



**MAKİNE ÖĞRENİMİ TEKNİKLERİ
KULLANILARAK ORMAN YANGINI RİSK
HARİTALARININ OLUŞTURULMASI: ANTALYA
ŞELALE ŞEFLİĞİ ÖRNEĞİ**

**2024
YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ**

Aybike Göksu BEKTAŞ

**Tez Danışmanı
Prof. Dr. İsmail Rakıp KARAŞ**

**MAKİNE ÖĞRENİMİ TEKNİKLERİ KULLANILARAK ORMAN
YANGINI RİSK HARİTALARININ OLUŞTURULMASI: ANTALYA
ŞELALE ŞEFLİĞİ ÖRNEĞİ**

Aybike Göksu BEKTAŞ

**Tez Danışmanı
Prof. Dr. İsmail Rakıp KARAŞ**

**T.C.
Karabük Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalında
Yüksek Lisans Tezi
Olarak Hazırlanmıştır**

**KARABÜK
Ağustos 2024**

Aybike Göksu BEKTAŞ tarafından hazırlanan “MAKİNE ÖĞRENİMİ TEKNİKLERİ KULLANILARAK ORMAN YANGINI RİSK HARİTALARININ OLUŞTURULMASI: ANTALYA ŞELALE ŞEFLİĞİ ÖRNEĞİ” başlıklı bu tezin Yüksek Lisans Tezi olarak uygun olduğunu onaylarım.

Prof. Dr. İsmail Rakıp KARAŞ

Tez Danışmanı, Bilgisayar Mühendisliği Anabilim Dalı

Bu çalışma, jürimiz tarafından Oy Birliği ile Bilgisayar Mühendisliği Anabilim Dalında Yüksek Lisans tezi olarak kabul edilmiştir. 29/08/2024

Ünvanı, Adı SOYADI (Kurumu)

İmzası

Başkan : Prof. Dr. İsmail Rakıp KARAŞ (KBÜ)

Üye : Doç. Dr. Adib M. Monzer HABBAL (KBÜ)

Üye : Doç. Dr. Zennure UÇAR (İKÇÜ)

ONLINE

KBÜ Lisansüstü Eğitim Enstitüsü Yönetim Kurulu, bu tez ile, Yüksek Lisans derecesini onamıştır.

Doç. Dr. Zeynep ÖZCAN

Lisansüstü Eğitim Enstitüsü Müdürü



“Bu tezdeki tüm bilgilerin akademik kurallara ve etik ilkelere uygun olarak elde edildiğini ve sunulduğunu; ayrıca bu kuralların ve ilkelerin gerektirdiği şekilde, bu çalışmadan kaynaklanmayan bütün atıfları yaptığımı beyan ederim.”

Aybike Göksu BEKTAŞ

ÖZET

Yüksek Lisans Tezi

MAKİNE ÖĞRENİMİ TEKNİKLERİ KULLANILARAK ORMAN YANGINI RİSK HARİTALARININ OLUŞTURULMASI: ANTALYA ŞELALE ŞEFLİĞİ ÖRNEĞİ

Aybike Göksu BEKTAŞ

Karabük Üniversitesi

Lisansüstü Eğitim Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı:

Prof. Dr. İsmail Rakıp KARAŞ

Ağustos 2024, 43 sayfa

Ana yangın faktörlerini analiz ederek orman yangını olasılık seviyelerini belirlemek, orman yöneticilerine yangın önleme stratejileri, yakıt yönetimi, yangın güvenlik önlemleri, acil durum planlaması ve yangınla mücadele ekiplerinin yerleştirilmesi gibi konularda kritik kararlar almaları için temel sağlayabilir. Orman yangını olasılık haritalarını oluşturmak için bitki örtüsü faktörleri, topoğrafik faktörler, iklim faktörleri, yollar ve yerleşim alanları gibi bazı özelliklere yakınlığı içeren ana yangın faktörleri dikkate alınmıştır. Son yıllarda, makine öğrenimi algoritmaları orman yangını olasılığını tahmin etmede etkili bir araç haline gelmiştir. Bu çalışmada, yaygın olarak kullanılan iki makine öğrenmesi modeli olan Lojistik Regresyon (LR) ve Destek Vektör Makineleri (DVM), Coğrafi Bilgi Sistemi (CBS) teknikleri ile entegre edilerek bir orman yangını olasılık haritası oluşturulması amaçlanmıştır. Çalışma alanı, Türkiye'nin Akdeniz şehri Antalya'da bulunan Antalya Orman Bölge Müdürlüğü

sınırları içerisindeki Şelale Orman İşletme Şefliği'dir (OİM). Çalışmada, orman yangını olasılığını etkileyen ana faktörler ağaç türleri, taç kapallığı, ağaç evresi, eğim, bakı ve yollara uzaklık olarak belirlenmiştir. Modellerin eğitim aşamasında Şelale OAK'da 2001'den 2021'e kadar meydana gelen tarihi orman yangınlarının mekânsal dağılımı dikkate alınmıştır. Yangın olasılığı haritalarının doğruluğu Eğri Altındaki Alan (AUC) değeri kullanılarak doğrulanmıştır. Makine öğrenmesi modellerinin gerçekleştirilmesi sonucunda, harita üzerinde beş yangın olasılığı seviyesine (çok yüksek, yüksek, orta, düşük ve çok düşük) kategorize edilen 47.086 nokta için tahminler yapılmıştır. Sonuçlar, LR modeli tarafından üretilen yangın olasılığı haritasının doğruluğunun (AUC = 0,845) SVM modeli tarafından üretilen yangın olasılığı haritasının doğruluğundan (AUC = 0,748) daha iyi olduğunu göstermiştir. Modeller tarafından oluşturulan olasılık haritalarına göre, çalışma alanındaki ormanların yarısından fazlası çok yüksek/yüksek yangın olasılığı seviyeleri olarak sınıflandırılmıştır.

Anahtar Sözcükler : Makine öğrenmesi, Orman yangını, Risk haritası, Coğrafi bilgi sistemleri.

Bilim Kodu : 92432

ABSTRACT

Master Thesis

CREATING FOREST FIRE RISK MAPS USING MACHINE LEARNING TECHNIQUES: THE EXAMPLE OF ANTALYA ŞELALE ŞEFLİĞİ

Aybike Göksu BEKTAŞ

**Karabük University
Institute of Graduate Programs
Department of Computer Engineering**

Thesis Advisor:

Prof. Dr. İsmail Rakıp KARAŞ

August 2024, 43 pages

Determining wildfire probability levels by analyzing the main fire factors can provide the basis for forest managers to make critical decisions on fire prevention strategies, fuel management, fire safety measures, emergency planning and deployment of firefighting teams. To create wildfire probability maps, the main fire factors have been considered, including vegetation factors, topographic factors, climatic factors, proximity to certain features such as roads and residential areas. In recent years, machine learning algorithms have become an effective tool for predicting wildfire probability. In this study, two widely used machine learning models, Logistic Regression (LR) and Support Vector Machines (SVM), are integrated with Geographic Information System (GIS) techniques to create a forest fire probability map. The study area is the Şelale Forest Management Directorate (FMD) within the borders of the Antalya Regional Directorate of Forestry located in the Mediterranean city of Antalya, Turkey. In the study, the main factors affecting the probability of forest fire were

identified as tree species, crown closure, tree stage, slope, aspect and distance to roads. In the training phase of the models, the spatial distribution of historical forest fires that occurred from 2001 to 2021 in the Şelale Şefliği FMD was taken into account. The accuracy of the fire probability maps was verified using the Area Under the Curve (AUC) value. The implementation of machine learning models resulted in predictions for 47,086 points on the map categorized into five fire probability levels (very high, high, medium, low and very low). The results showed that the accuracy of the fire probability map produced by the LR model (AUC = 0.845) was better than the accuracy of the fire probability map produced by the SVM model (AUC = 0.748). According to the probability maps generated by the models, more than half of the forests in the study area were classified as very high/high fire probability levels.

Key Word : Machine larning, Forest fire, Risk map, Geographic information system.

Science Code : 92432

TEŐEKKÜR

Çalıőmalarımın her aőamasında yanımda olan deęerli hocam ve tez danıőmanım Prof. Dr. İsmail Rakıp KARAAŐ'a yardımları, önerileri ve yönlendirmeleri için teőekkür ederim.

Bu tez, 2210309 numaralı TÜBİTAK projesi ile desteklenmiőtir. Projenin yürütücüsü olan sayın Prof. Dr. Abdullah Emin AKAY'a tez ile ilgili katkılarından ve yardımlarından dolayı teőekkürlerimi sunarım.

Bu süreçte bana inanan ve desteklerini hiçbir zaman esirgemeyen annem ve ablama teőekkür ederim.

İÇİNDEKİLER

	<u>Sayfa</u>
KABUL.....	ii
ÖZET.....	iv
ABSTRACT.....	vi
TEŞEKKÜR.....	viii
İÇİNDEKİLER	ix
ŞEKİLLER DİZİNİ.....	xi
ÇİZELGELER DİZİNİ	xii
KISALTMALAR	xiv
BÖLÜM 1	1
GİRİŞ	1
1.1. LOJİSTİK REGRESYON MODELİ	3
1.2. DESTEK VEKTÖR MAKİNELERİ MODELİ.....	4
1.3. LOJİSTİK REGRESYON MODELİ VE DESTEK VEKTÖR MODELİNİN KARŞILAŞTIRILMASI.....	4
BÖLÜM 2	6
LİTERATÜR TARAMASI.....	6
BÖLÜM 3	9
MATERYAL VE METOT	9
3.1. VERİ KÜMESİ OLUŞTURMA	9
3.2. VERİ ÖN İŞLEME	13
3.3. MODEL OLUŞTURMA.....	14
3.3.1. Lojistik Regresyon için Model Parametreleri.....	17
3.3.1.1. Regülerizasyon Parametresi (C)	17
3.3.1.2. Solver (Çözücü)	17
3.3.1.3. Regülerizasyon Türü (L1, L2)	17

	<u>Sayfa</u>
3.3.1.4. Maksimum Yineleme (max_iter)	17
3.3.1.3. Regülerizasyon Türü (L1, L2)	17
3.3.2. Destek Vektör Makineleri için Model Parametreleri.....	18
3.3.2.1. C parametresi (Ceza Parametresi).....	18
3.3.2.2. Kernel (Çekirdek) Fonksiyonu	18
3.3.2.3. Gamma (γ) Parametresi.....	18
3.3.2.4. Sınıf Ağırlıkları (class_weight).....	18
3.4. ORMAN YANGINI OLASILIK HARİTALAMASI	19
BÖLÜM 4	20
BULGULAR VE TARTIŞMA	20
4.1. YANGINI ETKİLEYEN FAKTÖRLER	20
4.2. ML MODELLERİNİN SONUÇLARI.....	26
BÖLÜM 5	32
SONUÇLAR VE ÖNERİLER.....	32
KAYNAKLAR	33
EK AÇIKLAMALAR A.	39
ÖZGEÇMİŞ	43

ŞEKİLLER DİZİNİ

	<u>Sayfa</u>
Şekil 3.1. Metodolojinin akış şeması.....	10
Şekil 3.2. Çalışma alanı.....	11
Şekil 4.1. Ağaç türleri haritası.....	21
Şekil 4.2. Taç kapalılık haritası.....	22
Şekil 4.3. Ağaç aşamaları haritası.....	23
Şekil 4.4. Eğim sınıfları haritası.....	24
Şekil 4.5. Bakı haritası.....	25
Şekil 4.6. Yol tampon bölgesi haritası.....	26
Şekil 4.7. Yangın durumu (yangın var/yangın yok) veri noktaları.....	27
Şekil 4.8. Modellerde kullanılan veri kümesinin bir kısmı.....	27
Şekil 4.9. LR modeliyle tahmin edilen orman yangını olasılık haritası.....	28
Şekil 4.10. SVM modeliyle tahmin edilen orman yangını olasılık haritası.....	29
Şekil 4.11. LR modeli için ROC eğrisi.....	31
Şekil 4.12. SVM modeli için ROC eğrisi.....	31

ÇİZELGELER DİZİNİ

Sayfa

Çizelge 3.1. Analizde dikkate alınan yangını etkileyen faktörlerin sınıfları.....	12
Çizelge 4.1. Ağaç türlerinin alansal dağılımı.....	20
Çizelge 4.2. Taç kapalılığının alansal dağılımı.....	22
Çizelge 4.3. Ağaç evresinin alansal dağılımı.....	23
Çizelge 4.4. Eğim sınıflarının alansal dağılımı.....	23
Çizelge 4.5. Bakı sınıflarının alansal dağılımı.....	24
Çizelge 4.6. En boy sınıflarının alansal dağılımı.....	25
Çizelge 4.7. Yangın olasılık seviyelerinin alansal dağılımı.....	29
Çizelge 4.8. LR modeliyle seçilen değişkenlerin önemi.....	30

SİMGELER VE KISALTMALAR DİZİNİ

KISALTMALAR

- ML : Makine Öğrenmesi
DVM : Destek Vektör Makineleri
LR : Lojistik Regresyon
CBS : Coğrafi Bilgi Sistemleri
AUC : Area Under the Curve (Eğrinin Altındaki Alan)
ROC : Receiver Operating Characteristic (Alıcı Çalışma Özelliği)
H : Hipotez
SYM : Sayısal Yükseklik Modeli
DT : Decision Tree (Karar Ağacı)
RF : Random Forest (Rastgele Orman)
RBF : Radial Basis Function (Radyal Taban Fonksiyonu)

BÖLÜM 1

GİRİŞ

Orman yangını, doğal kaynakların tahrip olmasına neden olan, ekosistemleri geniş ölçüde etkileyen ve hatta insan hayatının kaybına neden olan dünya çapında ciddi bir tehdittir (Kala, 2023). Küresel ölçekte her yıl milyonlarca hektar orman alanı yangınlardan zarar görmekte ve bu yangınlar önemli ekonomik kayıplara da neden olmaktadır (Erkan vd., 2024). Orman yangınlarının sıklığı ve şiddeti son zamanlarda kuraklık ve sıcak hava dalgalarının yaşandığı bölgelerde artış göstermiştir, çünkü bu bölgelerde yangın tutuşması ve yangının yayılma olasılığı çok yüksektir (Duine vd., 2022).

Akdeniz ormanları, özellikle Cezayir, Yunanistan, İtalya, Portekiz, İspanya ve Türkiye'de bitki türlerinin kompozisyonu, yüksek sıcaklıklar ve yaz aylarında yaşanan kuraklıklar nedeniyle orman yangınlarına karşı oldukça hassastır. 2021 yazında, yangınlar Akdeniz havzasında 620.000 hektardan fazla alana ve 102.000 hektardan fazla korunan alana zarar vermiştir (EFFIS, 2021; San-Miguel-Ayanz vd., 2022). Çok yüksek hava sıcaklıkları ve düşük bağıl nem gibi aşırı iklim koşullarından kaynaklanan ve 2021 yılında meydana gelen mega orman yangınları nedeniyle Türkiye'nin Akdeniz bölgelerinde yaklaşık 140.000 hektar orman alanı yanmıştır (OGM, 2022). Bu mega orman yangınları sırasında Antalya bölgesinde yaklaşık 60.000 hektar ormanlık alan yanmıştır (Türkan vd., 2023).

Yangına hassas ormanların bulunduğu bölgelerde, orman yangını olasılığını tahmin etmek için gelişmiş ve proaktif yaklaşımlara duyulan ihtiyaç giderek daha zorunlu hale gelmektedir. Orman yangınlarının ekolojik ve ekonomik etkilerini azaltmak için yangın olasılığı yüksek orman alanlarının önceden belirlenmesi ve bu bölgelerde yangınlara karşı gerekli önlemlerin alınması gerekmektedir (Akay ve Erdoğan, 2017). Yangın olasılığı yüksek orman alanlarının doğru bir şekilde tahmin edilebilmesi için,

yangını etkileyen faktörlerin sayısal verilerine dayalı olarak CBS teknikleri kullanılarak yangın olasılığı dağılım haritası üretilebilir (Jaiswal vd., 2002). Orman yangını olasılığı haritalaması, orman vejetasyon yapısı, topoğrafik özellikler, iklim faktörleri ve yollar ve yerleşim alanları gibi bazı özelliklere yakınlık gibi yangını etkileyen çeşitli faktörleri içerir (Carmel vd., 2009).

Yangın olasılığı haritalarında genellikle dikkate alınan başlıca orman vejetasyon yapıları ağaç türleri, taç kapallılığı ve ağaç evresidir. İğne yapraklı ağaçların yüzdesi arttıkça yangın olasılığı seviyesi yükselirken, yaprak döken ağaçlar yangına daha dayanıklıdır (Güney vd., 2016). Taç kapallılığı ile orman yangını olasılığı arasında pozitif bir ilişki vardır. Genç meşcerelerde yangın olasılığı yüksek, erken gelişim evrelerinde ise düşüktür (Sağlam vd., 2008; Küçük vd., 2009). Topoğrafik özellikler açısından, zemin eğimi arttıkça yangın olasılığı artar ve yangın yamaçlarda hızla ilerler. Orman yangını olasılığı, düşük nem ve yüksek sıcaklığın bir sonucu olarak güney yönlerde yüksektir (Lin ve Rinaldi, 2009). Sıcaklık, rüzgâr ve yağış, yangına neden olan ana iklim parametreleri olarak kabul edilir. Hava sıcaklığı yükseldikçe ve yağış miktarı azaldıkça orman yangını olasılığı artmaktadır. Rüzgâr hızı ve yönü, yayılma hızı ve yangın yönü üzerinde önemli etkilere sahiptir (Çanakçıoğlu, 1993). Yollara ve yerleşim alanlarına olan uzaklık da orman yangını oluşumunda önemli bir rol oynamaktadır. Antropolojik faktörler nedeniyle orman bu nesnelere yakın olduğunda orman yangını olasılığı artmaktadır (Güney vd., 2016; Jaiswal vd., 2002). Yalnızca CBS araçları kullanılarak yapılan yangın olasılığı haritalaması, yangını etkileyen faktörlerin karmaşıklığı ve dinamik doğası ile başa çıkmada sınırlamalara sahiptir. Orman yangınlarıyla ilgili çalışmalarda kullanılan verilerdeki hızlı artış, araştırmacıların çok büyük veri kümeleriyle uğraşmasını gerektirmektedir. CBS tekniklerinin gelişmiş matematiksel modellerle entegrasyonu, orman yangını olasılığını doğru bir şekilde tahmin etmek için etkili bir şekilde kullanılabilir. Son yıllarda, makine öğrenimi (ML) modelleri, yangını etkileyen faktörlere dayalı olarak orman alanlarını modelleyerek orman yangını olasılığının tahmini için önemli bir potansiyel sunmuştur (Pourghasemi vd., 2020). Orman yangını olasılığı haritalama ve yönetimi, makine öğrenmesi modellerinin başlıca uygulamalarından biridir (Khan vd., 2014). Başlıca makine öğrenimi modelleri arasında lojistik regresyon (LR), destek vektör makineleri (SVM), Naive Bayes (NB), karar ağacı (DT) ve rastgele orman (RF)

yer almaktadır. (Arpacı vd., 2014; Güney vd., 2016). LR ve SVM modelleri en sık kullanılan ML modelleri arasındadır (Kaur vd., 2014). Her iki algoritma da sınıflandırma problemlerinde etkilidir. Lojistik regresyon, yangın olup olmayacağını belirlemede basit ve anlaşılır bir model sunarken, SVM, verilerin doğrusal olarak ayrışmadığı durumlarda bile yüksek doğruluk sağlar. Bu algoritmalar, yangın riskini tahmin etmek için farklı çevresel değişkenleri dikkate alarak güçlü performans gösterir. Bu çalışmada, orman yangını olasılık haritalarının geliştirilmesi için bu iki ML modelinin kullanılması amaçlanmıştır. Olasılık haritalarının doğruluğu AUC değerleri kullanılarak test edilmiştir.

1.1. LOJİSTİK REGRESYON MODELİ

Lojistik regresyon, doğrusal regresyon modelinin bir uzantısı olarak ikili sınıflandırma problemleri için kullanılan bir makine öğrenimi algoritmasıdır. Bu algoritma, özellikle olayın gerçekleşme olasılığını tahmin etmek amacıyla yaygın olarak kullanılır. Lojistik regresyon, bağımlı değişkenin ikili olduğu (örn. yangın çıktı/çıkmadı) durumlarda kullanılır ve bu durumda bağımlı değişkenin aldığı değerler 0 ve 1 olarak kodlanır.

Lojistik regresyon, bir bağımsız değişkenler kümesi ile hedef değişken arasında ilişki kurar ve bu ilişkiyi doğrusal bir model olarak ifade eder. Ancak bu doğrusal model, logit fonksiyonu adı verilen bir dönüştürme fonksiyonuyla olasılık değerine dönüştürülür. Bu sayede, bağımlı değişkenin değeri 0 ile 1 arasında bir olasılık olarak yorumlanabilir. Lojistik regresyonun avantajı, modelin açıklanabilirliğinin yüksek olması ve model parametrelerinin yorumlanabilir olmasıdır. Ayrıca, aşırı uyum (overfitting) sorunlarına karşı genellikle daha dayanıklıdır.

Orman yangını risk haritası oluştururken lojistik regresyon, yangın çıkma olasılığını tahmin etmek için kullanılabilir. Yangına etki eden faktörler (örneğin, sıcaklık, nem, rüzgâr hızı, bitki örtüsü tipi) bağımsız değişkenler olarak modele dahil edilir. Model, bu değişkenler kullanılarak bir alanın yangın riski taşıyıp taşımadığını belirler. Lojistik regresyon, özellikle risk haritalarının anlaşılabilirliği ve uygulanabilirliği açısından güçlü bir araçtır.

1.2. DESTEK VEKTÖR MAKİNELERİ MODELİ

Destek Vektör Makineleri (SVM), özellikle sınıflandırma ve regresyon analizlerinde kullanılan güçlü bir makine öğrenimi algoritmasıdır. SVM, veriyi farklı sınıflara ayıran en iyi hiperdüzlemi bulmaya çalışır. Bu hiperdüzlem, iki sınıfı birbirinden en iyi şekilde ayıran düzlemdir ve sınıflar arasındaki marjı maksimum yapacak şekilde seçilir.

SVM algoritması, sadece doğrusal ayrıştırma yapabilen bir model olarak başlasa da, çekirdek (kernel) hileleri kullanarak doğrusal olmayan sınıflandırma problemlerini de çözebilir. Çekirdek hileleri, orijinal verileri daha yüksek boyutlu bir uzaya projeler ve bu sayede doğrusal olmayan verilerin ayrıştırılması mümkün hale gelir. Yaygın kullanılan çekirdek fonksiyonları arasında lineer, polinomial, radyal baz fonksiyonu (RBF) ve sigmoid yer alır.

Orman yangını risk haritası oluştururken, SVM algoritması yangın çıkma riski taşıyan ve taşımayan bölgeleri sınıflandırmak için kullanılabilir. Örneğin, SVM modeli, yangına etkisi olabilecek çevresel ve meteorolojik verileri (örneğin, sıcaklık, nem, yakıt tipi, eğim) kullanarak, yangın çıkma riskini taşıyan alanları belirleyebilir. SVM'nin en büyük avantajlarından biri, yüksek boyutlu veri setlerinde de etkili performans göstermesi ve doğrusal olmayan sınıflandırma problemlerini başarılı bir şekilde çözebilmesidir.

1.3. LOJİSTİK REGRESYON MODELİ VE DESTEK VEKTÖR MAKİNELERİ MODELİNİN KARŞILAŞTIRILMASI

Her iki algoritma da orman yangını risk haritası oluşturmak için etkili araçlar olmakla birlikte, kullanım alanları ve performansları farklılık gösterebilir. Lojistik regresyon, modelin açıklanabilirliği ve yorumlanabilirliği açısından öne çıkarken, SVM genellikle daha karmaşık veri setlerinde daha iyi performans gösterir. Ayrıca, SVM, veri setinde doğrusal olmayan ayrımların bulunduğu durumlarda da başarılı sonuçlar verirken, lojistik regresyon doğrusal ilişkilerin baskın olduğu durumlarda daha uygun olabilir.

Orman yangını risk haritası oluşturulurken, hangi algoritmanın daha iyi sonuç verdiđi, verinin yapısına ve hedeflenen amaca bađlı olarak deđiřebilir. Bu nedenle, her iki algoritmanın da performansını deđerlendirmek ve sonuçları karřılařtırmak, en uygun yöntemin sečilmesi ačilından önemlidir.



BÖLÜM 2

LİTERATÜR TARAMASI

Orman yangınları, doğal ekosistemler ve insan yaşamı için büyük tehditler oluşturduğundan, yangın riski haritalaması, yangınları önceden tahmin etmek ve bu tür felaketlerin etkilerini en aza indirmek için kritik bir araçtır. Yangın riski haritalarının oluşturulması, yangın yönetim stratejilerinin geliştirilmesinde önemli bir rol oynar. Pyne (2001), orman yangınlarının tarihsel gelişimini inceleyerek, yangın risk haritalamasının önemini vurgulamıştır.

Justice vd. (2002), MODIS uydu verilerinin yangın tespiti ve izlenmesi için nasıl kullanıldığını ele almıştır. Bu tür uydu verileri, yangınların zamanında ve doğru bir şekilde tespit edilmesine yardımcı olmaktadır.

Orman yangınları, çevresel ve ekonomik zararları nedeniyle küresel ölçekte büyük bir sorun olarak kabul edilmektedir. Bu yangınların etkili bir şekilde önlenmesi ve yönetilmesi, yangın riskini doğru bir şekilde tahmin etmeye dayanmaktadır. Geleneksel yöntemler, genellikle meteorolojik veriler ve coğrafi bilgi sistemleri (CBS) tabanlı olsa da, bu yöntemler yangın riskini tahmin etmede yetersiz kalabilmektedir. Son yıllarda, makine öğrenimi algoritmalarının gelişimi, yangın risk haritalarının oluşturulmasında yeni ve daha etkili yaklaşımlar sunmuştur.

Yangına duyarlılık haritaları, orman yangınlarını kontrol etmek ve yönetmek için olası yangın oluşum noktalarını yönlendirmeye yardımcı olabilir. Bununla birlikte, orman yönetimi yetkililerinin güney yamacı boyunca ve yüksek değerli kaynakların çevresinde yakıt molaları uygulamaya başlayabilecekleri öne sürüldü (Bar vd., 2023). Orman yangınlarından kaynaklanan çevresel kayıp düzeyi, hava ve iklim faktörleri, topografya, bitki örtüsü türleri, yakıt özellikleri ve insan yapımı özellikler dahil olmak üzere çeşitli çevresel parametrelerle ilişkilidir. Bu faktörler orman yangınlarının

yoğunluğunu ve büyüklüğünü etkiler ve genellikle katkıda bulunan parametreler olarak adlandırılır (Naderpour vd., 2021).

Syphard vd. (2007), Kaliforniya'da insan faaliyetlerinin orman yangını rejimleri üzerindeki etkisini incelemiştir. Bu çalışma, lojistik regresyon modellerini kullanarak insan etkisinin yangın riskini nasıl artırdığını göstermiştir.

Orman yangını analizinin çok değişkenli bir yapıya sahip olması hem birçok değişkeni hem de büyük veri kümelerini ele alabilecek teknikler gerektirdiği anlamına gelir. Bu nedenle orman yangınlarına ilişkin konumsal veri setlerinin analiz edilmesi görevinde makine öğrenmesi algoritmaları önemli rol oynamaktadır. Dahası, makine öğreniminin çıkarım yeteneği, orman yangınının mekansal dağılımının tahmin edilmesine son derece yardımcı olur (Tien Bui vd., 2018).

Uzaktan algılama ve CBS yapılarını makine öğrenimi modelleri ile kullanarak orman yangını riski olasılığının nasıl değerlendirileceğini çeşitli çalışmalar araştırmıştır. SVM, ilk olarak denetimli öğrenme ikili sınıflandırıcısına sahip popüler bir istatistiksel makine öğrenimi algoritmasıdır. Temel fikri, farklı örnek sınıfları arasındaki boşluğu, bir örnek bularak en üst düzeye çıkarmaktır (Tan ve Feng, 2023).

Zaman içinde, özellikle sensör gelişimindeki artışla birlikte ve artan veri kullanılabilirliği, SVM'nin yeryüzü gözlemleri ile ilgili birçok alanda başarıyla kullanılmasını sağlamıştır (Kalantar vd., 2020).

SVM'ler küçük eğitim veri kümelerini etkili bir şekilde yönetebilir ve sıklıkla yüksek doğruluk sağlar. Bununla birlikte, büyük ölçekli eğitim örneklerinin uygulanması zordur ve parametrelerin ve çekirdek fonksiyonlarının seçimine duyarlı olan çoklu sınıflandırma problemlerini çözmek zordur (Shao vd., 2022).

Cortes ve Vapnik (1995), SVM'nin teorik temelini oluşturan orijinal çalışmayı yayımlamışlardır ve SVM'nin sınıflandırma problemlerindeki üstünlüğünü kanıtlamışlardır (Cortes vd., 1995).

Cortez ve Morais (2007), Portekiz'deki yangın verilerini kullanarak SVM'nin yangın büyüklüğünü tahmin etmedeki etkinliğini araştırmıştır. Sonuçlar, SVM'nin diğer modellerle karşılaştırıldığında daha yüksek doğruluk sağladığını ortaya koymuştur. SVM'ler hem eğitim hem de test aşamalarında yavaştır. SVM'ler çok karmaşıktır ve büyük ölçekli görevlerde kapsamlı bellek gerektirir. SVM, en iyi sınıflandırma yöntemini belirlemeyi amaçlamaktadır (Sayad vd., 2019).

Modellerin orman yangını tehlikesi tahmininde sağladığı büyük performans artışına paralel olarak, tahminlerin itici güçlerinin ardındaki fiziksel mantık, orman yangını tehlikesini ve dinamiklerini tahmin etmek ve anlamak için güvenilir ve güvenilir bir yaklaşıma yönelik veriye dayalı öğrenmenin rolünü vurgulamaktadır (Kondylatos vd., 2022).

Lojistik regresyon (LR), 1990'lı ve 2000'li yıllarda yangının varlığı veya yokluğu gibi ikili sonuçlarla uğraşırken kullanılan baskın yöntemdi (Milanović vd., 2023). Bu model, yangın riskini tahmin etmek için sıklıkla kullanılan bir algoritmadır ve genelleştirilmiş doğrusal modeller ailesinin bir parçası olan bir modelleme tekniğidir (Phelps ve Woolford, 2021).

Martell vd. (1987), Ontario bölgesinde insan kaynaklı orman yangınlarını tahmin etmek için bir lojistik regresyon modeli geliştirmiştir. Bu model, meteorolojik veriler ve insan faaliyetlerine dayalı olarak günlük yangın olaylarını başarıyla tahmin etmiştir. Bu çalışma, günlük insan kaynaklı orman yangınlarını tahmin etmek için meteorolojik ve demografik verilerin nasıl işlenebileceğini göstermiştir. Lojistik regresyon, orman yangını riskinin belirlenmesinde kullanılan ilk makine öğrenmesi algoritmalarından biridir ve bu çalışmada modelin başarısı, ML tekniklerinin potansiyelini vurgulamaktadır.

Özellikle ikili sınıflandırma problemlerinde (örneğin, yangın çıktı/çıkmadı) etkin bir şekilde kullanılabilir. Adab vd. (2013) tarafından yapılan bir çalışmada, lojistik regresyon kullanılarak İran'ın Mazandaran eyaletinde yangın riski haritası oluşturulmuş ve modelin doğru tahminlerde bulunmada etkili olduğu görülmüştür.

Logistik regresyon modelini yorumlamak, diđer karmaşık modelleri yorumlamaktan çok daha kolaydır (Pham vd., 2020).



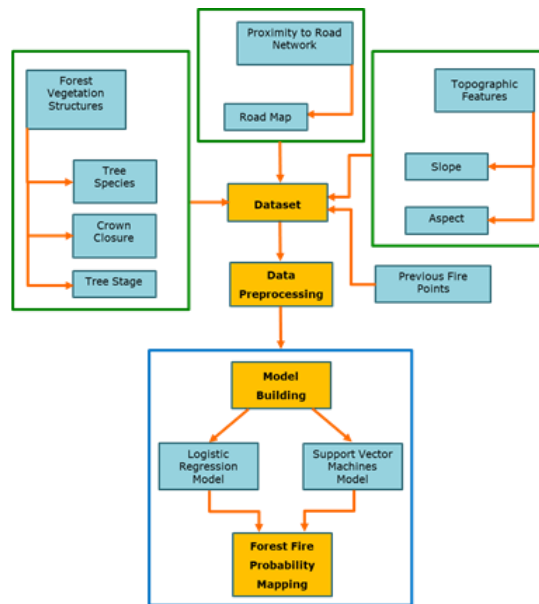
BÖLÜM 3

MATERTAL VE METOT

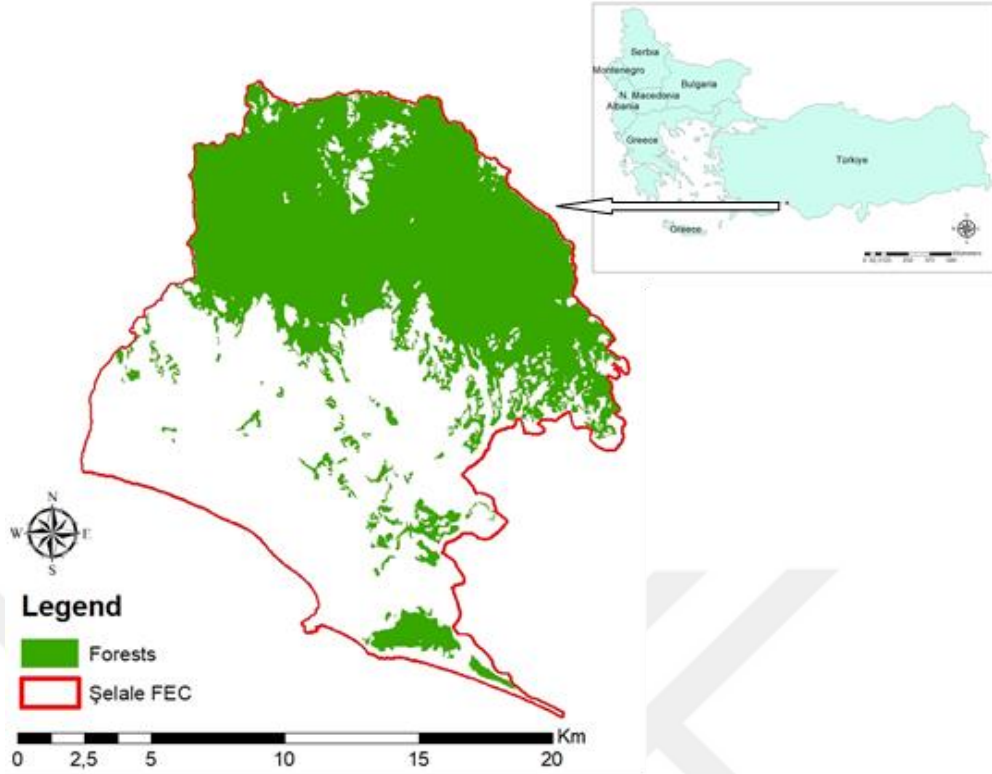
Orman yangınlarının olasılık haritaları, LR ve SVM gibi CBS entegre makine öğrenimi modelleri kullanılarak üretilmiştir. Yöntemin dört aşaması vardır: veri kümesi oluşturma, veri ön işleme, model oluşturma ve orman yangını olasılık haritalaması. Şekil 1 metodolojinin akış şemasını göstermektedir.

3.1. VERİ KÜMESİ OLUŞTURMA

Çalışma, Antalya Orman Bölge Müdürlüğü sınırlarında yer alan Şelale OÇEM'de gerçekleştirilmiştir. Şelale OÇM yaklaşık 25000 ha olup, %48'i (12218 ha) orman alanlarıyla kaplıdır (Şekil 3.2). Baskın tür Brutian çamı (*Pinus brutia* Ten.) olmakla birlikte alanda Sedir, Taş çamı, Defne ve Keçiboynuzu gibi diğer ağaç türleri de bulunmaktadır. Yıllık ortalama yağış miktarı 1050 mm, sıcaklık ise 19 °C'dir (GDM, 2024).



Şekil 3.1. Metodolojinin akış şeması



Şekil 3.2. Çalışma alanı

Veri setinde, orman vejetasyon yapıları, topoğrafik özellikler ve yol ağına yakınlık, yangın olasılığı haritasının oluşturulmasında yangını etkileyen faktörler olarak kabul edilmiştir. İklimsel parametreler, nispeten küçük çalışma alanında tekdüze oldukları ve bitki örtüsü yapıları ve topoğrafik faktörler tarafından kısmen temsil edildikleri varsayılarak doğrudan dahil edilmemiştir. ArcGIS 10.8 kullanılarak yangını etkileyen faktörler için veri katmanları (50 m x 50 m), Şelale FEC'den elde edilen orman yönetimi, topografya ve yol ağı temel haritalarına bağlı olarak oluşturulmuştur.

Her bir yangını etkileyen faktörün sınıfları Çizelge 3.1'de belirtilmiştir. Sekiz sınıf içeren ağaç türü katmanları, orman yönetimi haritalarına dayalı olarak oluşturulmuştur. Taç kapalılığı sınıfları çıplak arazi, seyrek, orta ve yoğun kapalıdır. Çalışma alanında genç (A), orta yaşlı (B), olgun (C) ve daha olgun (D) ağaç evreleri mevcuttu. Topoğrafik harita tabanlı Sayısal Yükseklik Modelinden üretilen eğim haritası beş sınıfa ayrılmıştır (Jaiswal vd., 2002). Yine SYM'den üretilen bakı haritası dokuz bakı sınıfına ayrılmıştır. Son olarak, yol yakınlığı için tampon bölgeler (0-100

m, 100-200 m, 200-300 m, 300-400 m ve 400 m'den fazla) belirlenmiştir (Sivrikaya vd., 2011).

Çizelge 3.1. Analizde dikkate alınan yangını etkileyen faktörlerin sınıfları

Tree Species	Crown Closure	Slope Classes	Aspect
Çz (Brutian pine)	Bare-land (0-10%)	Gentle (0-5%)	Flat
ÇzS (Brutian pine, Cedar)	Sparse (11-40%)	Low (5-15%)	N
ÇzDf (Brutian pine, Daphne)	Moderate (41-70%)	Medium (15-25%)	NE
Çf (Stone pine)	Dense (0>70%)	High (25-35%)	E
Df (Daphne)		Steep (>35%)	SE
DfÇz (Daphne, Brutian pine)	Tree Stages	Proximity to Roads	S
HrÇz (Carob, Brutian pine)	A-Young	0 - 100 m	SW
OT (Forest soil)	B-Middle-aged	100 m - 200 m	W
	C-Maturing	200 m - 300 m	NW
	D-Mature	300 m - 400 m	
	0-Degraded areas	> 400 m	

2001-2021 yılları arasında 230 orman yangını gösteren “yangın noktaları” adlı dijital katman, Şelale OGM'den elde edilen mekânsal verilere dayanarak oluşturulmuştur. Ardından, ArcGIS 10.8'de orman alanlarında “yangın olmayan noktaları” temsil eden yeni bir katman oluşturulmuştur. Bu süreçte öncelikle yangın noktalarının etrafında 500 m yarıçaplı dairesel tamponlar oluşturulmuş ve bu tamponlar orman katmanından elenerek rastgele “yangın olmayan noktalar” ile “yangın noktalarının” kesişmesi engellenmiştir (Tan ve Feng, 2023). Son olarak, “Veri Yönetimi” altındaki “Örnekleme” aracı kullanılarak bu katmandan 247 rastgele nokta (yangın noktası olmayan) oluşturulmuştur.

3.2 VERİ ÖN İŞLEME

Veri ön işleme, veri analitiği ve makine öğrenimi projelerinin temel bir aşamasıdır. Ham verinin temizlenmesi, dönüştürülmesi ve düzenlenmesi veriyi daha anlamlı hale getirir ve analiz edilmesini sağlar. Doğru veri ön işleme, başarılı sonuçlar elde etmenin

anahtarıdır. Veri analitiği ve makine öğrenimi projeleri, veriye dayalı kararlar almak için çeşitli sektörlerde kullanılmaktadır. Ancak ham veriler genellikle düzensiz, eksik veya hatalı olabilir. Bu nedenle, bu verilerin doğru şekilde hazırlanması ve işlenmesi son derece önemlidir.

ML modelinin güvenilir, doğru ve başarılı tahminler yapabilmesi için aşağıdaki adımlar uygulanmıştır. Özniteliklerin yaklaşık yüzde yirmisi boştu ve veri setinden çıkarılmıştır. ML modelleri oluşturmak için hem girdi hem de çıktı değişkenlerinin sayısal olması gerekmektedir. Orman yangını veri kümesinde, yangın durumunu (yangın var/yok) gösteren kategorik bağımlı değişken (çıktı) ikili sayısal forma dönüştürülmüştür. Bir ile etiketlenen kategorik bağımlı değişken yangın noktalarını temsil ederken, sıfır ile etiketlenen veriler yangın noktalarının olmadığını temsil etmektedir. Daha sonrasında, yangını etkileyen faktörlerin sayısal değerleri ArcGIS 10.8'de "Spatial Analysis" altındaki "Extract Multi Values to Points" aracı kullanılarak tüm noktalara (yani yangın noktaları ve yangın olmayan noktalar) atanmıştır. Son olarak, model oluşturulmadan önce veri setine Min-Max Normalizasyonu uygulanmıştır. Min-Max Normalizasyonu, veri setindeki özelliklerin (değişkenlerin) belirli bir aralığa ölçeklenmesini sağlar. Genellikle bu aralık 0 ile 1 arasında olur, ancak bazı durumlarda farklı aralıklar da kullanılabilir. Bu normalizasyon yöntemi, özelliklerin minimum ve maksimum değerlerini temel alarak her bir değeri bu aralık içinde yeniden ölçeklendirir. Bu sayede, veriler aynı ölçeğe getirilir ve modelin her bir özelliği eşit derecede değerlendirebilmesi sağlanır ve aşağıdaki formülle hesaplanır:

$$X_{scaled} = (X - X_{min} / X_{max} - X_{min}) \times (b - a) + a \quad (3.1)$$

Burada:

- X, orijinal veri değeri,
- X_{min}, ilgili özelliğin veri setindeki minimum değeri,
- X_{max}, ilgili özelliğin veri setindeki maksimum değeri,
- A ve b, hedef aralık değerleridir (genellikle 0 ve 1).

3.3 MODEL OLUŞTURMA

Makine öğrenimi, veri analitiği ve yapay zekâ alanında hızla büyüyen bir disiplindir. Bu alandaki en temel görevlerden biri verileri sınıflandırmaktır. Hangi sınıflandırma algoritmasının kullanılacağı projenin özelliklerine ve gereksinimlerine bağlıdır. Veri kümesinin özellikleri ve hedefler doğru algoritmanın seçimini etkiler. En iyi sonuçları elde etmek için seçim sürecinde farklı algoritmalar test edilebilir. Bu çalışmada, orman yangını olasılığını tahmin etmede SVM modeli ve LR modelinin performansları test edilmiştir.

LR, genellikle maksimum olasılık yöntemini kullanarak olasılık tahminini çözmek için geliştirilmiş istatistiksel bir modeldir (Türkan vd., 2023). LR, belirli bir sınıfa ait olma olasılığının (iki sınıf arasında) tahmin edildiği ikili sınıflandırma problemlerini çözmek için yaygın olarak kullanılır. Diğer karmaşık sınıflandırma algoritmalarına göre daha basit ve daha doğru bir ML modelidir (Madhav vd., 2023). Bu çalışmada, bir ikili bağımlı değişken (yangın var = 1 ve yangın yok = 0) ve altı bağımsız değişkene (ağaç türleri, taç kapalılığı, ağaç evresi, eğim, bakı ve yollara yakınlık) dayalı olarak orman yangını olasılığını tahmin etmek için LR modeli seçilmiştir. Lojistik fonksiyon aşağıdaki formülle tanımlanabilir (McGrew ve Monroe, 1993):

$$P_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_P X_P)}} \quad (3.2)$$

P_i = Yangın meydana gelme olasılığı

β = Regresyon katsayısı

X = Bağımsız değişken değeri

Vapnik ve Vapnik (1998) tarafından önerilen SVM modeli, karmaşık sınıflandırma ve regresyon problemlerini çözmek için kullanılan bir denetimli öğrenme modelidir (Shafapourtehrany, 2022). SVM, bir veri kümesini iki homojen alt kümeye ayırmak için daha yüksek boyutlu bir uzayda bir karar sınırı (optimum hiper düzlem) tanımlar (Zaidi, 2023). Veri noktası hiper düzlemin üzerindeyse + 1 olarak sınıflandırılır. Bu çalışmada SVM'nin amacı, veri kümesini yangın noktaları (- 1) ve yangın olmayan (+

1) noktalar olmak üzere iki sınıfa ayırmaktır. Çekirdek fonksiyonları olarak adlandırılan matematiksel fonksiyonlar, özellikle küçük veri kümelerinde doğru ve hassas sınıflandırmalar sağlamak için SVM yöntemine dahil edilir (Patle ve Chouhan, 2013). Çekirdek fonksiyonları, orijinal girdi uzayını yeni bir özellik uzayına eşlemek ve çekirdek uzayında doğrusal olarak ayırt edilemeyen örnekler oluşturmak için kullanılır (Tan ve Feng, 2023). Bu çalışmada, bir dizi deneme yapıldıktan sonra modeli oluşturmak için Radyal Taban Fonksiyonu (RBF) kullanılmıştır.

Modeller, kullanıcıların Python programlama dilinde kod yazmasına olanak tanıyan ücretsiz bir web tabanlı geliştirme ortamı olan Google Colaboratory (Colab) üzerinde oluşturulmuştur (Google, 2018). Modeller çalıştırılıp doğrulukları hesaplandıktan sonra, ikinci veri seti test verisi olarak daha önce eğitilen modele tekrar verilmiştir. Modelin bu noktalar için verdiği tahmin değerleri analiz edilmiş ve Python'da `predict_proba` ile sıfır ile bir arasında bir olasılık değeri elde edilmiştir. `predict_proba` yöntemi, eğitim örneklerinin belirli bir sınıfa ait olma olasılığını hesaplamak için kullanılır (Raschka ve Mirjalili, 2017). Bu işlemin amacı, eğitilen modellerin tüm harita üzerinde elde edilen maksimum nokta sayısı için bir yangın salgını olasılığını tahmin etmesiydi. Böylece tahminlerin daha kolay görselleştirilmesi ve sınıflandırılması sağlanmıştır. Bu çalışmada, elde edilen tahminler Python'da bilimsel hesaplama için gerekli bir kütüphane olan “`matplotlib`” kullanılarak görselleştirilmiştir (Raschka ve Mirjalili, 2017).

Modellerin başarısı, yalnızca kullanılan algoritmaya değil, aynı zamanda hiperparametrelerin doğru bir şekilde seçilmesine de bağlıdır. Hiperparametreler, bir modelin performansını ve genel başarısını belirleyen, model eğitimi sırasında önceden belirlenmiş değerlerdir. Bu nedenle, hiperparametrelerin seçimi ve ayarlanması, makine öğrenmesi sürecinin kritik bir aşamasını oluşturur. Hiperparametreler, modelin öğrenme sürecine doğrudan etki eden ve kullanıcı tarafından önceden belirlenen parametrelerdir. Bu parametreler, modelin karmaşıklığını, öğrenme hızını, katman sayısını, düğüm sayısını, düzenlileştirmenin miktarını ve diğer birçok faktörü kontrol eder. Örneğin, bir yapay sinir ağı modelinde, katman sayısı, her katmandaki nöron sayısı ve öğrenme hızı gibi parametreler hiperparametre olarak kabul edilir. Uygun hiperparametreler seçilmezse, model ya aşırı uyum sağlayarak (overfitting) eğitim

verisine fazla odaklanır ya da yetersiz uyum göstererek (underfitting) genelleme yeteneğini kaybeder. Aşırı uyum, modelin eğitim verisine çok iyi uyum sağladığı ancak yeni verilere karşı başarısız olduğu durumları ifade eder. Yetersiz uyum ise, modelin hem eğitim hem de test verilerine karşı düşük performans gösterdiği durumlardır.

Hiperparametrelerin optimizasyonu, makine öğrenmesi sürecinde en zorlayıcı ve zaman alıcı adımlardan biridir. Farklı hiperparametre kombinasyonları deneyerek, modelin en iyi performansı göstermesi hedeflenir. Bu çalışmada, uygun hiperparametrelerin seçimi için kullanılan en yaygın yöntemlerden biri olan Grid Search yöntemi kullanılmıştır. Grid Search yöntemi önceden belirlenmiş hiperparametre aralıklarını sistematik bir şekilde tarayarak en iyi kombinasyonu bulmayı amaçlamaktadır. Çalışmada, bu yöntem kullanılsa da daha iyi sonuçların alınabilmesi durumunu değerlendirmek için farklı değerler kullanılarak da denemeler yapılmıştır.

3.3.1. Lojistik Regresyon için Model Parametreleri

3.3.1.1. Regülerizasyon Parametresi (C)

Regülerizasyon, modelin aşırı uyumunu (overfitting) engellemek için kullanılan bir tekniktir. Lojistik regresyon modelinde C parametresi, modelin ceza terimini kontrol eder. C'nin düşük olması, daha yüksek ceza anlamına gelir ve modelin daha basit olmasına yol açar. C'nin yüksek olması ise ceza terimini azaltarak modelin daha karmaşık olmasına izin verir. Dengeli bir model için, C parametresi dikkatle seçilmelidir. Bu çalışmada C parametresi “10.0” olarak belirlenmiştir.

3.3.1.2. Solver (Çözücü)

Lojistik regresyon modelinin eğitimi sırasında kullanılan optimizasyon algoritmasını belirler. Yaygın kullanılan çözücüler arasında liblinear, saga, lbfgs, ve newton-cg yer alır. Bu çözücüler, modelin eğitim süresi ve doğruluğu üzerinde etkili olabilir. Verinin boyutuna ve yapısına bağlı olarak, uygun bir çözücü seçmek model performansını artırabilir. Bu çalışmada çözücü “liblinear” olarak belirlenmiştir.

3.3.1.3. Regülarizasyon Türü (L1, L2)

Lojistik regresyonda kullanılan regülarizasyon türü, modelde yer alan özelliklerin (features) seçilmesine yardımcı olabilir. L1 regülarizasyonu, modelde daha fazla sıfır ağırlık oluşturduğu için özellik seçimi ile ilgili durumlarda faydalıdır. L2 regülarizasyonu ise tüm özelliklere küçük, ama pozitif veya negatif değerler atayarak modelin genel doğruluğunu artırabilir. Bu çalışmada regülasyon türü olarak “L2” seçilmiştir.

3.3.1.4. Maksimum Yineleme (max_iter)

Bu parametre, model eğitimi sırasında gerçekleştirilecek maksimum yineleme sayısını belirler. Modelin optimizasyon sürecinde sapma yaşanmaması için yeterli sayıda yineleme yapılması gereklidir. Bu parametrenin çok düşük olması durumunda, modelin tam olarak eğitilememesi söz konusu olabilir. Bu çalışmada, maksimum yineleme değeri “200” olarak belirlenmiştir.

3.3.2 Destek Vektör Makineleri (SVM) için Model Parametreleri

SVM modelini oluştururken dikkate alınması gereken başlıca parametreler şunlardır:

3.3.2.1. C Parametresi (Ceza Parametresi)

SVM modelinde C parametresi, yanlış sınıflandırmaların cezalandırılması derecesini belirler. C'nin düşük bir değeri, modelin daha geniş bir marja sahip olmasına izin verir, bu da daha fazla yanlış sınıflandırmaya neden olabilir ancak modelin genel olarak daha basit olmasını sağlar. C'nin yüksek bir değeri ise modelin daha az hata yapmasına olanak tanır, fakat modelin aşırı uyum riskini artırabilir. C parametresi, veri setine ve problem türüne bağlı olarak dikkatle optimize edilmelidir. Bu çalışmada C parametresi “1” olarak seçilmiştir.

3.3.2.2. Kernel (Çekirdek) Fonksiyonu

SVM modeli, verilerin doğrusal olarak ayrılmadığı durumlarda çekirdek fonksiyonları kullanarak verileri daha yüksek boyutlu bir uzaya dönüştürür. Yaygın olarak kullanılan çekirdek fonksiyonları arasında lineer, polinomial, radyal baz fonksiyonu (RBF) ve sigmoid yer alır. Çekirdek fonksiyonunun doğru seçimi, modelin doğruluğunu ve genel performansını önemli ölçüde etkileyebilir. Bu çalışmada, Kernel fonksiyonu “lineer” olarak seçilmiştir.

3.3.2.3. Gamma (γ) Parametresi

Gamma, çekirdek fonksiyonunun etkisini kontrol eden bir parametredir ve genellikle RBF, polinomial ve sigmoid çekirdek fonksiyonlarıyla birlikte kullanılır. Gamma'nın yüksek bir değeri, her veri noktası için daha yakın komşularına odaklanırken, düşük bir gamma değeri daha geniş bir alandaki verileri dikkate alır. Bu çalışmada, Gamma değeri “scale” olarak belirlenmiştir.

3.3.2.4. Sınıf Ağırlıkları (class_weight)

Dengesiz veri setlerinde (örneğin, bir sınıfın diğerine göre daha fazla örneğe sahip olduğu durumlar), sınıf ağırlıkları parametresi kullanılarak modelin belirli sınıflara daha fazla önem vermesi sağlanabilir. Bu, özellikle orman yangını risk haritası gibi dengesiz veri dağılımlarına sahip problemlerde faydalıdır. Bu çalışmada, sınıf ağırlıkları parametresi “balanced” olarak belirlenmiştir.

3.4 ORMAN YANGINI OLASILIK HARİTALAMASI

Her iki model için üretilen tahmin değerleri ayrı ayrı tablolştırılmış ve “*.csv” dosyalarına dönüştürülmüştür. Daha sonra ArcGIS 10.8 kullanılarak veri enterpolasyonunda “Kriging” yöntemi uygulanmıştır. “Kriging”, yakındaki noktaların değerlerini kullanarak düzenli bir ızgara üzerinde bir yüzeyi tahmin etmeye yardımcı olan güçlü bir araçtır (Shad vd., 2009). Tüm nokta değerleri bir yüzeye dönüştürüldükten sonra, ArcGIS 10.8'de yangın olasılığı değerleri belirlenen beş

seviyeye kategorize edilerek orman yangını olasılık haritası oluşturulmuştur. Daha sonra, yangın olasılık haritalarının doğruluğu Alıcı İşletim Karakteristiği (ROC) eğrisi yöntemi kullanılarak değerlendirilmiştir. Özgüllük ve duyarlılık arasındaki ilişkiyi yorumlamak için kullanılan grafiksel bir teknik olan ROC eğrisi, yangın olasılığı haritalarının doğruluğunu değerlendirmek için başarıyla kullanılmıştır (Güney vd., 2016; Sivrikaya ve Küçük, 2022). Modelin ROC eğrisi sol üste doğru kaydıkaç model daha etkili olmaktadır. Tahmin yeteneğinin başarısı AUC değeri ile gösterilir. İdeal olarak, optimum performans için değeri mümkün olduğunca 1,0'a yakın olmalıdır. AUC değeri 0,5 - 0,6 (zayıf), 0,6 - 0,7 (orta), 0,7 - 0,8 (iyi), 0,8 - 0,9 (çok iyi) ve 0,9 - 1,0 (mükemmel) olmak üzere beş sınıfa ayrılır (Gheshlaghi vd., 2020). Bu çalışmada, yangın olasılığı haritasının doğruluğunu test etmek için ArcGIS 10.8'de "ROC_ArcSDM" uzantısı kullanılmıştır (Demir ve Akay, 2024).

BÖLÜM 4

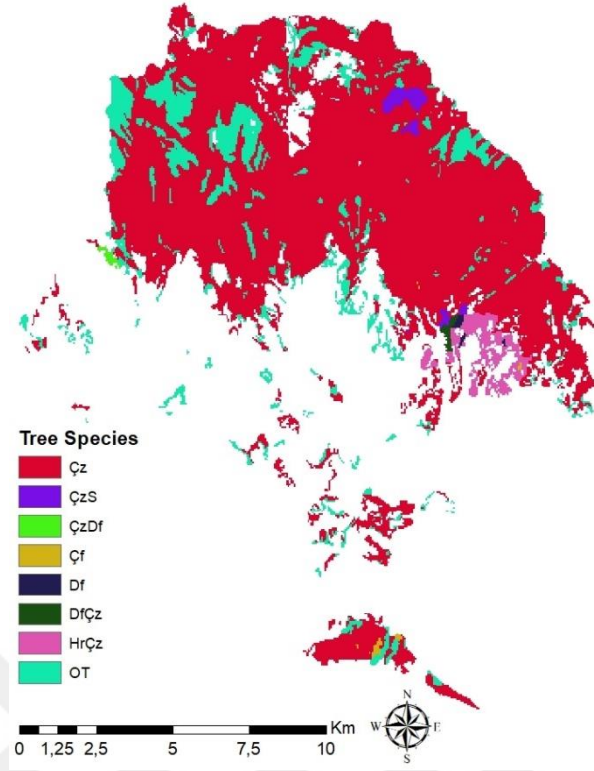
BULGULAR VE TARTIŞMA

4.1. YANGINI ETKİLEYEN FAKTÖRLER

Sonuçlar, çalışma alanında sekiz tür bulunduğunu göstermiştir (Çizelge 4.1, Şekil 4.1). En yaygın bitki türü, hızlı tutuşma kabiliyeti nedeniyle orman yangınlarına karşı oldukça hassas olan Brutian çamıdır (%81,93) (Akay ve Erdoğan, 2017; Güney vd., 2022).

Çizelge 4.1. Ağaç türlerinin alansal dağılımı

Tree Species	Area (%)
Çz (Brutian pine)	81.93
ÇzS (Brutian pine, Cedar)	0.92
ÇzDf (Brutian pine, Daphne)	0.16
Çf (Stone pine)	0.18
Df (Daphne)	0.14
DfÇz (Daphne, Brutian pine)	0.21
HrÇz (Carob, Brutian pine)	2.34
OT (Forest soil)	14.12



Şekil 4.1. Ağaç türlerinin haritası.

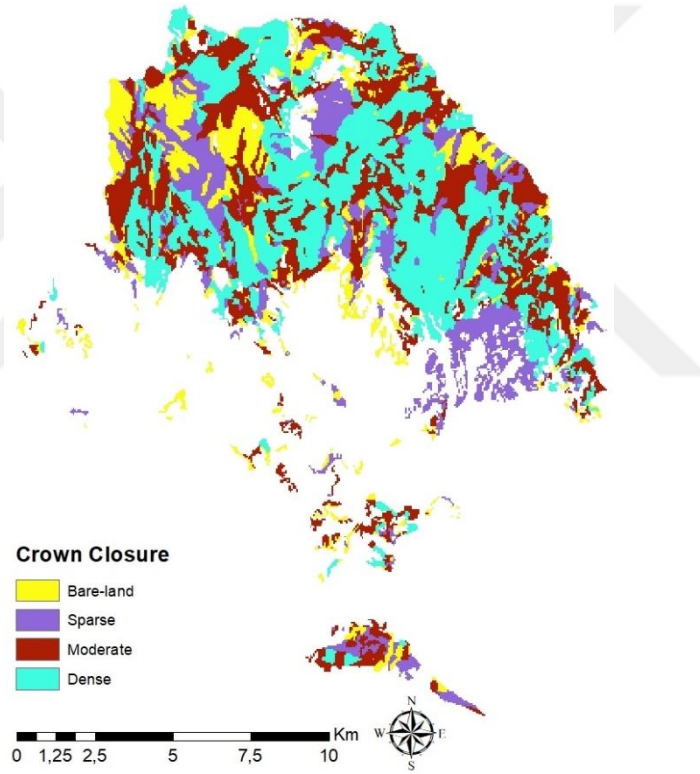
Çalışma alanının yaklaşık %39'unun sık ormanlarla kaplı olduğu tespit edilmiştir (Çizelge 4.2, Şekil 4.2). Önceki çalışmalar orman alanlarında taç kapalılığı arttıkça yangın olasılığının da arttığını ortaya koymuştur (Durmaz, 2004; Küçük vd., 2009). Ağaç evreleri haritası, ormanın %31,15'inin orta yaşlı ve %14,12'sinin genç meşcerelerde olduğunu göstermiştir (Çizelge 4.3, Şekil 4.3). Sağlam ve diğerleri (2008) yangın olasılığının genç meşcerelerde olgun ve olgunlaşmakta olan meşcerelere kıyasla önemli ölçüde daha yüksek olduğunu bildirmiştir.

Eğim haritasına göre, çalışma alanının yarısı dik ve yüksek eğimli arazilerde, %17,73'ü ise orta eğimli arazilerde yer almaktadır (Çizelge 4.4, Şekil 4.4). Benzer çalışmalardan elde edilen sonuçlar, sarp arazilerde bulunan orman meşcerelerinde yangın olasılığının çok yüksek olduğunu göstermiştir (Jaiswal vd., 2002). Çalışma alanında ormanların yaklaşık %37'si güney bakılarda (S, SE ve SW) yer alırken, bunu doğu (%16,52) ve batı bakılar (%16,14) takip etmektedir (Çizelge 4.5, Şekil 4.5). Bakının önemli bir topografik faktör olduğu ve düşük nem ve yüksek sıcaklık nedeniyle güney bakılardaki

orman meşcerelerinde yangın olasılığının oldukça yüksek olduğu bildirilmiştir (Lin ve Sergio, 2009).

Çizelge 4.2. Taç kapalılığının alansal dağılımı

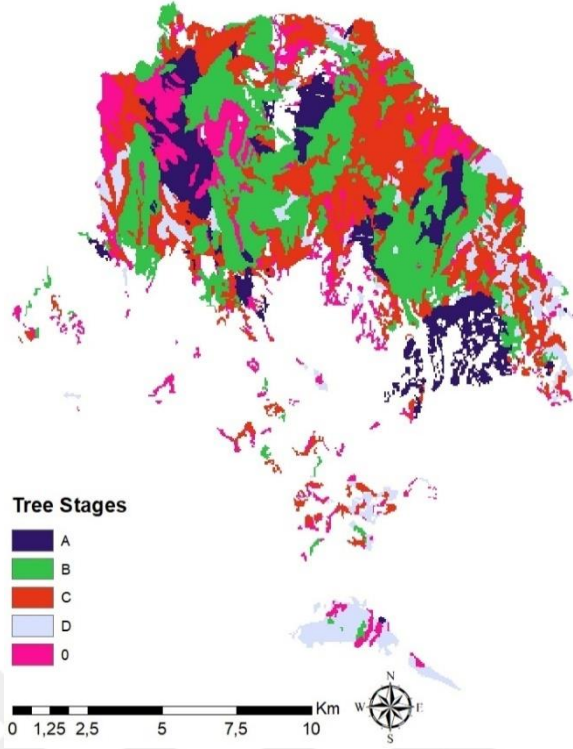
Crown Closure	Area (%)
Bare-land (0-10%)	14.12
Sparse (11-40%)	18.01
Moderate (41-70%)	29.18
Dense (0>70%)	38.69



Şekil 4.2. Taç kapalılık haritası.

Çizelge 4.3. Ağaç evresinin alansal dağılımı

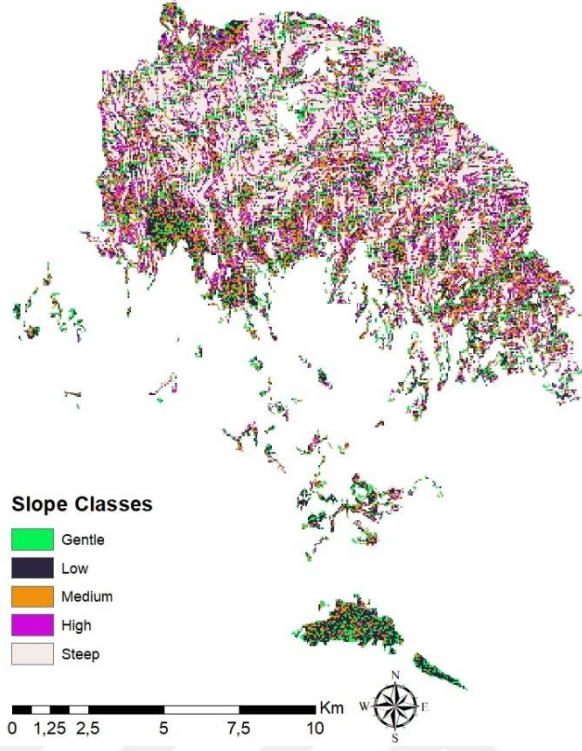
Tree Stages	Area
A-Young	14.03
B-Middle-aged	31.15
C-Maturing	30.53
D-Mature	10.17
0-Degraded	14.12



Şekil 4.3. Ağaç aşamalarının haritası.

Çizelge 4.4. Eğim sınıflarının alansal dağılımı

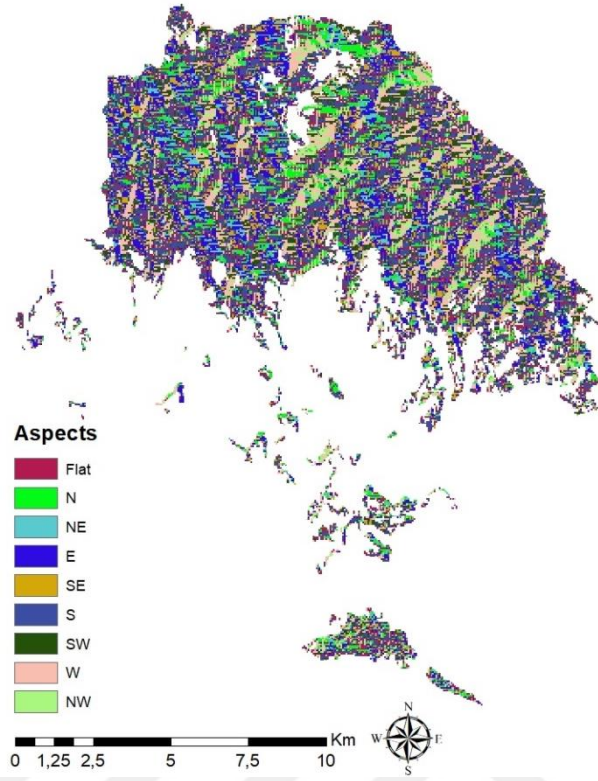
Slope Classes	Area (%)
Gentle (0-5%)	12.66
Low (5-15%)	17.71
Medium (15-25%)	17.73
High (25-35%)	16.55
Steep (>35%)	35.35



Şekil 4.4. Eğim sınıflarının haritası.

Çizelge 4.5. Bakı sınıflarının alansal dağılımı.

Aspect	Area (%)
Flat	11.93
N	9.77
NE	4.75
E	16.52
SE	7.53
S	20.23
SW	8.81
W	16.14
NW	4.32

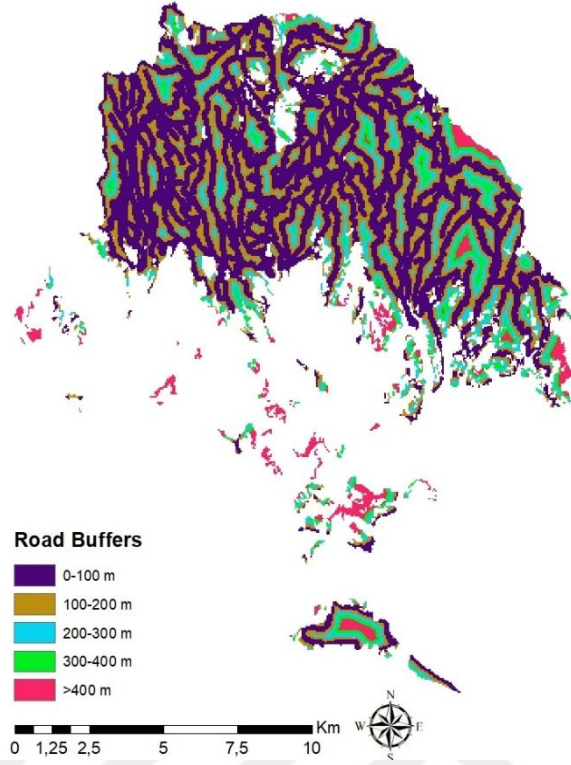


Şekil 4.5. Bakı haritası.

Ormanların yarısından fazlasının yol ağlarının 100 m içinde yer aldığı ve %28,08'inin yollardan 100 ila 200 m uzakta olduğu tespit edilmiştir (Çizelge 4.6). Önceki çalışmalar, yol ağları boyunca insan faaliyetlerinin yüksek olması nedeniyle orman alanlarının yollara yakın olduğu yerlerde orman yangını olasılığının arttığını göstermiştir (Güney vd., 2016; Jaiswal vd., 2002). Yollar boyunca oluşturulan tampon bölgeler Şekil 4.6'da gösterilmektedir.

Çizelge 4.6. En boy sınıflarının alansal dağılımı.

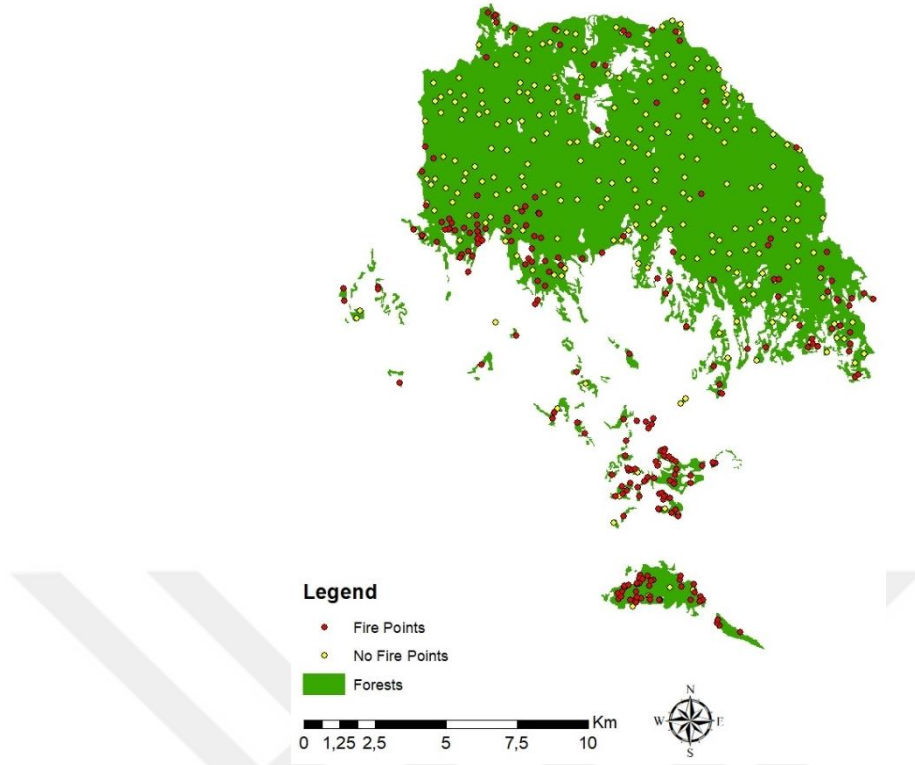
Proximity to Roads	Area (%)
0 - 100 m	51.36
100 m - 200 m	28.08
200 m - 300 m	11.53
300 m - 400 m	4.43
> 400 m	4.60



Şekil 4.6. Yol tampon bölgesi haritası.

4.2. ML MODELLERİNİN SONUÇLARI

Orman yangını veri seti, yangın durumu (yangın var/yok) veri noktaları ve bu noktalara atanan yangını etkileyen faktörlerin sayısal değerleri temel alınarak oluşturulmuştur. Yangın durumu veri noktalarının mekânsal dağılımı Şekil 4.7'de gösterilmektedir. Şekil 4.8, kimlik numarası, X ve Y koordinatları, altı yangını etkileyen faktör ve yangın durumu dahil olmak üzere 10 öznitelik içeren orman yangını veri kümesinin bir bölümünü göstermektedir.



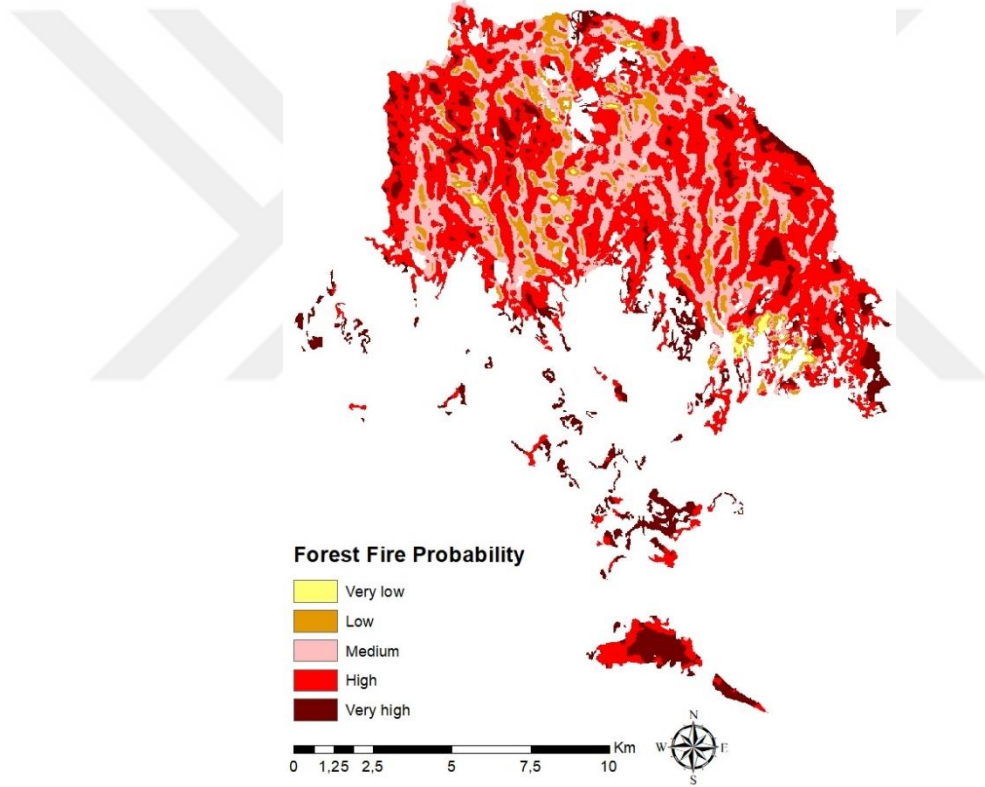
Şekil 4.7. Yangın durumunu (yangın var/yangın yok veri noktaları).

index	id	Xcoord	Ycoord	species	closure	age	slope	aspect	distance	output
0	7758	364293.618815	4080079.32964	8	2	1	3	6	1	1
1	7763	363846.669852	4078291.70214	1	2	1	4	9	2	1
2	7767	353789.286423	4081992.49227	1	3	3	4	6	2	1
3	7770	353986.605237	4082476.69506	1	3	3	2	4	1	1
4	7775	353861.151555	4082064.38559	1	3	3	2	5	1	1
5	7867	362212.885108	4076994.94725	1	3	3	5	5	3	1
6	8013	357225.80169	4087104.36234	1	4	1	1	1	1	1
7	8039	351754.337005	4084485.59865	1	3	4	5	6	2	1
8	8041	353716.311053	4082336.3895	1	3	3	3	6	1	1
9	8065	359279.010544	4069375.8109	1	3	4	2	8	2	1
10	8091	357486.355464	4075263.54585	1	3	3	2	4	5	1
11	8153	360270.289412	4074494.16689	1	3	4	3	3	5	1
12	8191	366889.388207	4080024.55422	9	1	5	2	7	2	1
13	8220	356546.791472	4081459.57251	1	3	3	5	9	2	1
14	8226	367005.069043	4077267.24676	1	3	4	2	9	1	1
15	8274	366819.990991	4078406.45161	1	2	1	4	6	2	1

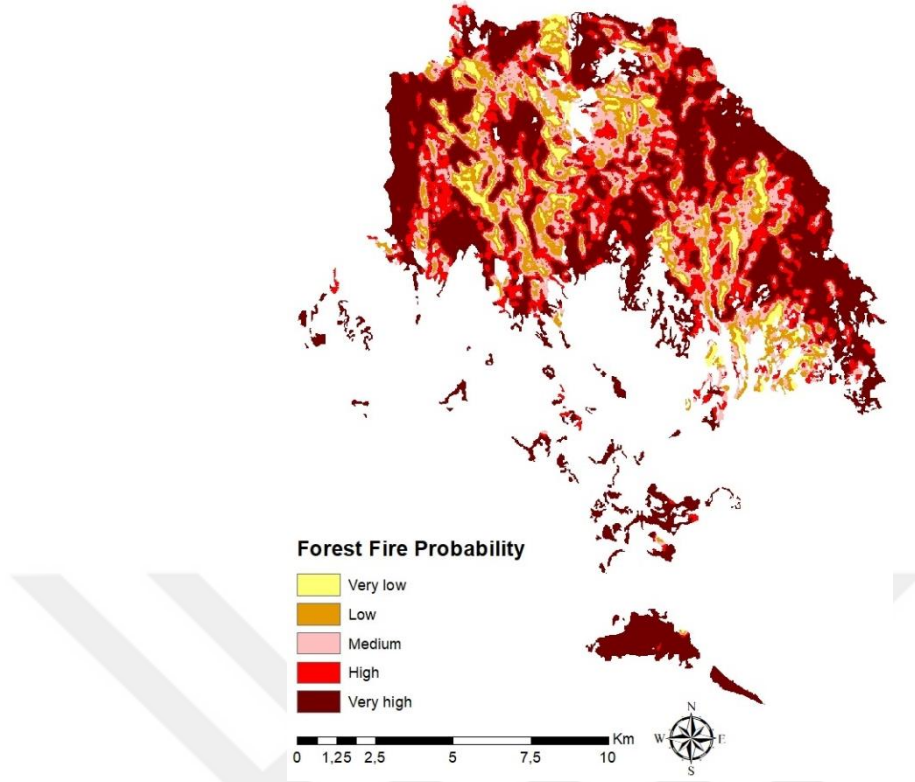
Şekil 4.8. Modellerde kullanılan veri kümesinin bir kısmı.

Model oluşturmada, orijinal örnek veri noktalarının %70'i model oluşturma ve %30'u model testi için rastgele seçilmiştir (Tan ve Feng, 2023). ML modellerinin tahminleri, olasılık seviyelerine (yani çok yüksek, yüksek, orta, düşük ve çok düşük) bölünmüş olasılık haritası üzerindeki 47.086 nokta için yapılmıştır. Modeller bu noktalara sıfır ile bir arasında bir yangın olasılığı değeri atamış ve daha sonra bunlar bir yüzey verisine dönüştürülmüştür. LR ve DVM modelleri tarafından oluşturulan orman yangını olasılık haritaları Şekil 4.9 ve a.10'da gösterilmektedir.

Her iki model için yangın olasılığı seviyelerinin dağılımı Çizelge 4.7'de gösterilmektedir. LR modeli tarafından oluşturulan olasılık haritasına göre, çalışma alanının %50'sinden fazlası yüksek olasılık seviyesindeyken, sadece %12'si çok yüksek seviyededir. DVM modelinden elde edilen sonuçlara göre, çalışma alanının %46'sı çok yüksek olasılık düzeyinde, %22'si ise yüksek düzeyde sınıflandırılmıştır. Modellerin doğruluğu LR ve DVM modelleri için sırasıyla 0.64 ve 0.71 olarak hesaplanmıştır. Rodrigues ve Riva (2006) tarafından yapılan önceki bir çalışmada, insan kaynaklı orman yangınlarının araştırılması için ML modellerinin (LR, SVM, RF) ortalama performans doğruluğu %74'tür.



Şekil 4.9. LR modeliyle tahmin edilen orman yangını olasılık haritası.



Şekil 4.10. SVM modeliyle tahmin edilen orman yangını olasılık haritası.

Çizelge 4.7. Yangın olasılık seviyelerinin alansal dağılımı

Fire Probability Levels	Area (%)	
	LR	SVM
Very high	12	46
High	53	22
Medium	28	17
Low	6	11
Very low	1	4

Python'da kullanılan kod, özellikle listeler veya veri çiftleri üzerinde işlem yapmak için kullanılan bir döngü yapısıdır. Bu kod, iki farklı listeyi paralel olarak işlemek için “zip ()” fonksiyonunu kullanır ve her bir özellik adını ve bir katsayı çiftini ekrana yazdırır. Bu kod parçası kullanılarak LR modelinden elde edilen yangını etkileyen faktörlerin önemi Çizelge 4.8'de belirtilmiştir. Ağaç evresi, yollara yakınlık ve bakı faktörlerinin orman yangını çıkma olasılığı üzerinde en yüksek etkiye sahip olduğu bulunmuştur. Benzer çalışmalarda da (Güney vd., 2016) yangını etkileyen en önemli faktörler bakı, ağaç evresi ve insan faaliyetleri ile ilgili nesnelere (tarım alanları, yollar, yerleşim alanları gibi) yakınlık olmuştur. Şelala FEC'de, meşcerelerin yaklaşık %75 'i

genç, orta yaşlı ve olgunlaşmış meşcerelerden oluşmakta olup, bu da ağaç evresinin yangın çıkması durumunda en etkili faktör olduğunu açıklamaktadır. Ayrıca, önceki çalışmalar ağaç evresinin yangın başlangıç noktalarını ve davranışını belirleyen en önemli faktörlerden biri olduğunu göstermiştir (Durmaz, 2004; Kavlak vd., 2021).

Çizelge 4.8. LR modeliyle seçilen değişkenlerin önemi.

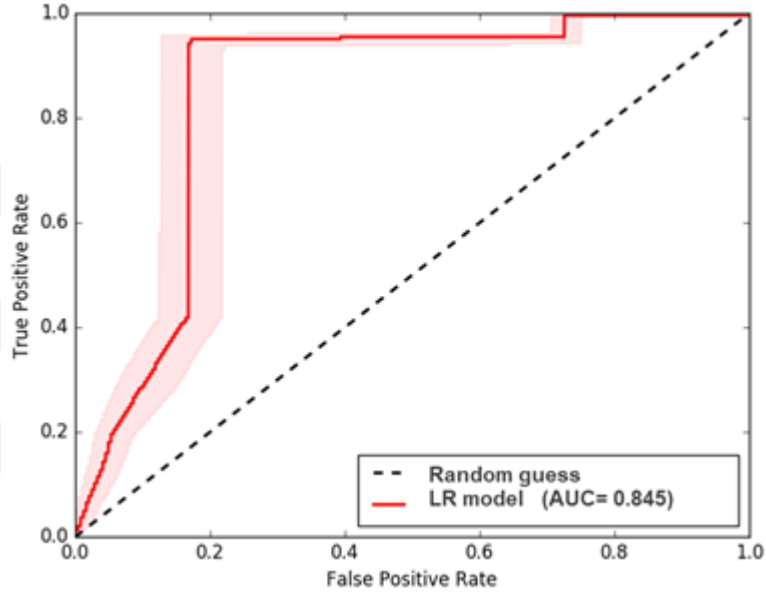
Fire Influencing Factors	LR	
	Coefficient	Odds Ratio
Tree Species	-1.643	0.193
Crown Closure	-1.132	0.322
Tree Stage	1.451	4.267
Slope	-1.430	0.239
Aspect	0.199	1.220
Proximity to Roads	1.387	4.003

Yol ağlarına yakınlık, yangını etkileyen en etkili ikinci faktör olmuştur. Çalışma alanındaki yol ağlarının yerel halk ve kamu tarafından oldukça sık kullanıldığı, bunun da yol ağları boyunca potansiyel insan faaliyetlerini artırdığı ve bölgede daha yüksek yangın olasılığına yol açtığı öngörülmektedir (Fekir vd., 2022). Ayrıca, çalışma alanındaki yol yoğunluğu (yaklaşık 32 m/ha) Türkiye' deki mevcut koşullar göz önüne alındığında nispeten yüksektir. Son olarak, bakı yanıcı maddelerin güneş ışığına maruz kalma miktarında önemli bir rol oynadığı için yangını etkileyen en etkili üçüncü faktör olmuştur (Xiangwei vd., 2011).

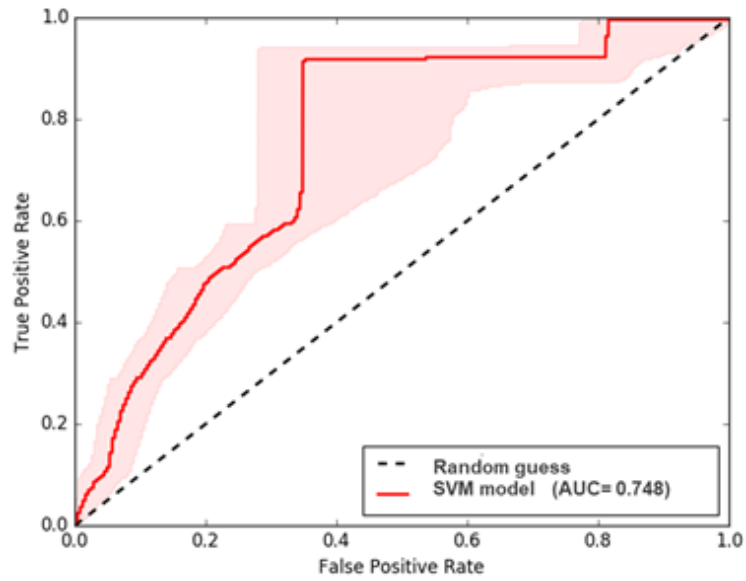
Modellerin doğrulama aşamasında, yangın olasılık haritalarının doğruluğu, doğal tehlike çalışmalarında en önemli doğrulama parametreleri arasında yer alan AUC değeri kullanılarak değerlendirilmiştir (Shafapourtehrany, 2022).

Jaiswal vd., 2002, çalışmalarında uydu görüntüleri ve Coğrafi Bilgi Sistemleri (GIS) ile makine öğrenmesi algoritmalarını birleştirerek orman yangını risk zonları oluşturmuştur. AUC ve diğer değerlendirme metrikleri kullanılarak farklı algoritmalar karşılaştırılarak algoritmaların AUC değerleri 0.70–0.90 aralığında değişmektedir ve orman yangını tahmininde kabul edilen bir doğruluk seviyesi sağlanmıştır. Bu çalışmada, LR modelinin 0,845 AUC ile en yüksek performansı elde ettiği, bunu 0,748

AUC ile DVM modelinin izlediği görülmüştür (Şekil 4.11, 4.12). Böylece DVM modeli kullanılarak geliştirilen yangın olasılık haritası oldukça güvenilir bulunurken LR modeli de iyi bir tahmin performansı sağlamıştır. Birkaç ML modeli kullanılarak orman yangını olasılık haritalarının geliştirildiği bir çalışmada, LR modeli SVM modeline kıyasla nispeten daha iyi tahmin performansına sahip olmuştur (Rihan vd., 2023).



Şekil 4.11. LR modeli ile için ROC eğrisi.



Şekil 4.12. SVM modeli için ROC eğrisi.

BÖLÜM 5

SONUÇLAR VE ÖNERİLER

Dünyanın birçok yerinde, özellikle de yazları çok yüksek hava sıcaklıkları ve düşük bağıl nem oranına sahip Akdeniz havzasında yer alan ülkelerde, son derece yıkıcı orman yangınları küresel ısınmanın kaçınılmaz bir sonucu haline gelmiştir. Bu nedenle, orman yangınlarının doğru bir şekilde meydana gelme olasılığı, orman yangını yönetiminin önemli bir bileşenidir. Orman yangını olasılık haritaları, orman yöneticilerine mevcut tüm kaynakların önceden en iyi şekilde organize edilmesi için yararlı bilgiler sağlar. Bölgesel yangın olasılık haritaları ile yangın riski yüksek alanlara daha fazla odaklanılacak ve daha etkili yangınla mücadele önlemleri alınacaktır. CBS teknikleri ile entegre edilmiş makine öğrenimi modelleri, doğal tehlikeler için doğru olasılık haritaları geliştirmek için etkin bir şekilde kullanılabilir. Bu çalışmada, Akdeniz'in Antalya kentinde bulunan Şelale OÇEM'deki ormanların orman yangını olasılık haritalarını oluşturmak için CBS teknikleriyle entegre edilmiş iki iyi bilinen ML modeli (LR ve SVM) kullanılmıştır. Sonuçlar, her iki model tarafından geliştirilen yangın olasılığı haritalarının oldukça güvenilir olduğunu (LR: %84,5; DVM: %74,8) ve ormanların %65'inden fazlasının “çok yüksek/yüksek” yangın olasılığı seviyeleri olarak sınıflandırıldığını göstermiştir. Yangın olasılığı yüksek orman alanlarının bilinmesi, tüm yönetim kademelerinin bilinçlenmesi ve hedefe yönelik proaktif önlemlerin alınması yangınla mücadele faaliyetlerinde oldukça destekleyici bir rol oynamaktadır. Bu çalışmada sunulan modeller (LR ve SVM) doğru sonuçlar vermiştir, ancak metodolojiyi geliştirmek için hala bazı potansiyeller bulunmaktadır. Gelecekteki çalışmalarda, iklim faktörleri ve yerleşim alanları ve tarım alanları da dahil olmak üzere bazı yerlere yakınlık gibi yangını etkileyen ek faktörler modellere dahil edilmelidir. Ayrıca, diğer etkili ML modellerinin orman yangını olasılığı haritalamasındaki performansı daha büyük çalışma alanları için değerlendirilebilir.

KAYNAKLAR

Abid, F. 2021. A survey of machine learning algorithms-based forest fires prediction and detection systems, *Fire Technol*, 57:559–590.

Adab, H., Kanniah, K. D., & Solaimani, K. (2013). Modeling forest fire risk in the northeast of Iran using remote sensing and GIS techniques. *Natural Hazards*, 65(3), 1685-1708.

Afifi, A.A., Clark, V. 1990. Computer-aided multivariate analysis (2nd ed.). Van Nostrand Reinhold Co.

Akay, A.E., Erdoğan, A. 2017. GIS-based multi-criteria decision analysis for forest fire risk mapping. In 4th International Geoadvances Workshop-Geoadvances 2017: ISPRS Workshop on Multi-Dimensional & Multi-Scale Spatial Data Modeling. October 14-15. Safranbolu, Türkiye.

Arpaci, A., Malowerschnig, B., Sass, O., Vacik, H. 2014. Using multi variate data mining techniques for estimating fire susceptibility of Tyrolean forests. *Appl. Geogr.* 53:258–270.

Atmaca, İ., Derakhshandeh, M., Işık-Pekkan, Ö., Özenen-Kavlak, M, Tunca, Y.S., Çabuk, S.N. 2022. Forest Fire Risk Modeling Using Logistic Regression and Geographic Information Systems: A Case Study in Muğla - Mila. *Journal of Natural Hazards and Environment*. 8:66-75.

Bar, S., Parida, B. R., Pandey, A. C., Shankar, B. U., Kumar, P., Panda, S. K., & Behera, M. D. (2023). Modeling and prediction of fire occurrences along an elevational gradient in Western Himalayas. *Applied Geography*, 151. <https://doi.org/10.1016/j.apgeog.2022.102867>.

Bonora, L., Conese, C., Marchi, E. 2013. Wildfire occurrence: integrated model for risk analysis and operative suppression aspects management. *Am. J. Plant Sci.* 4:705–710.

Carmel, Y., Paz, S., Jahashan, F., Shoshany, M. 2009. Assessing fire risk using Monte Carlo simulations of fire spread, *Forest Ecology and Management*, 257(1):370-377.

Cortes, C., Vapnik, V., & Saitta, L. (1995). Support-Vector Networks Editor. In *Machine Learning* (Vol. 20). Kluwer Academic Publishers.

Cortez, P., & Morais, A. (2007). A data mining approach to predict forest fires using meteorological data. Proceedings of the 13th EPIA Conference on Artificial Intelligence.

Çanakçioğlu, H. 1993. Forest Conservation. Istanbul University, Faculty of Forestry Press. Istanbul, Türkiye.

Demir, A., Akay, A.E. 2024. Forest fire risk mapping using GIS based Analytical Hierarchy Process approach, *European Journal of Forest Engineering*, 10(1): 1-9.

Duine, G.J., Carvalho, L.M.V., Jones, C. 2022. Mesoscale patterns associated with two distinct heatwave events in coastal Santa Barbara, California, and their impact on local fire risk conditions, *Weather and Climate Extremes*, 37: 100482.

Durmaz, B.D. 2004. Effects of stand characteristics on fire potential, Karadeniz Technical University, Science Institute, MSc. Thesis, Trabzon. 53 p.

EFFIS, 2021. Current Statistics Portal: European Forest Fire Information System (EFFIS) Estimates. Available at <https://effis.jrc.ec.europa.eu/apps/effis.statistics/estimates>

Erkan, N., Akay, A.E., Bilici, E., Uçar, Z., Güney, C.O. 2024. Estimating the economic value of timber products potentially saved from wildfires by improving forest road standards. *International Journal of Forest Engineering*, DOI: 10.1080/14942119.2024.2317106.

Fekir, Y., Hamadouche, M.A., Anteur, D. 2022. Integrated approach for the assessment of forest fire risk and burn severity mapping using GIS, AHP method, and Google Earth Engine in Western Algeria. *Euro-Mediterranean J. Environ. Integ.*, (2022) 1-14.

Gao, X., Fei, X., Xie, H. 2011. Forest fire risk zone evaluation based on high spatial resolution RS image in Liangyungang Huaguo Mountain Scenic Spot. *International Conference on Spatial Data Mining and Geographical Knowledge Service*, pp. 593-596. China: Fuzhou.

GDF, 2022. General Directorate of Forestry, Forestry Statistics 2022, <https://www.ogm.gov.tr/tr/e-kutuphane/resmi-istatistikler>. Last visit: 29-01-2024.

GDM, 2024. General Directorate of Meteorology, Weather Forecast, <https://www.mgm.gov.tr/eng/forecast-cities.aspx>

Gheshlaghi, H. A., Feizizadeh, B., Blaschke, B. 2020. GIS-based forest fire risk mapping using the analytical network process and fuzzy logic. *Journal of Environmental Planning and Management*, 63(3):481-499.

Google, 2024. Colaboratory: frequently asked questions, Last visit: 02-03-2024. <https://research.google.com/colaboratory/faq.html>

Güney, C.O., Özkan, K., Şentürk, Ö. 2016. Modelling of spatial prediction of fire ignition risk in the Antalya-Manavgat district. *Journal of the Faculty of Forestry Istanbul University*, 66(2): 459-470.

Güney, C.O., Sarı, A., Cekim, H.O., Küçüksille, E.U., Sentürk, Ö., Gülsoy, S., Özkan, K. 2022. An advanced approach for leaf flammability index estimation. *Int J Wildland Fire*, 31(3): 277-290.

Jaiswal, R.K., Mukherjee, S., Raju, K.D., Saxena, R. 2002. Forest fire risk zone mapping from satellite imagery and GIS, *International Journal of Applied Earth Observation and Geoinformation*, 4(1):1-10.

Justice, C.O., Giglio, L., Korontzi, S., et al. (2002). The MODIS fire products. *Remote Sensing of Environment*, 83(1-2), 244–262.

Kala, C.P. 2023. Environmental and socioeconomic impacts of forest fires: A call for multilateral cooperation and management interventions, *Natural Hazards Research*, 3(2): 286-294.

Kalantar, B., Ueda, N., Idrees, M. O., Janizadeh, S., Ahmadi, K., & Shabani, F. (2020). Forest fire susceptibility prediction based on machine learning models with resampling algorithms on remote sensing data. *Remote Sensing*, 12(22), 1–24. <https://doi.org/10.3390/rs12223682>

Kaur, M., Gulati, H., Kundra, H. 2014. Data mining in agriculture on crop price prediction: techniques and applications. *International Journal of Computer Applications*, 99(12): 1-3.

Kavlak, M.O., Çabuk, S.N., Çetin, M. 2021. Development of forest fire risk map using geographical information systems and remote sensing capabilities: Ören case, *Environmental Science and Pollution Research*, 28:33265–33291.

Khan, N., Yaqoob I., Hashem, I.A.T., Inayat, Z., Kamaleldin, W., Ali, M., Alam, M., Shiraz, M., Gani, A. 2014. Big data: survey, technologies, opportunities, and challenges. *The Scientific World Journal*, 2014: 712826.

Kondylatos, S., Prapas, I., Ronco, M., Papoutsis, I., Camps-Valls, G., Piles, M., Fernández-Torres, M. A., & Carvalhais, N. (2022). Wildfire Danger Prediction and Understanding With Deep Learning. *Geophysical Research Letters*, 49(17). <https://doi.org/10.1029/2022GL099368>

Küçük Ö., Bilgili, E., Durmaz, B.D., Sağlam, B., Baysal, İ. 2009. The Effect Factors on Transition from Surface Fire to Crown Fire. Kastamonu Univ., *Journal of Forestry Faculty*, 9(2):80-85.

Lin, J., Rinaldi, S. 2009. A derivation of the statistical characteristics of forest fires. *Ecological Modelling*, 220(7):898-903.

Madhav, A.D.V., Gargeya, A., Sravan, A., Suman, C. 2023. Prediction of Forest Fires Using Logistic Regression, EasyChair Preprint, No: 9901.

Milanović, S., Kaczmarowski, J., Ciesielski, M., Trailović, Z., Mielcarek, M., Szczygieł, R., Kwiatkowski, M., Bałazy, R., Zasada, M., & Milanović, S. D. (2023). Modeling and

Mapping of Forest Fire Occurrence in the Lower Silesian Voivodeship of Poland Based on Machine Learning Methods. *Forests*, 14(1). <https://doi.org/10.3390/f14010046>

Naderpour, M., Rizeei, H. M., & Ramezani, F. (2021). Forest fire risk prediction: A spatial deep neural network-based framework. *Remote Sensing*, 13(13). <https://doi.org/10.3390/rs13132513>

Oliveira, S., Oehler, F., San-Miguel-Ayanz, J., Camia, A., Pereira, J.M., Pereira, J.M.C. 2012. Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. *Ecol. Manag.* 275: 117–129.

Patle, A., Chouhan, D.S. 2013. SVM kernel functions for classification, 2013 International Conference on Advances in Technology and Engineering (ICATE), IEEE, pp. 1-9.

Pham, B. T., Jaafari, A., Avand, M., Al-Ansari, N., Du, T. D., Hai Yen, H. P., Phong, T. Van, Nguyen, D. H., Van Le, H., Mafi-Gholami, D., Prakash, I., Thuy, H. T., & Tuyen, T. T. (2020). Performance evaluation of machine learning methods for forest fire modeling and prediction. *Symmetry*, 12(6). <https://doi.org/10.3390/SYM12061022>

Phelps, N., & Woolford, D. G. (2021). Comparing calibrated statistical and machine learning methods for wildland fire occurrence prediction: a case study of human-caused fires in Lac La Biche, Alberta, Canada. *International Journal of Wildland Fire*, 30(11), 850–870. <https://doi.org/10.1071/WF20139>

Pourghasemi, H.R., Gayen, A., Lasaponara, R., Tiefenbacher, J.P. 2020. Application of learning vector quantization and different machine learning techniques to assessing forest fire influence factors and spatial modelling. *Environ. Res.* 184: 109321.

Pourtaghi, Z.S., Pourghasemi, H.R., Aretano, R., Semeraro, T. 2016. Investigation of general indicators influencing on forest fire and its susceptibility modeling using different data mining techniques. *Ecol. Indic.* 64:72–84.

Pyne, S.J. (2001). *Fire: A Brief History*. University of Washington Press.

Raschka S., Mirjalili, V. 2017. *Python Machine Learning—Second Edition*, London, U.K., Packt, ISBN: 978-1-78712-593-3. 622 p.

Rihan, M., Bindajam, A., Talukdar, S., Shahfahad, Naikoo, M., Mallick, J., Rahman, A. 2023. Forest Fire Susceptibility Mapping with Sensitivity and Uncertainty Analysis Using Machine Learning and Deep Learning Algorithms, *Advances in Space Research*, 72(2):426-443.

Rodrigues, M., Riva, J. 2014. An insight into machine-learning algorithms to model human-caused wildfire occurrence. *Environ. Model. Softw.* 57:192–201.

Sağlam, B., Bilgili, E., Durmaz, B. D., Kadioğulları, A. İ., Küçük, Ö. 2008. Spatio-temporal analysis of forest fire risk and danger using LANDSAT imagery. *Sensors*, 8(6):3970-3987.

San-Miguel-Ayanz, J., Durrant, T., Boca, R., Maianti, P., Liberta`, G., Artes Vivancos, T., Jacome Felix Oom, D., Branco, A., De Rigo, D., Ferrari, D., Pfeiffer, H., Grecchi, R. and Nuijten, D. 2022. Advance report on wildfires in Europe, Middle East and North Africa 2021, EUR 31028 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-49633-5.

Sayad, Y. O., Mousannif, H., & Al Moatassime, H. (2019). Predictive modeling of wildfires: A new dataset and machine learning approach. *Fire Safety Journal*, 104, 130–146. <https://doi.org/10.1016/j.firesaf.2019.01.006>

Sen, J. 2021. Machine Learning - Algorithms, Models and Applications. Artificial Intelligence. IntechOpen. London, UK. 154 p.

Shad, R., Mesgari, M. S., abkar, A., Shad, A. 2009. Predicting air pollution using fuzzy genetic linear membership kriging in GIS. *Computers, Environment and Urban Systems*, 33(6): 472-481.

Shafapourtehrany, M. 2022. Evaluation of machine learning performance in wildfire susceptibility mapping under limited training data condition. *Journal of Engineering Sciences and Researches*, 4:317-327.

Shao, Y., Feng, Z., Sun, L., Yang, X., Li, Y., Xu, B., & Chen, Y. (2022). Mapping China's Forest Fire Risks with Machine Learning. *Forests*, 13(6). <https://doi.org/10.3390/f13060856>

Sivrikaya, F., Akay, A.E., Oğuz, H., Yenilmez, N. 2011. Mapping Forest Fire Danger Zones Using GIS: A Case Study from Kahramanmaraş. 6th International Symposium on Ecology and Environmental Problems. 17-20 November, Antalya.

Sivrikaya, F., Küçük, Ö. 2022. Modeling forest fire risk based on GIS-based analytical hierarchy process and statistical analysis in Mediterranean region. *Ecological Informatics*, 68(2022):101537.

Sutanto, S.J., Vitolo, C., Di Napoli, C., D'Andrea, M. 2020. Heatwaves, droughts, and fires: Exploring compound and cascading dry hazards at the pan-European scale. *Environment International*, 134: 105276.

Syphard, A.D., Radeloff, V.C., Keeley, J.E., et al. (2007). Human influence on California Fire Regimes. *Ecological Applications*, 17(5), 1388–1402.

Tan, C., Feng, Z. 2023. Mapping Forest Fire Risk Zones Using Machine Learning Algorithms in Hunan Province, China. *Sustainability*, 15(7): 6292.

Tien Bui, D., Le, H.V., Hoang, N.D. 2018. GIS-based spatial prediction of tropical forest fire danger using a new hybrid machinelearning method. *Ecol. Inform.* 48: 104–116.

Türkan, S., Özel, G., Güney, C.O., Ünal, C., Şentürk, Ö., Özkan, K. 2023. A novel approach to determine the influence of weather conditions on forest fire risk in the Mediterranean Region of Türkiye, *Mugla Journal of Science and Technology*, 9(2): 1-10.

Vapnik, V.N., Vapnik, V. 1998. *Statistical Learning Theory*. Wiley, New York.
Zaidi, A. 2023. Predicting wildfires in Algerian forests using machine learning models, *Heliyon*, 9(7): e18064.

Wang, D., Guan, D., Zhu, S., Kinnon, M.M., Geng, G., Zhang, Q., Zheng, H., Lei, T., Shao, S., Gong, P., Davis, S.J. 2021. Economic footprint of California wildfires in 2018. *Nature Sustainability*. 4(3):252-260.

Xiangwei C., Long S., Qiangxue W., Shujing W., Dao G., Haiqing H. 2011. Response characteristics and prospect of forest fire disasters in the context of climate change in China, College of Forestry, Northeast Forestry University.





EK AÇIKLAMALAR A.

ARAŞTIRMA HAKKINDA

ARAŞTIRMA KONUSU

Bu araştırmanın temel konusu, orman yangını risk haritalarının oluşturulmasında makine öğrenimi algoritmalarının kullanımını incelemektir. Orman yangınları, çevresel ve ekonomik etkileri nedeniyle dünya genelinde önemli bir sorun teşkil etmektedir. Bu çalışmada, farklı makine öğrenimi algoritmalarının (lojistik regresyon, destek vektör makineleri, derin öğrenme, vb.) orman yangını riskinin tahmininde ne kadar etkili olduğu değerlendirilecektir. Bu değerlendirme, yangın riskini etkileyen çeşitli çevresel ve meteorolojik faktörlerin analizi üzerine odaklanacaktır.

ARAŞTIRMANIN AMACI VE ÖNEMİ

Araştırmanın temel amacı, orman yangını risk haritalarının oluşturulmasında kullanılan makine öğrenimi algoritmalarının performansını incelemek ve bu algoritmaların yangın tahminlerinde nasıl daha etkili hale getirilebileceğini belirlemektir. Bu bağlamda, Lojistik Regresyon ve Destek Vektör Makineleri algoritmalarının doğruluk oranları, genel performansları ve uygulanabilirlikleri değerlendirilecektir.

Araştırmanın önemi, orman yangınlarının etkili bir şekilde önlenmesi ve yönetilmesine katkı sağlayacak yenilikçi yöntemler geliştirilmesinde yatmaktadır. Makine öğrenimi algoritmalarının yangın risk haritalarının doğruluğunu artırma potansiyeli, yangın yönetim stratejilerinin iyileştirilmesi açısından büyük bir önem taşımaktadır. Bu çalışmanın sonuçları, yangın yönetim birimlerine ve karar vericilere, orman yangınlarını daha iyi tahmin etme ve önleme konusunda değerli bilgiler sunabilir.

ARAŞTIRMANIN YÖNTEMİ

Bu araştırmada kullanılan yöntem, makine öğrenimi algoritmalarının orman yangını risk haritalarının oluşturulmasındaki etkisini analiz etmeye yöneliktir. Araştırma sürecinde, aşağıdaki adımlar izlenecektir:

Veri Toplama: Orman yangını riski ile ilgili çevresel ve meteorolojik veriler çeşitli kaynaklardan toplanacaktır. Bu veriler, eğim, bakı, ağaç türü, yola olan uzaklık gibi faktörleri içerecektir.

Veri Ön İşleme: Toplanan veriler üzerinde temizleme, normalleştirme ve eksik verilerin doldurulması gibi veri ön işleme adımları uygulanacaktır. Bu adım, veri setinin makine öğrenimi algoritmalarına uygun hale getirilmesini sağlayacaktır.

Model Geliştirme: Lojistik regresyon ve destek vektör makineleri (SVM) yöntemleri gibi farklı makine öğrenimi algoritmaları kullanılarak modeller geliştirilecektir. Bu modeller, yangın riski tahmininde kullanılacak ve performansları karşılaştırılacaktır.

Model Değerlendirme: Geliştirilen modeller, doğruluk, hassasiyet, özgüllük gibi performans ölçütleri ile değerlendirilecektir. Ayrıca, modelin genelleme kapasitesi ve aşırı uyum riski de dikkate alınacaktır.

Sonuçların Analizi: Elde edilen sonuçlar analiz edilerek, hangi algoritmanın orman yangını riski tahmininde daha etkili olduğu belirlenecektir. Bu analizler, yangın yönetim stratejilerine yönelik öneriler sunacaktır.

ARAŞTIRMA HİPOTEZLERİ / PROBLEM

Araştırmanın temel problemi, orman yangını risk haritalarının doğruluğunu artırmak için en uygun makine öğrenimi algoritmalarının belirlenmesidir. Bu bağlamda, şu hipotezler test edilecektir:

H1: Lojistik regresyon algoritması, orman yangını riski tahmininde yüksek doğruluk oranı sunmaktadır.

H2: Destek vektör makineleri (SVM), orman yangını riski tahmininde lojistik regresyondan daha yüksek performans göstermektedir.

KAPSAM VE SINIRLILIKLAR

Araştırmanın kapsamı, orman yangını riski tahmininde makine öğrenimi algoritmalarının uygulanabilirliği ile sınırlıdır. Çalışma, belirli bir bölge veya ülke üzerine odaklanmamaktadır; genel olarak algoritmaların performansı üzerine odaklanmaktadır. Ancak, veri kaynaklarının çeşitliliği ve kalitesi, araştırmanın kapsamını ve elde edilen sonuçların genelleme kapasitesini sınırlayabilir. Ayrıca, kullanılan algoritmaların parametre ayarlamaları ve model optimizasyonu, sonuçların doğruluğunu etkileyebilecek önemli faktörlerdir.

KARŞILAŞILAN GÜÇLÜKLER

Araştırma sürecinde karşılaşılan güçlükler, veri toplama ve ön işleme aşamalarında ortaya çıkmıştır. Özellikle, çevresel ve meteorolojik verilerin doğruluğu ve tamlığı, modelin performansını doğrudan etkilemektedir. Ayrıca, makine öğrenimi algoritmalarının optimize edilmesi ve aşırı uyum (overfitting) riskinin azaltılması da zorluklar arasında yer almaktadır. Çeşitli algoritmaların karşılaştırılması ve en uygun yöntemin belirlenmesi, karmaşık ve zaman alıcı bir süreç olmuştur.

ÖZGEÇMİŞ

Aybike Göksu BEKTAŞ, Karabük Üniversitesi Bilgisayar Mühendisliği Bölümü'nden mezun olmuştur. Karabük Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı'nda Yüksek Lisans eğitimine devam etmektedir. Bursa'da GCS Test Denetim firmasında Siber Güvenlik ve AR-GE mühendisi olarak çalışmaktadır.

