doğrusal olmayan davranış sergiler ve bağımlı değişkenleri etkileyen birçok bağımsız değişken bulunmaktadır, bu nedenle parametrik olmayan regresyon modelleri yeryüzü araştırmalarına adapte edildiğinde daha başarılı sonuçlar vermektedir. MARS, bağımsız değişkenlerin belirli aralıkları için farklı ilişkiler ortaya koyarak, daha esnek modeller oluşturulmasını sağlar. MARS algoritmasının bu güçlü yönlerini başka algoritmaların avantajları ile birleştirerek, sınıflandırma doğruluğunu artırılabilir. Bu bağlamda, bu tez çalışmasında MARS algoritmasına dayalı hibrit yöntemler geliştirilmiş ve MARS-RF, MARS-PSO ve MARS-BO olarak adlandırılan üç yaklaşım sunulmuştur.

MARS ve RF algoritmalarının sonuçları incelendiğinde üç doğruluk ölçütü de birbirine çok yakın olarak elde edilmiştir. Literatürde arazi örtüsü/kullanımı analizi açısından MARS ve RF algoritmalarının karşılaştırılması olmasa da, tıp, çevre alanlarında regresyon analizi karşılaştırmalarında çoğunlukla MARS algoritması daha iyi performans sağlamıştır, ama bazılarında ise RF daha iyi performans göstermiştir (Yao et al., 2013; Zabihi et al., 2016; Hamidi et al., 2018).

Bu çalışmada önerilen ilk hibrit model olan MARS-RF modeli, hibrit sınıflandırma tekniklerinden paralel kombinasyon yöntemi ile oluşturulmuş olup modelin performansı toplam doğruluk, kappa istatistiği ve F1 skor ile değerlendirilmiştir. Hem iki çalışma alanında hem de iki veri setinde MARS-RF modelinin daha yüksek doğruluğa sahip olduğu görülmüştür. Her ne kadar MARS ve

RF tekil sınıflandırma sonuçları toplam doğruluk, kappa ve F1 skor değerleri bakımından birbirine yakın olsa da, belirlenen karar kuralına göre her bir sınıfın performansı ayrı olarak dikkate alındığı için MARS-RF hibrit modeli sınıflandırma doğruluğunu arttırmış, farklı sınıflandırıcıların hibrit kullanımına dair örnek oluşturmuştur.

Ayrıca görüntü sınıflandırma analizinde iyi performansa sahip MARS algoritması, optimizasyon yöntemleri ile birlikte kullanımı da performansı artırıcı bir etki yapmaktadır. Birçok yüksek performansı sınıflandırıcı gibi MARS algoritmasında da kullanıcının belirlediği parametreler bulunmaktadır, algorithmada bu parametrelerin seçimi performansı doğrudan etkilemektedir. Bu nedenle parametre seçimi önemli bir unsurdur ve optimizasyon yöntemleri en uygun parametre seçiminde kullanılan ve iyi sonuçlar veren yöntemlerdir. MARS algoritması parametreleri kullanıcı tarafından belirlenmektedir ve zamanla literatürde bu parametreler için alt ve üst sınırlar oluşmaya başlamıştır. Fakat temel fonksiyon sayısı ve ceza parametresi gibi değerlerin kullanıcı tarafından belirlenmesi her zaman en uygun sonucu vermemektedir. Bu nedenle, bu parametrelerin belirlenmesinde ek bir yöntem olarak optimizasyon yöntemi kullanılmasıyla algorithmadaki bu kısıtın giderilmesi hedeflenmiştir. Çalışma kullanılan PSO ve BO tekniklerine parametreler için alt ve üst sınırlar verilerek, iterasyonlarla en uygun değerlerin bulunması sağlanmıştır. Bu tez çalışmasında, MARS algoritmasının PSO ve BO optimizasyon teknikleri ile sırasıyla MARS-PSO ve MARS-BO modelleri geliştirilmiştir. Parametrelerin optimizasyonuna yönelik iki algoritmanın performansı karşılaştırıldığında MARS-BO modelinin daha iyi sonuç verdiği görülmektedir. Her iki algoritmanın da performansı iyi olmakla beraber literatür incelendiğinde BO algoritması PSO algoritmasına göre daha iyi performansa sahip olduğu birçok çalışma görülmektedir (Snoek et al., 2012; Frazier, 2018; Tani and Veelken, 2024).

MARS algoritması, bağımsız değişkenlerin karmaşık ilişkilerini modelleme kabiliyeti ve doğrusal olmayan yapıları yakalama yeteneği ile öne çıkar. Bu özellikleri sayesinde, uydu görüntülerinin sınıflandırılmasında yüksek doğruluk oranları elde edilir. MARS algoritmasının avantajları ve kısıtları, uydu görüntülerinin hibrit sınıflandırılmasında yeni modeller oluşturma potansiyelini göstermektedir.

Gelecekteki çalışmalarda, MARS algoritmasına dayalı hibrit modellerin, uydu görüntü analizinde sıklıkla kullanılan bir araç haline gelmesi beklenmektedir. Bu da araştırmacılara daha kapsamlı ve güvenilir analizler yapma olanağı sunacaktır.

MARS tabanlı hibrit modellerin uydu görüntü analizinde yaygınlaşması, özellikle değişkenlik gösteren ve karmaşık yapıları olan coğrafi alanlarda daha doğru sonuçlar elde edilmesine katkı sağlayacaktır. Böylece, arazi kullanım değişikliklerinin izlenmesi, çevresel etkilerin değerlendirilmesi ve doğal kaynak yönetimi gibi alanlarda daha etkili çözümler sunulacaktır.



5. SONUÇ VE ÖNERİLER

Mekânsal analizlerde uydu görüntüsü sınıflandırması, yaygın ve etkili bir yöntem olarak öne çıkmaktadır. Mevcut durumun tespiti, değişimlerin izlenmesi, planlama ve modelleme gibi pek çok araştırmada sınıflandırma analizi önemli bir rol oynamaktadır. Günümüzde uydu görüntülerinin sınıflandırma analizine yönelik araştırmalar hız kesmeden devam etmekte ve bu alandaki gelişim sürekli olarak ilerlemektedir. Son yıllarda, daha güvenilir ve doğru sonuçlar elde etmek amacıyla farklı algoritmaların güçlü yönlerini birleştiren hibrit yaklaşımlar geliştirilmektedir. Hibrit yaklaşımlar, çeşitli sınıflandırıcıların birlikte kullanılması veya optimizasyon tekniklerinin sınıflandırıcılara entegre edilmesi yoluyla elde edilir. Bu tür hibrit yöntemler, farklı sınıflandırma algoritmalarının avantajlarını bir araya getirerek sınıflandırma doğruluğunu arttırmakta ve mekânsal analizlerde daha kapsamlı sonuçlara ulaşılmasına katkı sağlamaktadır.

Hibrit modeller, farklı algoritmaların güçlü yönlerini birleştirerek sınıflandırma doğruluğunu artırır ve karmaşık arazi örtüsü türlerinin daha hassas bir şekilde tanımlanmasını sağlar. Bu amaç doğrultusunda, MARS algoritması hibrit model geliştirmede yüksek potansiyele sahip bir yöntem olarak öne çıkmaktadır. Bu çalışmada hibrit yöntemlerin oluşturulmasında iki temel yaklaşım benimsenmiştir: farklı sınıflandırma yöntemlerinin birleştirilmesi ve optimizasyon tekniklerinin kullanılması. Bu amaçla MARS algoritmasına dayalı olarak MARS-RF, MARS-PSO ve MARS-BO olmak üzere üç hibrit model geliştirilmiştir.

MARS-RF hibrit modeli, MARS algoritması ile RF algoritmasının paralel kombinasyon tekniğiyle entegre edilmesiyle oluşturulmuştur. MARS ve RF algoritmalarının avantajları ve kısıtları incelenerek hibrit kullanımda doğruluğun artırılabilceği ön görülmüştür. MARS algoritmasında görülen aşırı uyum problemi, RF algoritmasında bulunmamaktadır. Aynı zamanda RF algoritmasının karşılaştığı uygun değişken seçimi zorluğu, MARS algoritmasının değişkenler arasındaki etkileşimi dikkate alarak en uygun değişkenleri seçme özelliği ile aşılabilmektedir. Bu avantaj ve kısıtlar doğrultusunda, MARS ve RF algoritmalarının sınıflandırma sonuçları birleştirilerek kısıtların etkisini en aza indirmek ve daha yüksek doğruluk sağlamak amacıyla bir hibrit model geliştirilmiştir.

MARS algoritması incelendiğinde en uygun temel fonksiyon sayısı ve ceza parametresi belirleme kısıtı bulunmaktadır. Bu parametreler optimizasyon

yöntemleriyle belirlenebilir. Optimizasyon, amaç fonksiyonunu maksimum veya minimum değere ulaştıran parametreleri belirlemek için bir dizi işlemin uygulanmasıdır. Bu çalışmada iki farklı optimizasyon tekniği kullanılarak hibrit model oluşturulmuştur. MARS algoritmasının parametrelerini optimize etmek amacıyla PSO ve BO teknikleri kullanılarak sırasıyla MARS-PSO ve MARS-BO hibrit sınıflandırma modelleri geliştirilmiştir.

MARS, RF, MARS-RF, MARS-PSO ve MARS-BO olmak üzere beş modelin performanslarını değerlendirmek amacıyla iki farklı çalışma alanı ve iki ayrı uydu veri seti kullanılmıştır. Sınıflandırma işlemleri, RStudio ve MATLAB yazılımları aracılığıyla gerçekleştirilmiştir. Modellerin performansını detaylı bir şekilde karşılaştırmak için, her sınıf özelinde kullanıcı doğruluğu/kesinlik, üretici doğruluğu/duyarlılık ve F1 skoru gibi ölçütler dikkate alınmıştır. Genel model karşılaştırmasında ise toplam doğruluk, kappa değeri ve makro ortalama F1 skoru gibi ölçütler kullanılmıştır.

Sonuçlar, hibrit modellerin tek başına kullanılan algoritmalara göre daha yüksek doğruluk sağladığını göstermektedir. Özellikle optimizasyon teknikleriyle desteklenen hibrit modellerin, sınıflandırma doğruluğunu önemli ölçüde artırdığı gözlemlenmiştir. Bu durum, algoritma parametrelerinin optimize edilmesinin sınıflandırma sonuçları üzerindeki olumlu etkisini vurgulamaktadır. Özellikle MARS-BO modeli, toplam doğruluk, kappa değeri ve F1 skoru gibi kritik ölçütlerde en yüksek performansa ulaşarak, MARS algoritmasının optimizasyon teknikleri ile entegre edilmesinin uydu görüntüleri sınıflandırma analizlerinde güçlü ve etkili bir yöntem olduğunu ortaya koymuştur. Bu sonuçlar, MARS algoritmasının hibrit modellerle desteklenmesinin, uydu görüntülerinin sınıflandırmasında daha doğru ve güvenilir sonuçlar elde etmede potansiyel bir çözüm sunduğunu kanıtlamaktadır.

Hibrit yaklaşımlar daha doğru ve kapsamlı bilgi elde etmek için araştırmacılara ve uygulayıcılara önemli bir araç sağlanmıştır. Hibrit model geliştirme hem farklı sınıflandırıcı kullanılabilmesi hem de birden fazla sınıflandırıcı/optimizasyon kullanılabilmesi açısından esneklik sunmaktadır. Hibrit modellerin farklı coğrafi alanlar ve arazi örtüsü/kullanımı çeşitlilikleri üzerinde test edilmesi, yüksek doğruluk sağlama potansiyelini desteklemektedir. Bu bağlamda ilerleyen çalışmalarda MARS algoritmasına dayalı farklı sınıflandırıcı ve/veya optimizasyon teknikleri kullanılarak sınıflandırma doğruluğu artırma amacıyla çalışmalar devam etmelidir.

KAYNAKLAR

- Akyol M. (2011). *Yaşam Çözümlemesine Yeni Bir Yaklaşım: MARS*, Doktora Tezi, Ankara Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Ana Bilim Dalı, 100, Ankara.
- Alade, I. O., Abd Rahman, M. A., & Saleh, T.A., (2019). Predicting the specific heat capacity of alumina/ethylene glycol nanofluids using support vector regression model optimized with Bayesian algorithm. *Solar Energy*, 183, 74-82.
- Alam, M. N. (2004). Particle swarm optimization: algorithm and its codes in MATLAB. *Control systems engineering*. Wiley, Hoboken.
- Alpar, R. (1997) . *Uygulamalı çok değişkenli istatistiksel yöntemlere giriş*. Nobel Yayıncılık.
- Anderson, J. R. Hardy, E. E., Roach, J. T., & Witmer, R. E. (1976). *A land use and land cover classification system for use with remote sensor data* (Vol. 964). Geological Survey Professional Paper, Washington.
- Anonim (2019). Biodiversity and Climate Change Virtual Laboratory (BCCVL).
<https://support.bccvl.org.au/support/solutions/articles/6000118097-multivariate-adaptive-regression-splines> (Erişim tarihi: 04.09.2022)
- Anonim (2021). R-Project. <https://www.r-project.org/about.html> (Erişim tarihi: 08.03.2021)
- Archer, K. J., & Kimes, R. V. (2008). Empirical characterization of random forest variable importance measures. *Computational statistics & data analysis*, 52(4), 2249-2260.
- Assal, T. J., Anderson, P. J., & Sibold, J. (2015). Mapping forest functional type in a forest-shrubland ecotone using SPOT imagery and predictive habitat distribution modelling. *Remote Sensing Letters*, 6(10), 755-764.
- Aziz, R., Verma, C. K., & Srivastava, N. (2017). Dimension reduction methods for microarray data: a review. *AIMS Bioengineering*, 4(1), 179-197.
- Banos, O., Damas, M., Pomares, H., Rojas, F., Delgado-Marquez, B., & Valenzuela, O. (2013). Human activity recognition based on a sensor weighting hierarchical classifier. *Soft Computing*, 17, 333-343.
- Bazaraa, M. S., Jarvis, J. J., & Sherali, H. D. (1990). *Linear programming and network flows*. John Wiley & Sons.
- Belgiu, M., & Dragut, L. (2016). Random forest in remote sensing: a review of applications and future directions. *ISPRS journal of photogrammetry and remote sensing*, 114, 24-31.
- Benemaran, R. S., & Esmaeili-Falak, M. (2020). Optimization of cost and mechanical properties of concrete with admixtures using MARS and PSO. *Computers and concrete*, 26(4), 309-316.
- Bergstra, J., Bardenet, R., Bengio, Y., & Kegl, B. (2011). Algorithms for hyper-parameter optimization. *Advances in neural information processing systems*, 24.
- Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS journal of photogrammetry and remote sensing*, 65(1), 2-16.
- Breiman, L. (1996). Bagging predictors. *Machine learning*, 24, 123-140.
- Breiman, L. (2001). Random forests. *Machine learning* 45, 15-32.
- Breiman, L., Friedman, J. H., Olshen, R. A. & Stone C. J. (1984). *Classification and regression trees*. Wadsworth and Brooks/Cole, New York.
- Brochu, E., Cora, V. M. & Freitas, N. (2010). A tutorial on Bayesian optimization of expensive cost functions with application to active user modeling and hierarchical

- reinforcement learning. *arXiv preprint arXiv:1012.2599*.
- Brodley, C. E. (1995). Recursive automatic bias selection for classifier construction. *Machine Learning* 20, 63-94.
- Bo, Y., & Wang, J., (2005). Combining multiple classifiers for thematic classification of remotely sensed data. *National Remote Sensing Bulletin*. 9 (5), 555-564.
- Boyd, S., & Parrilo, P. (2009). Introduction to convex optimization. MIT Open Course Ware.
- Bui, D. T., Bui, Q. T., Nguyen, Q. P., Pradhan, B., Nampak, H., & Trinh, P. T. (2017). A hybrid artificial intelligence approach using GIS-based neural-fuzzy inference system and particle swarm optimization for forest fire susceptibility modeling at a tropical area. *Agricultural and forest meteorology*, 233, 32-44.
- Bui, D. T., Hoang, N. D., Pham, T. D., Ngo, P. T. T., Hoa, P. V., Minh, N. Q., Tran, X. T. & Samui, P. (2019). A new intelligence approach based on GIS-based multivariate adaptive regression splines and metaheuristic optimization for predicting flash flood susceptible areas at high-frequency tropical typhoon area. *Journal of Hydrology*, 575, 314-326.
- Büttner, G. (2014). "CORINE land cover and land cover change products," in *Land use and land cover mapping in europe: practices & trends remote sensing and digital image processing*. Editors I. Manakos, & M. Braun. Springer Netherlands, 55-74. doi: <http://dx.doi.org/10.1007/978-94>.
- Chi, M., Kun, Q., Benediktsson, J. A., & Feng, R. (2009). Ensemble classification algorithm for hyperspectral remote sensing data. *IEEE Geoscience remote Sensing* 6(4), 762-766.
- Chinnaswamy, A., & Srinivasan, R. (2016). Hybrid feature selection using correlation coefficient and particle swarm optimization on microarray gene expression data. In *Innovations in Bio-Inspired Computing and Applications: Proceedings of the 6th International Conference on Innovations in Bio-Inspired Computing and Applications (IBICA 2015)* Kochi, India.
- Congalton, R. G., & Green, K. (2009). Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. 2nd Edition, Lewis Publishers, Boca Raton.
- Craven, P., & Wahba, G. (1978). Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, 31, 377-403.
- Çatal Reis, H., & Bayram, B., (2016). Parçacık Sürü Optimizasyonu (PSO) Yöntemi İle Kıyı Çizgisi Çıkarılması: Erçek Gölü Örneği. VI. *Uzaktan Algılama ve Coğrafi Bilgi Sistemleri Sempozyumu*, 5-7 Ekim, Adana, Türkiye.
- Dribault, Y., Chokmani, K., & Bernier, M. (2012). Monitoring seasonal hydrological dynamics of minerotrophic peatlands using multi-date GEOEYE-1 very high resolution imagery and object-based classification. *Remote Sensing*, 4, 1887-1912. doi: <http://dx.doi.org/10.3390/rs4071887>.
- Du, P. Xia, J. Zhang, W. Tan, K. Liu, Y. & Liu, S. (2012). Multiple classifier system for remote sensing image classification: a review. *Sensors*, 12(4), 4764-4792.
- Eberhart, R., & Kennedy, J. (1995). "A new optimizer using particle swarm theory". In: MHS 95., Proceedings of the Sixth International Symposium on Micro Machine and Human Science, 39-43.
- Ekici, O. (2009). İstatistikte Bayesyen ve Klasik yaklaşımın kavramsal farklılıkları. *Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 12(21), 89-101.
- Erar M. A. (çeviri editörü) (2013). Doğrusal Regresyon Analizine Giriş – Introduction to

Linear Regression Analysis Montgomery D. C., Peck E. A., Vining G. G., Nobel Akademik Yayıncılık.

- Fallah, S. N. Deo, R. C. Shojafar, M. Conti, M., & Shamshirband, S. (2018). Computational intelligence approaches for energy load forecasting in smart energy management grids: State of the art, future challenges, and research directions, *Energies* 11, 596.
- Fauvel, M., Chanussot, J., & Benediktsson, J. A. (2006). Decision fusion for the classification of urban remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 44(10), 2828-2838.
- Feurer, M., & Hutter, F. (2019). Hyperparameter Optimization. *Automated Machine Learning: Methods, Systems, Challenges*, 3-33.
- Frazier, P. I. (2018). A tutorial on Bayesian optimization. *arXiv preprint arXiv:1807.02811*.
- Foody, G. M., (2009). Classification accuracy comparison: hypothesis tests and the use of confidence intervals in evaluations of difference, equivalence and non-inferiority. *Remote Sensing of Environment*, 113(8), 1658-1663.
- Foody, G. M., Boyd, D. S., & Sanchez-Hernandez, C., (2007). Mapping a specific class with an ensemble of classifiers. *International Journal of Remote Sensing*, 28(8), 1733-1746.
- Friedl, M. A., & Brodley, C. E., (1997). Decision tree classification of land cover from remotely sensed data, *Remote Sensing of Environment*, 61, 399-409.
- Friedman, J. H. (1991). Multivariate adaptive regression splines. *The Annals of Statistics*, 19, 1-67.
- Friedman J. H. (1993). Fast MARS. Technical Report No:110.
- Frohn, R. C., Reif, M., Lane, C., & Autrey, B. (2009). Satellite remote sensing of isolated wetlands using object-oriented classification of Landsat-7 data. *Wetlands* 29, 931. doi: <https://doi.org/10.1672/08-194.1>.
- Gad, A. G. (2022). Particle swarm optimization algorithm and its applications: a systematic review. *Arch. Comput. Methods Eng.* 29, 2531-2561. doi: <https://doi.org/10.1007/s11831-021-09694-4>.
- Gamon, M., Aue, A., Corston-Oliver, S., & Ringger, E. (2005). Pulse: Mining customer opinions from free text. In *Advances in Intelligent Data Analysis VI: 6th International symposium on intelligent data analysis, IDA 2005, Madrid, Spain, September 8-10, 2005. Proceedings* 6, 121-132. Springer Berlin Heidelberg.
- Garai, S., Paul, R. K., Yeasin, M., & Paul, A. K. (2024). CEEMDAN-Based Hybrid Machine Learning Models for Time Series Forecasting Using MARS Algorithm and PSO-Optimization. *Neural Processing Letters*, 56(2), 92.
- Garnett, R., Osborne, M. A., & Roberts S. J. (2010). Bayesian optimization for sensor set selection. In *9th ACM/IEEE International conference on information processing in sensor networks*, April 12-16, Stockholm, Sweden.
- Geneletti, D., & Gorte, B. G. H. (2003). A method for object-oriented land cover classification combining Landsat TM data and aerial photographs, *International Journal of Remote Sensing*, 24(6), 1273-1286.
- Goepp, V., Bouaziz, O., & Nuel G., (2018). Spline Regression with Automatic Knot Selection. *Computational Statistics & Data Analysis*, 202, 108043.
- Gonzalez, J., Longworth, J., James, D. C., & Lawrence N. D. (2015). Bayesian optimization for synthetic gene design. *arXiv preprint arXiv:1505.01627*.
- Gregorio, A. D., & Jansen, L. J. M. (1998). A new concept for a land cover classification

- system. The *Land* 2(1), 55–65.
- Grullon, Y. R., Alhaddad, B. & Cladera, J. R. (2009). The analysis accuracy assessment of CORINE land cover in the Iberian coast. In *Remote Sensing for Environmental Monitoring, GIS Applications, and Geology IX*, 7478, 443-450.
- Hai, T., Theruvil Sayed, B., Majdi, A., Zhou, J., Sagban, R., Band, S. S., & Mosavi, A. (2023). An integrated GIS-based multivariate adaptive regression splines-cat swarm optimization for improving the accuracy of wildfire susceptibility mapping. *Geocarto International*, 2167005.
- Hamidi, O., Tapak, L., Abbasi, H., & Maryanaji, Z. (2018). Application of random forest time series, support vector regression and multivariate adaptive regression splines models in prediction of snowfall (a case study of Alvand in the middle Zagros, Iran). *Theoretical and Applied Climatology*, 134(3), 769-776.
- Hansen L., & Salamon P.. (1990). Neural network ensembles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(10):993–1001.
- Hanteh, M., Malek, H., & Kheyroddin, A. (2023). Proposing the combined MARS–PSO and ELM–PSO models for estimating the compressive strength of circular concrete columns wrapped with FRP sheets. *Soft Computing*, 27(21), 15937-15953.
- Hasanoğlu, M. S. (2019). An Advanced evolutionary programming method for mechanical system design: Feasibility enhanced particle swarm optimization. Middle East Technical University, Institute of Science and Technology, Department of Mechanical Engineering, 233, Ankara.
- Hashem, S., (1994). Optimal linear combinations of neural networks. *Neural Networks*, 10(4), 599–614.
- Hastie, T. Tibshirani, R., & Friedman, J. H., (2001). *The elements of statistical learning*, Springer series in Statistics Springer, Berlin.
- Hekimoğlu, B., Altindeğer, M., & Demirbaş, A. R. (2007). Samsun İlinin Fiziki Durumu Ve Avantajları. Samsun Valiliği Tarım İl Müdürlüğü.
- Herold, M., Scepan, J., & Clarke, K. C. (2002). The use of remote sensing and landscape metrics to descibe structures andchanges in urban land uses. *Environment and Planning* 34(8): 1443–145.
- Herzog, R. (2022). Introduction to optimization. Lecture Notes Interdisciplinary Center for Scientific Computing, Heidelberg University.
- Harita Genel Müdürlüğü – HGM. İl ve ilçe yüzölçümleri: <https://www.harita.gov.tr/il-ve-ilce-yuzolcumleri> (Erişim tarihi: 29.10.2024).
- Hillier, F., & Lieberman, G. (2005). *Introduction to operations research*, McGraw-Hill, Boston.
- Hoang, N. D., Nguyen, Q. L., & Tien Bui, D. (2018). Image processing–based classification of asphalt pavement cracks using support vector machine optimized by artificial bee colony. *Journal of Computing in Civil Engineering*, 32(5), 04018037.
- Hoffman, M., Brochu, E., & De Freitas, N. (2011). Portfolio Allocation for Bayesian Optimization. In *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence*, 14-17 July, Barcelona, Spain.
- Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. *International Journal of remote sensing*, 23(4), 725-749.
- Inti, S., & Tandon, V. (2017). Integration of data envelopment analysis-based preference aggregation method and α particle swarm optimization technique into group decision

- model. *Journal of Computing in Civil Engineering*, 31(1), 04016043.
- Jaison, B., Chilambuchelvan, A., & Mohamed Junaid, K. A. (2015). Hybrid classification techniques for microarray data. *National Academy Science Letters*, 38, 415-419.
- Japkowicz, N., & Shah, M. (2011). *Evaluating learning algorithms: a classification perspective*. Cambridge University Press.
- Jiang, M., Luo, Y., & Yang, S. (2007). Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm. *Information Processing Letters*, 102(1), 8-16. doi: <https://doi.org/10.1016/j.ipl.2006.10.005>
- Jones, D. R. (2001). A taxonomy of global optimization methods based on response surfaces. *Journal of Global Optimization*, 21(4):345–383.
- Kalkan, K., & Maktav, D. (2010). Nesne tabanlı ve piksel tabanlı sınıflandırma yöntemlerinin karşılaştırılması (IKONOS örneği). *III. Uzaktan Algılama ve Coğrafi Bilgi Sistemleri Sempozyumu*, 12-15 Eylül, Gebze, Kocaeli.
- Kennedy, J., & Eberhart, R., (1995). Particle swarm optimization. *IEEE International Conference on Neural Networks*, 27 November - 1 December, Perth, WA, Australia. doi: <https://doi.org/10.1109/ICNN.1995.488968>.
- Kennedy J., & Eberhart R. (2007). Defining a Standard for Particle Swarm Optimization. 2007 IEEE Swarm Intelligence Symposium, 1-5 April, Honolulu, HI, USA.
- Koç E. K., (2012). An algorithm for the forward step of adaptive regression splines via mapping approach, Doctoral Thesis, Middle East Technical University, Institute of Natural and Applied Sciences, Department of Statistics, 148, Ankara.
- Krogh A., & Vedelsby J., (1995). Neural network ensembles, cross validation, and active learning. *Advances in Neural Information Processing Systems*, 231-238.
- Kuter S., (2014). Atmospheric correction and image classification on MODIS images by nonparametric regression splines. Doctoral Thesis, Middle East Technical University, Institute of Science and Technology, Department of Geodetic and Geographic Information Technologies, 199, Ankara.
- Kuter, S., Weber, G. W. & Karasozen, B., (2015). Modern Uygulamalı Matematik ve Modern Sürekli Optimizasyon Bağlamında Parametrik Olmayan Regresyon Eğrilerinin Bilim, Mühendislik ve Finans Alanlarındaki Güncel Uygulamaları. *Akademik Bilişim Konferansı*, 4-6 Şubat, Eskişehir, Türkiye.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models*. McGraw-hill.
- Le, C. A., Huynh, V. N., Shimazu, A., & Nakamori, Y. (2007). Combining classifiers for word sense disambiguation based on Dempster–Shafer theory and OWA operators. *Data & Knowledge Engineering*, 63(2), 381-396.
- Leathwick, J. R., Elith, J., & Hastie, T. (2006). Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. *Ecological Modelling*, 199(2), 188–196.
- Lee, T. H., Ullah, A., & Wang, R. (2020). Bootstrap aggregating and random forest. *Macroeconomic forecasting in the era of big data: Theory and practice*, 389-429.
- Li, M., Zang, S., Zhang, B., Li S., & Wu C., (2014). A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information, *European Journal of Remote Sensing*, 47(1), 389-411.
- Liao, K. W., Hoang, N. D., & Gitomarsano, J. (2018). A probabilistic safety evaluation framework for multi-hazard assessment in a bridge using SO-MARS learning model. *Journal of Civil Engineering*, 22, 903-915.

- Liberti, L., & Maculan N. (2006). *Global optimization: from theory to implementation*. Springer Science & Business Media.
- Lillesand, T. M., & Kiefer, R. W. (1994). *Remote sensing and image interpretation*. 3rd ed.; John Wiley and Sons: New York, USA.
- Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2015) *Remote Sensing and Image Interpretation*. 7th Edition, Wiley, New York, USA.
- Lipton, Z. C., Elkan, C., & Naryanaswamy, B. (2014). Optimal thresholding of classifiers to maximize F1 measure. *Machine Learning and Knowledge Discovery in Databases: European Conference*, 15-19 September, Nancy, France.
- Liu, J. G., & Mason, P. J. (2009). *Essential image processing and GIS for remote sensing*. UK: John Wiley & Sons Ltd.
- Liu, Y., & Ge, Z. (2018). Weighted random forests for fault classification in industrial processes with hierarchical clustering model selection. *Journal of Process Control*, 64, 62-70.
- Lu, D. & Weng, Q., (2007) A survey of image classification methods and techniques for improving classification performance, *International Journal of Remote Sensing*, 28(5), 823-870.
- Luenberger, D. G., & Ye, Y. (2008) *Linear and nonlinear programming*, fourth edition. Springer, Amsterdam.
- Lv, F., Han, M., & Qiu, T. (2017). Remote sensing image classification based on ensemble extreme learning machine with stacked autoencoder. *IEEE Access*, 5, 9021-9031.
- Mantovani, R. G., Horváth, T., Cerri, R., Vanschoren, J., & De Carvalho, A. C. (2016). Hyper-parameter tuning of a decision tree induction algorithm. *5th Brazilian Conference on Intelligent Systems (BRACIS)*, 9-12 October, Recife, Brazil.
- Mather, P. M., & Koch, M. (2011) *Computer processing of remotely-sensed images: an introduction*. Fourth Edition, Chichester: Wiley-Blackwell.
- Mather, P., & Tso, B. (2016). *Classification Methods for Remotely Sensed Data*, CRC Press.
- MATLAB Help Center, (2024). https://www.mathworks.com/help/pdf_doc/matlab/index.html?s_tid=mwa_osa_a
- Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., & Gascon, F. (2017). Sen2Cor for sentinel-2. *Image and signal processing for remote sensing XXIII* 10427, 37-48.
- Mockus, J., Tiesis, V. & Zilinskas, A. (1978). The application of Bayesian methods for seeking the extremum. *Optimization techniques IFIP technical conference: Novosibirsk*, 1-7 July, Heidelberg, Berlin.
- Nedd, R., Light, K., Owens, M., James, N., Johnson, E., & Anandhi, A. (2021). A synthesis of land use/land cover studies: Definitions, classification systems, meta-studies, challenges and knowledge gaps on a global landscape. *The Land*, 10(9), 994.
- Nguyen, H., Bui, X. N., Tran, Q. H., Nguyen, H. A., Nguyen, D. A., Hoa, L. T. T., & Le, Q. T. (2022). Prediction of ground vibration intensity in mine blasting using the novel hybrid MARS-PSO-MLP model. *Engineering with Computers*, 1-19.
- Omran, M. G., Engelbrecht, A. P., & Salman, A. (2004). Image classification using particle swarm optimization. *Recent advances in simulated evolution and learning*, 347-365.
- Opitz, D., & Shavlik, J. (1996). Actively searching for an effective neural-network ensemble. *Connection Science*, 8(3), 337-353.
- Orhunbilge, N., (1996). *Uygulamalı Regresyon ve Korelasyon Analizi*, İstanbul Üniversitesi

- Özdarıcı Ok, A., Akar, Ö., ve Güngör, O., (2011). Rastgele orman sınıflandırma yöntemi yardımıyla tarım alanlarındaki ürün çeşitliliğinin sınıflandırılması. VI. TUFUAB Teknik Sempozyumu, 21-25 Şubat, Antalya, Türkiye.
- Özalcı, Y., (2008). *Çok Değişkenli Uyarlanabilir Regresyon Kesitleri: MARS*. Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Ana Bilim Dalı, 69, Ankara.
- Özkurt Y. F., (2013). *Refinements, extensions and modern applications of conic multivariate adaptive regression splines*. Doctoral Thesis, Middle East Technical University, Institute of Applied Mathematics, Department of Scientific Computing, 114, Ankara.
- Özmen A., (2010). *Robust conic quadratic programming applied to quality improvement – a robustification of CMARS*. Master Thesis, Middle East Technical University, Institute of Applied Mathematics, Department of Scientific Computing, 155, Ankara.
- Öztürk, D. (2002). *Uydu görüntüleri kullanılarak arazi kullanımı/arazi örtüsündeki değişimlerin belirlenmesi ve Samsun örneği*. Yüksek Lisans Tezi, Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü Jeodezi ve Fotogrametri Mühendisliği Anabilim Dalı, 82, Samsun.
- Öztürkcan M., (2009). *Regresyon Analizi*, Maltepe Üniversitesi Yayınları.
- Pal, M. & Mather, P. M., (2005). Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, 26, 1007-1011.
- Park, S., Hamm, S., Jeon, H. & Kim J., (2017) Evaluation of logistic regression and multivariate adaptive regression spline models for groundwater potential mapping using R and GIS, *Sustainability*, 9, 1157.
- Phiri, D., & Morgenroth J., (2017). Developments in landsat land cover classification methods: a review. *Remote Sensing*. 9(9) 967.
- Phyu, T. N., (2009). Survey of classification techniques in data mining. *International Multi Conference of Engineers and Computer Scientists*, 18-20 March, Hong Kong.
- Piao, Y., Piao, M., Jin, C. H., Shon, H. S., Chung, J. M., Hwang, B. & Ryu, K. H. (2015). A new ensemble method with feature space partitioning for high-dimensional data classification. *Mathematical Problems in Engineering*. 1, 590678.
- Quirós, E., Felicísimo, Á. M., & Cuartero, A., (2009). Testing multivariate adaptive regression splines (MARS) as a method of land cover classification of TERRA-ASTER satellite images. *Sensors*, 9(11), 9011-9028.
- Rahman, A. F. R., & Fairhurst, M. C. (1999). Serial combination of multiple experts: a unified evaluation. *Pattern Analysis & Applications*, 2, 292-311.
- Ranawana, R., & Palade, V. (2006). Multi-classifier systems: Review and a roadmap for developers. *International journal of hybrid intelligent systems*, 3(1), 35-61.
- Rendell, L. & Cho, H. (1990). Empirical learning as a function of concept character, *Machine Learning* 5, 267-298.
- Rozenstein, O., & Karnieli, A. (2011). Comparison of methods for land-use classification incorporating remote sensing and GIS inputs. *Applied Geography*, 31(2), 533-544.
- Saini, R., & Ghosh, S. K. (2017). Ensemble classifiers in remote sensing: a review. *International Conference on Computing, Communication and Automation (ICCCA)*, 5-6 May, Greater Noida, India. doi: 10.1109/CCAA.2017.8229969.
- Salah, M., (2017). A survey of modern classification techniques in remote sensing for improved image classification. *Journal of Geomatics*, 11(1), 21.

- Sanders, S., & Giraud-Carrier, C. (2017). Informing the use of hyperparameter optimization through metalearning. *2017 IEEE International Conference on Data Mining (ICDM)*, 18-21 November, New Orleans, LA, USA.
- Santoso, N., & Wulandari, S. P., (2018). Hybrid Support Vector Machine to Preterm Birth Prediction. *Indonesian Journal of Electronics and Instrumentation Systems (IJEIS)*, 8, 191-200.
- Sarker, R. A., & Newton, C. S. (2007). Optimization Modelling: A Practical Approach. CRC Press.
- Sasena, M. J. (2002). *Flexibility and efficiency enhancement for constrained global design optimization with kriging approximations*. PhD thesis, University of Michigan, Department of Mechanical Engineering, 239, USA.
- Savaş, S. (2019). *Karotis arter intima media kalınlığının derin öğrenme ile sınıflandırılması*. Doktora Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Ana Bilim Dalı, 156, Ankara.
- Sephton, P. (2001). Forecasting Recessions: Can We Do Better on MARS. *Federal Reserve Bank of St. Louis Review*, 39-49.
- Sevimli, Y. (2009). *Çok değişkenli uyarlanabilir regresyon uzanımlarının bir split-mouth çalışmasında uygulaması*. Yüksek Lisans Tezi. Marmara Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı, 87, İstanbul.
- Sevgenler H., (2019). *Keçilere ait kimi özelliklerin canlı ağırlık üzerindeki etkilerini belirlemek amacıyla kullanılan veri madenciliği algoritmalarının (CART, CHAID ve MARS) karşılaştırılması*. Yüksek Lisans Tezi, Iğdır Üniversitesi, Fen Bilimleri Enstitüsü, Zooteknik Anabilim Dalı, Iğdır.
- Sharma, A., Tewari, R., Rana, S. S, Soni, R., & Soni, S. K. (2016). Classification, methods of determination and industrial applications. *Applied biochemistry and biotechnology*, 179, 1346-1380.
- Sharma, A., Liu, X., & Yang, X. (2018). Land cover classification from multi-temporal, multi-spectral remotelysensed imagery using patch-based recurrent neural networks. *Neural Networks*, 105, 346-355.
- Shavlik, J. W. Mooney, R. J. & Towell, G. G. (1991). Symbolic and neural learning algorithms: an experimental comparison. *Machine Learning* 6, 111-144.
- Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical Bayesian optimization of machine learning algorithms. *Advances in Neural Information Processing Systems*, 4, 2951-2959.
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information processing & management*, 45(4), 427-437.
- Stehman, S. V. (1999). Basic probability sampling design for thematic map accuracy assessment. *International Journal of Remote Sensing*, 20 (12) 2423-2441.
- Stehman, S. V., & Czaplewski, R. L. (2003). Introduction to special issue on map accuracy. *Environmental and Ecological Statistics* 10, 301-308 doi: <https://doi.org/10.1023/A:1025138423071>
- Stehman, S. V. & Wickham, J. D. (2011). Pixels, block of pixels, and polygons: choosing a spatial unit for thematic accuracy assessment, *Remote Sensing of Environment*, 115, 3044-3055. doi: <https://doi.org/10.1016/j.rse.2011.06.007>.
- Stevens, J. G., (1991). An investigation of multivariate adaptive regression splines for modeling and analysis of univariate and semi-multivariate time series systems, Doctoral Thesis, Naval Postgraduate School, 207, USA.

- Tan K., Jin X., Plaza A., Wang X., Xiao L., & Du P. (2016). Automatic change detection in high-resolution remote sensing images by using a multiple classifier system and spectral_spatial features. *IEEE Journal of Selected topics in applied Earth observations and Remote Sensing*, 9(8), 3439-3451. doi: <https://doi.org/10.1109/JSTARS.2016.2541678>
- Tan K, Zhang Y, Wang X, & Chen Y. (2019). Object-Based Change Detection Using Multiple Classifiers and Multi-Scale Uncertainty Analysis. *Remote Sensing*. 11(3). doi: <https://doi.org/10.3390/rs11030359>
- Tani, L., & Veelken, C. (2024). Comparison of Bayesian and particle swarm algorithms for hyperparameter optimisation in machine learning applications in high energy physics. *Computer Physics Communications*, 294, 108955.
- Temel G. O., Ankaralı H. & Yazıcı A. C., (2010). Regresyon Modellerine Alternatif Bir Yaklaşım: MARS, VIII. Ulusal Biyoistatistik Kongresi, 20 Eylül, Bursa
- Thakur, N., & Maheshwari, D., (2017). A review of image classification techniques. *International Research Journal of Engineering Technology*. 4, 1588-1591.
- Thornton, C., Hutter, F., Hoos, H. H., & Leyton-Brown, K. (2013). Auto-WEKA: combined selection and hyperparameter optimization of classification algorithms. *The 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, 847-855.
- Topaloğlu, R. H., Sertel, E., & Musaoğlu, N. (2016). Assessment of classification accuracies of Sentinel-2 and Landsat-8 data for land cover/use mapping. *The International archives of the photogrammetry, remote sensing and spatial information sciences*, 41, 1055-1059.
- Toprak S., (2011). *Çok değişkenli uyarlamalı regresyon eğrileri ve konik programlama ile zaman serilerinin modellenmesi*. Yüksek Lisans Tezi, Dicle Üniversitesi Fen Bilimleri Enstitüsü, Matematik Ana Bilim Dalı, 120, Diyarbakır, Türkiye.
- Tuia, D., Volpi, M., Copa, L., Kanevski, M., & Muñoz-Marí, J., (2011). A survey of active learning algorithms for supervised remote sensing image classification. *IEEE Journal of Selected Topics in Signal Processing*. 5(3), 606-617.
- Tukey, J., (1977). *Exploratory Data Analysis*. Addison-Wesley.
- TÜİK – Türkiye İstatistik Kurumu (2023). İstatistik Veri Portalı: Nüfus ve Demografi, <https://data.tuik.gov.tr/Kategori/GetKategori?p=Nufus-ve-Demografi-109> (Erişim tarihi: 29.10.2024)
- USGS, (2023). <https://www.usgs.gov/landsat-missions/landsat-collection-2-level-2-science-products>
- Ünal, B., (2009). *Çok değişkenli uyarlamalı regresyon uzanımları*. Yüksek Lisans Tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, 131, Ankara, Türkiye.
- Wang, M., Wan, Y., Ye, Z., & Lai, X. (2017). Remote sensing image classification based on the optimal support vector machine and modified binary coded ant colony optimization algorithm. *Information Sciences*, 402, 50-68.
- Wang, X., Liu, S., Du, P., Liang, H., Xia, J., & Li, Y. (2018). Object-based change detection in urban areas from high spatial resolution images based on multiple features and ensemble learning. *Remote Sensing*, 10(2), 276.
- Wickham, J., Stehman S. V., Gass, L., Dewitz, J. A., Sorenson, D. G., Granneman, B. J., Poss, V. R., & Baer, L. A. (2017). Thematic accuracy assessment of the 2011 National Land Cover Database (NLCD). *Remote Sensing of Environment*, 191, 328-341.
- Waske, B., Menz, G., & Benediktsson, J. A. (2007). Fusion of Support Vector Machines for

- Classifying SAR and Multispectral Imagery from Agricultural Areas. *IEEE Geoscience and Remote Sensing Symposium*, 23-28 July, Barcelona, Spain. doi: <https://doi.org/10.1109/IGARSS.2007.4423945>
- Weber, G. W., Batmaz, İ., Köksal, G., Taylan, P. & Yerlikaya, Ö. F. (2012). CMARS: a new contribution to nonparametric regression with multivariate adaptive regression splines supported by continuous optimization. *Inverse Problems in Science and Engineering*, 20(3), s. 371-400.
- Weisberg, S. (2005). *Applied Linear Regression*, John Wiley & Sons, Inc., Publication.
- Wilson, M. J., & Oreopoulos L. (2013) Enhancing a simple MODIS cloud mask algorithm for the Landsat data continuity mission. *IEEE Transactions on Geoscience and Remote*, 51(2), 723-731. doi: <https://doi.org/10.1109/TGRS.2012.2203823>.
- Xia, G. S., Hu, J., Hu, F., Shi, B., Bai, X., Zhong, Y., Zhang, L., & Lu, X., (2017). AID: a benchmark data set for performance evaluation of aerial scene classification. *IEEE Transactions on Geoscience and Remote Sensing*. 55(7), 3965-3981. doi: <https://doi.org/10.1109/TGRS.2017.2685945>.
- Xu, L., Krzyzak, A., & Suen, C. Y. (1992). Methods of combining multiple classifiers and their applications to handwriting recognition. *IEEE transactions on systems, man, and cybernetics*, 22(3), 418-435. doi: <https://doi.org/10.1109/21.155943>
- Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: theory and practice. *Neurocomputing*, 415, 295-316.
- Yazıcı C., (2011). *A computational approach to nonparametric regression: bootstrapping CMARS method*. Master Thesis Statistics, Middle East Technical University, Institute of Applied Mathematics, Department of Scientific Computing, 134, Ankara.
- Yao, D., Yang, J., & Zhan, X. (2013). A Novel Method for Disease Prediction: Hybrid of Random Forest and Multivariate Adaptive Regression Splines. *Journal of Computers*, 8(1), 170-177. doi: <https://doi.org/10.4304/jcp.8.1.170-177>.
- Ye, F., & Ding, F. (2018). Research and application of unbalanced data classification. *Computer Applications and Software*, 35(1), 132-136.
- Yu, T., & Zhu, H. (2020). Hyper-parameter optimization: A review of algorithms and applications. arXiv preprint arXiv:2003.05689.
- Zabihi, M., Pourghasemi, H. R., Pourtaghi, Z. S., & Behzadfar, M. (2016). GIS-based multivariate adaptive regression spline and random forest models for groundwater potential mapping in Iran. *Environmental Earth Sciences*, 75, 1-19.
- Zadeh, L. A., (1983). A computational approach to fuzzy quantifiers in natural languages. *Computers and Mathematics with Applications*, 9(1), 149-184.
- Zhang, C. (2018). *Deep learning for land cover and land use classification*. Doctoral Thesis, Lancaster University. Lancaster University, 267, England. doi: <https://doi.org/10.17635/lancaster/thesis/428>.
- Zhang, A., & Tang, P. (2013). Fusion algorithm of pixel-based and object-based classifier for remote sensing image classification. *IEEE International Geoscience and Remote Sensing Symposium*, 21-26 July, Melbourne, VIC, Australia. doi: <https://doi.org/10.1109/IGARSS.2013.6723390>.
- Zhang, C., Sargent, I., Pan, X., Gardiner, A., Hare, J., & Atkinson, P. M., (2018). VPRS-Based Regional Decision Fusion of CNN and MRF Classifications for Very Fine Resolution Remotely Sensed Images. *IEEE Transactions on Geoscience and Remote Sensing*, 56(8), 4507-4521. doi: <https://doi.org/10.1109/TGRS.2018.2822783>.
- Zhigljavsky, A., & Zilinskas, A. (2007). *Stochastic global optimization*. Springer Science &

Business Media.

Zhou, Z. H., Wu, J., & Tang, W. (2002). Ensembling neural networks: many could be better than all. *Artificial intelligence*, 137(1-2), 239-263.



ÖZ GEÇMİŞ

Azize Uyar, İstanbul Nevzat Ayaz Anadolu Lisesi'ni bitirdikten sonra İstanbul Teknik Üniversitesi İnşaat Fakültesi, Geomatik Mühendisliği bölümünden 28.06.2014 tarihinde mezun oldu. OMÜ LEE Harita Mühendisliği Ana Bilim Dalı Yüksek Lisans programını 2017 tarihinde bitirdi, Harita Mühendisliği ABD doktora programına başladı. Lisans mezuniyetinden bu yana Araştırma Görevlisi olarak görev yapmakta olup, iyi derecede İngilizce bilmektedir. (2024).

İletişim Bilgileri

ORCID ID : 0000-0003-2625-3690

Yayınlar:

1. Uyar, A., Yılmaz, İ., & Öztürk, D. (2023). Yol Ağlarının Fraktal Analiz İle Değerlendirilmesi Samsun Örneği. *19. Türkiye Harita Bilimsel ve Teknik Kurultayı*, 4-6 Ekim, Ankara.
2. Uyar A. ve Öztürk D. (2023). An Ensemble Multivariate Adaptive Regression Splines-Random Forest for Image Classification. *11. Uluslararası Mühendislik Mimarlık ve Tasarım Kongresi*, 10-11 Haziran.
3. Uyar A. ve Öztürk D. (2023). Performance Evaluation of Random Forest, Support Vector Machines and Multivariate Adaptive Regression Splines for Image Classification. *11. Uluslararası Mühendislik Mimarlık ve Tasarım Kongresi*, 10-11 Haziran.
4. Uyar, A., & Öztürk, D. (2019). Fraktal Analiz Kullanılarak Arnavutköy'ün Kentsel Yayılma Özelliklerinin Zamansal Değişiminin İncelenmesi. *Teknik Bilimler Dergisi*, 9(3), 28-31.
5. Uyar, A., & Öztürk, D. (2019). Morphometric Analysis of the Sakarya Basin Stream Network Using Linear Morphometric Indices and Fractal Analysis. *International Symposium on Applied Geoinformatics*, 7-9 Kasım, İstanbul.
6. Uyar, A., & Öztürk, D. (2019). Investigation of Shoreline Morphometry of Kızılırmak Delta s Lagoons Using Fractal Dimension and Shoreline Development Index. *16th International Conference on Environmental Science and Technology*, 4 - 7 September, Rhodes, Greece .
7. Bektas, S., Pakel, E., Yildirim, R. E., Uyar, A., Kirici, Ü. and Karaahmetoglu, N. (2018). Space Resection with Zenith Angles. *International Journal of Research - Granthaalayah*, 6(2), 331-338., doi:10.5281/zenodo.1194758
8. Bektas, S., Karaahmetoglu, N., Yildirim, R. E., Uyar, A., Kirici, Ü. and Pakel, E. (2018). Three Point Resection in 3d With Distances. *International Journal of Research - Granthaalayah*, 6(2), 292-298., doi: 0.5281/zenodo.11947
9. Uyar, A., & Öztürk, D. (2018). Kentsel Yayılmadaki Zamansal Değişimlerin Fraktal Analiz Kullanılarak İncelenmesi Arnavutköy Örneği. *2 nd International Symposium on Innovative Approaches in Scientific Studies*, 30 Kasım-2 Aralık, Samsun, Turkey.
10. Uyar, A., & Öztürk, D. (2018). Kızılırmak Havzası Drenaj Ağının ve Alt Havza Sınırlarının Uzaktan Algılama ve CBS İle Otomatik Çıkarımı. *2nd International Symposium on Innovative Approaches in Scientific Studies* 30 Kasım-2 Aralık, Samsun, Turkey.

11. Uyar, A., & Öztürk, D. (2018). Investigation of Coastline Changes of Samsun-Turkey Using Remote Sensing and GIS. *2nd International UNIDOKAP Black Sea Symposium*, 28-30 Kasım, Samsun, Turkey.
12. Uyar, A., & Öztürk, D. (2018). Determination of Land Use Cover Changes of Kucuk Menderes Basin in Turkey. *2nd International UNIDOKAP Black Sea Symposium*, 28-30 Kasım, Samsun, Turkey.
13. Uyar, A., & Öztürk, D. (2018). Determination Of Shoreline Of Kızılırmak Delta And Lagoons Using Pan-Sharpned NDWI And MNDWI Water Indexes. *3rd International Conference on Civil and Environmental Engineering (ICOCEE)*, 24-27 Nisan, İzmir, Turkey.
14. Uyar, A., & Öztürk, D. (2018). Natural Environment and Fractal Geometry. *3rd International Conference on Civil and Environmental Engineering (ICOCEE)*, 24-27 Nisan, İzmir, Turkey.
15. Uyar, A., & Öztürk, D. (2017). Fraktal analizin yeryüzü araştırmalarında kullanılması. *Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 17(4), 147-155.
16. Uyar, A., & Öztürk, D. (2017). Fractal Analysis of Coastal Pattern of the Kızılırmak Delta and Lagoons. *UCTEA International Geographical Information Systems Congress*, 15-18 November, Adana, Turkey.
17. Uyar, A., & Öztürk, D. (2017). From Euclidean to Fractal Geometry in Geography and Geosciences. *International Symposium on GIS Applications in Geography Geosciences*, 18-21 October, Çanakkale, Turkey.
18. Uyar, A., & Öztürk, D. (2017). Fractal Analysis in Geomorphology. *International Symposium on GIS Applications in Geography Geosciences*, 18-21 October, Çanakkale, Turkey.
19. Uyar, A., & Öztürk, D. (2017). Fractal Analysis of the Lagoons in the Kızılırmak Delta. *19th International Symposium on Environmental Pollution and its Impact on Life in the Mediterranean Region*, 4-6 October, Roma, Italy.
20. Uyar, A., & Uluğtekin, N. N. (2017). Generalization Problems of 3D Models and LoD Concept. *28th International Cartographic Conference*, 2-7 July, Washington DC, USA.
21. Uyar, A., & Uluğtekin, N. N. (2017). A Proposal for Generalization of 3D Models. *4th International Workshop on Geoinformation Science: GeoAdvances*, 14-15 October, Karabük, Turkey.
22. Uyar, A., & Öztürk, D. (2017). Fraktal Analizin Yeryüzü Araştırmalarında Kullanılması. *TUFUAB IX. Teknik Sempozyumu*, 27-29 Nisan, Afyon, Türkiye.
23. Uyar, A., & Uluğtekin, N. N. (2016). 3B Modellemede Genelleştirme Problemleri ve LoD Kavramı. *6. Uzaktan Algılama-CBS Sempozyumu (UZAL-CBS 2016)*, 17-19 Ekim, Adana, Türkiye.