

**ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

Metin ERSİN

**UYDU ve METEOROLOJİ VERİLERİNDEN MAKİNE
ÖĞRENMESİ YÖNTEMLERİ KULLANILARAK ADANA
BÖLGESİ İÇİN GÜNEŞ ENERJİSİ POTANSİYELİNİN
TAHMİNİ**

**UZAKTAN ALGILAMA VE COĞRAFİ BİLGİ SİSTEMLERİ
ANA BİLİM DALI**

ADANA-2021

ÖZ

YÜKSEK LİSANS TEZİ

UYDU ve METEOROLOJİ VERİLERİNDEN MAKİNE ÖĞRENMESİ YÖNTEMLERİ KULLANILARAK ADANA BÖLGESİ İÇİN GÜNEŞ ENERJİSİ POTANSİYELİNİN TAHMİNİ

Metin ERSİN

ÇUKUROVA ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ UZAKTAN ALGILAMA VE COĞRAFI BİLGİ SİSTEMLERİ ANABİLİM DALI

Danışman : Doç. Dr. Bekir Yiğit YILDIZ
Yıl: 2021, Sayfa: 87
Jüri : Doç. Dr. Nuri EMRAHOĞLU
: Doç. Dr. Bekir Yiğit YILDIZ
: Doç. Dr. Hüseyin TOPAKLI

Nüfus artışı, hızla gelişen sanayileşme, teknolojinin gelişmesi, insanların ihtiyaçlarının artması gibi birçok sebepten dolayı dünyada büyük miktarda enerji ihtiyacı ortaya çıkmıştır. Bu ihtiyacın ya geleneksel yöntemler olan fosil yakıtlara bağlı santrallerden ya da yeni nesil enerji kaynakları olan güneş enerjisi, rüzgâr enerjisi gibi temiz enerji olarak adlandırılan kaynaklardan sağlanması zaruri hale gelmiştir.

Bu çalışmada Adana meteoroloji istasyonunu temel alarak çevresini kapsayan bir alanda yer yüzeyine gelen güneş enerjisi miktarının makine öğrenmesi algoritmaları ile tahmin edilmesi amaçlanmıştır. Bu amaç doğrultusunda bu geniş bölgeyi en hızlı ve maliyetsiz olarak incelemek amacı ile uzaktan algılama verileri ve yöntemleri kullanılmıştır. NOAA-AVHRR uydu verileri NASA'nın resmi sitesi üzerinden temin edilip kalibrasyon ve coğrafik konumlandırma işlemleri yapıp yüzey sıcaklığı değerleri hesaplanmıştır. Makine öğrenmesi algoritmalarında çalıştırılmak üzere 8 farklı veri seti hazırlanmıştır. Bu veri setleri makine öğrenmesi algoritmaları ile çalıştırılmış ve WEKA yazılımı kullanılarak algoritmaların hangi veri setleri ile uygun sonuçlar verdiği araştırılmıştır.

Bu sonuçlar incelendiğinde lineer regresyon algoritmasının %79,9'luk, çok katmanlı algılayıcı algoritmasının %88,4'lük, destek vektör regresyon algoritmasının %89,2'lik başarımlarına sahip olduğu belirlenmiştir.

Anahtar Kelimeler: Uzaktan Algılama, Makine Öğrenmesi, NOAA, Güneş enerjisi

ABSTRACT

MSc THESIS

FORECASTING THE SOLAR ENERGY POTENTIAL FOR THE ADANA REGION USING MACHINE LEARNING METHODS FROM SATELLITE AND METEOROLOGY DATA

Metin ERSİN

**ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCE
DEPARTMENT OF REMOTE SENSING AND
GEOGRAPHICAL INFORMATION SYSTEM**

Supervisor : Assoc.Prof. Dr. Bekir Yiğit YILDIZ
Year: 2021, Pages: 87
Jury : Assoc.Prof. Dr. Nuri EMRAHOĞLU
: Assoc.Prof. Dr. Bekir Yiğit YILDIZ
: Assoc.Prof. Dr. Hüseyin TOPAKLI

Due to many reasons such as population growth, rapidly developing industrialization, development of technology, and increase in people's needs, the world's need for large amounts of energy has emerged. It has become necessary to meet this need either from power plants based on fossil fuels, which are traditional methods, or from sources called clean energy, such as solar energy, wind energy, which is a new generation energy source.

In this study, it is aimed to calculate the amount of solar energy, based on Adana meteorology station, in the area covering the self region in order to overcome this energy deficit. For this purpose, remote sensing data and machine learning algorithms have been used in order to examine the solar radiation in this wide region in the fastest and most cost-effective way. NOAA-AVHRR satellite data was obtained from NASA' web site, calibration and geographic processes were performed and land surface temperature values were calculated. 8 different data sets were prepared to be run in machine learning algorithms. These data sets were run with machine learning algorithms by using WEKA software, it was investigated with which data sets the algorithms gave appropriate results.

When these results are examined, it has been determined that the linear regression algorithm has 79.9%, the multilayer perceptron algorithm has 88.4% and the support vector regression algorithm has 89.2% success rates.

Key Words: Remote Sensing, Machine Learning, NOAA, Solar Radiation

GENİŞLETİLMİŞ ÖZET

Enerji her dönemde insanlık için vazgeçilmez olmuştur. Günümüzde fosil yakıtların tükenmekte olması ve ciddi çevre sorunları yaratması insanlığı yenilenebilir, temiz enerji kaynakları araştırmaya sevketmiştir. Jeotermal, dalga, rüzgar ve güneş enerjisi en çok araştırılan kaynaklardır. Bunlardan güneş enerjisi Türkiye açısından da stratejik bir enerji kaynağıdır. Fotovoltaik teknolojinin gelişimiyle güneş panellerinden oluşan santrallerde elektrik üretimi mümkün olmaktadır. Bu santrallerin kurulumu ciddi maliyetler içerdiğinden iyi planlama yapılmalıdır. Güneşten en çok faydalanan, enerji iletimi verimli yapılabilen yerler seçmek önemlidir. Yeryüzünde her konumun güneş enerjisi potansiyeli farklıdır. Coğrafi ve meteorolojik etkenler bunu etkilerken, güneş enerjisi potansiyelinin her noktada ölçümü son derece güçtür. Her yerde meteoroloji istasyonları kurmak son derece maliyetli olup güneş enerjisi potansiyelini ölçümler yerine değişik meteorolojik parametrelerden tahmin etme eğilimi son yıllarda literatürde giderek artmıştır (Zhou ve ark., 2021).

Günümüzde yapay uydular, başta haberleşme olmak üzere değişik amaçlar için kullanılır olmuştur. Dünyayı sürekli gözlemleyebilen bu uydulardan elde edilen değişik dalga boylarındaki görüntüler ile çeşitli meteorolojik veriler hesaplanabilmektedir. Uzaktan algılama disiplini başta uydular olmak üzere değişik hava cihazları ile yeryüzü ve atmosfer hakkında sürekli veriler toplayabilmektedir.

Bu tez çalışmasında Adana Meteoroloji İstasyonu baz alınarak, Adana bölgesi için güneş enerjisi potansiyeli tahmin edilmeye çalışılmıştır. Uydu verileri ve meteorolojik verilerden oluşan girdi setlerinin, güneş enerjisi miktarını belirlemede ağırlıkları araştırılmıştır. Verilerden tahminler yapmak için bir yapay zeka alanı olan denetimli makine öğrenmesi algoritmalarından faydalanılmıştır. Algoritmaların çalıştırılması için ücretsiz bir veri madenciliği - makine öğrenmesi yazılımı, WEKA kullanılmıştır. NOAA-AVHRR uydu görüntüleri NASA'nın

resmi sitesi üzerinden ikisi görünür bölge, üçü kızılötesi bölgeyi içeren beş bant verisi temin edilip kalibrasyon ve coğrafik konumlandırma işlemleri yapıp yüzey sıcaklığı değerleri hesaplanmıştır. Adana Meteoroloji Müdürlüğünden yoğunlaşabilir su buharı, günlük hava sıcaklığı ve radyometrik güneş enerjisi verileri alınmıştır. Makine öğrenmesi algoritmalarında çalıştırılmak üzere 8 farklı veri seti kaynağına bağlı olarak çeşitli varyasyonlarda hazırlanmıştır. Bu veri setleri, WEKA yazılımı üzerinde makine öğrenmesi algoritmalarında çalıştırılmış algoritmalara göre hangi veri setleri ile uygun sonuçlar verdiği araştırılmıştır.

Çalışmada en iyi değerler, uydu ham bant verilerinden hesaplanmış meteorolojik ürünler olan kızılötesi bant fark değeri (T_4-T_5) ve yer yüzey sıcaklığı (LST) ile atmosferdeki yoğunlaşabilir su buharı (PWAT) ve günlük hava sıcaklığı (TEMP) meteoroloji verilerinin kullanıldığı veri seti grubunda destek vektör regresyonu (SVM) algoritmasının $R=0,892$ ve $RMSE=111,97$ değerleri elde edilmiştir. Çok katmanlı algılayıcı (MLP) algoritmasının sonuçları da $R=0,817$ ve $RMSE=187,54$ değerleri ile nispeten SVM'ye yakın ve başarılı olarak belirlenmiştir. Bu veri grubu hiçbir ham uydu bant verisi barındırmamaktadır. Ham uydu verisi yerine ham verilerden hesaplanmış ürünlerin kullanılmasının daha kullanışlı olduğunu ortaya çıkarmıştır. Diğer veri setlerinin sonuçları ile karşılaştırıldığında tek başına ne meteoroloji verileri ne de uydu verileri daha iyi sonuçlar verebilmiştir. En düşük başarımlık değeri olan $R=0,753$ SVM algoritmasının sadece uydu verileri ve uydudan elde edilen ürünler ile çalıştırılmasından elde edilmiştir.

Bu çalışma, güneş enerjisi potansiyelini tahmin etme amacıyla yapılacak yeni araştırmalarda farklı veri setlerinin ve makine öğrenmesi yöntemlerinin kullanılması gerektiğini öneren bir altlık görevi üstlenmiştir.

TEŞEKKÜR

Başta danışmanım Doç. Dr. Bekir Yiğit YILDIZ olmak üzere, Uzaktan Algılama ve Coğrafi Bilgi Sistemleri anabilim dalının tüm hocalarına teşekkürü borç bilirim. Bu süreçte beni hep destekleyen eşim Sibel ve kızım Bilgesu'ya minnettarım.



İÇİNDEKİLER	SAYFA
ÖZ.....	I
ABSTRACT.....	II
GENİŞLETİLMİŞ ÖZET	III
TEŞEKKÜR.....	V
İÇİNDEKİLER	VI
ÇİZELGELER DİZİNİ	VIII
SİMGELER VE KISALTMALAR.....	XII
1. GİRİŞ	1
1.1. Uzaktan Algılama Uyduları ve Meteoroloji.....	2
1.2. Makine Öğrenmesi ve Güneş Enerjisi Tahmini	4
2. ÖNCEKİ ÇALIŞMALAR.....	7
3. MATERYAL VE METOT.....	15
3.1. NOAA Uyduları.....	15
3.1.1. Çok Yüksek Çözünürlüklü Gelişmiş Radyometre (AVHRR).....	16
3.1.2. NOAA-AVHRR Termal Kanal Verilerinin Kalibrasyonu	20
3.1.3. Bitki Örtüsü İndeksinin Hesaplanması	23
3.1.4. Temel Yer Yüzeyi Objelerinin Yayınırılığı.....	24
3.1.5. Split-Window Metodu ile Yer Yüzey Sıcaklığının Hesaplanması	26
3.2. Yapay Zeka / Makine Öğrenmesi	32
3.2.1. Makine Öğrenmesi Türleri	33
3.2.2. Yapay Sinir Ağları / Çok Katmanlı Algılayıcı.....	34
3.2.3. Destek Vektör Makinesi / Regresyonu	37
3.3. Klasik Regresyon Analizi	39
3.4. Değerlendirme Kriterleri.....	40
3.5. Uygulama Yazılımı (WEKA)	42
3.5.1. WEKA Arayüzleri ve İşlevleri.....	43
3.5.2. WEKA Algoritmaları.....	51
4. BULGULAR VE TARTIŞMA	53

4.1. Uydu Görüntü İşlemleri.....	53
4.2. Veri Setinin Düzenlenmesi.....	61
4.3. Çok Katmanlı Algılayıcı Uygulaması	66
4.4. Destek Vektör Regresyonu Uygulaması	69
4.5. Regresyon Analizi Uygulaması.....	71
4.6. Uygulama Sonuçları.....	73
5. SONUÇ VE ÖNERİLER.....	79
KAYNAKLAR	81
ÖZGEÇMİŞ	87

ÇİZELGELER DİZİNİ

SAYFA

Çizelge 3. 1. AVHRR/2-3 Algılayıcısının Bantları Ve Bant Özellikleri.....	17
Çizelge 3.2. NOAA-15 ve Sonrası AVHRR Uydu İçin Sıcaklık Düzeltme Ve Merkezi Dalgaboyu Katsayıları	63
Çizelge 3.3. Bazı Maddelerin Yayınrlık Değerleri	26
Çizelge 3.4. Weka Sonuçlarında Verilen Değerlendirme Kriterleri	41
Çizelge 4.1. Verilerin Betimsel İstatistikleri	61
Çizelge 4.2. Veri Setinin İçeriği	63
Çizelge 4.3. Veri Seti Kombinasyonları	64
Çizelge 4.4. Çok Katmanlı Algılayıcı Sonuçları.....	68
Çizelge 4.5. Destek Vektör Regresyonu Sonuçları	70
Çizelge 4.6. Lineer Regresyon Sonuçları	72
Çizelge 4.7. Veri Setlerine göre Algoritmaların Başarı ve Hata değerleri	73
Çizelge 4.8. Algoritmalara Göre En Başarılı Güneş Enerjisi Tahmin Sonuçları	74



ŞEKİLLER DİZİNİ SAYFA

Şekil 1.1. Aktif ve Pasif Uzaktan Algılama ile Gözlem Uydusunun Veri Toplaması ... 3

Şekil 1.2. Güneş Enerjisi Tahmininde Kullanılan Makine Öğrenmesi

Yöntemlerinin Sınıflandırması..... 5

Şekil 3.1. NOAA uydusu ve üzerinde yer alan algılayıcı sistemler..... 16

Şekil 3.2. Yayınlılık ile NDVI Değerleri Arasındaki Bağntı..... 25

Şekil 3.3. Feedforward Neural Network yapısı..... 35

Şekil 3.4. Çok Katmanlı Algılayıcı Nöron Yapı ve Bağlantıları. 36

Şekil 3.5. Destek Vektör Makinesi Vektör Sınıflaması 38

Şekil 3.6. WEKA Yazılım Ara Yüzü..... 44

Şekil 3.7. WEKA İçeriğine Erişim Ara Yüzü, Paket yöneticisi..... 45

Şekil 3.8. Explorer Ara Yüzü..... 46

Şekil 3.9. Experimenter Ara Yüzü. 47

Şekil 3.10. Knowledge Flow Ara Yüzü. 48

Şekil 3.11. Workbench Ara Yüzü. 49

Şekil 3.12. Simple CLI Ara Yüzü 50

Şekil 4.1. 4 Eylül 2011 Tarihli Ham NOAA–AVHRR Görüntüsü..... 54

Şekil 4.2. 4 Eylül 2011 Tarihli Ham NOAA–AVHRR Görüntüsüne Ait 5 Bant

Histogram Değerleri..... 54

Şekil 4.3. 4 Eylül 2011 Tarihli Ön-İşlemleri Yapılmış AVHRR 1.Bant Görüntüsü... 55

Şekil 4.4. 4 Eylül 2011 Tarihli AVHRR 1.Banta Ait Histogram..... 55

Şekil 4.5. 4 Eylül 2011 Tarihli Ön-İşlemleri Yapılmış AVHRR 2.Bant Görüntüsü... 56

Şekil 4.6. 4 Eylül 2011 Tarihli AVHRR 2.Banta Ait Histogram..... 56

Şekil 4.7. 4 Eylül 2011 Tarihli Ön-İşlemleri Yapılmış AVHRR 3.Bant Görüntüsü .. 57

Şekil 4.8. 4 Eylül 2011 Tarihli AVHRR 3.Banta Ait Histogram..... 57

Şekil 4.9. 4 Eylül 2011 Tarihli Ön-İşlemleri Yapılmış AVHRR 4.Bant Görüntüsü.... 58

Şekil 4.10. 4 Eylül 2011 Tarihli AVHRR 4.Banta Ait Histogram..... 58

Şekil 4.11. 4 Eylül 2011 Tarihli Ön-İşlemleri Yapılmış AVHRR 5.Bant Görüntüsü..	59
Şekil 4.12. 4 Eylül 2011 Tarihli AVHRR 5.Banta Ait Histogram.....	59
Şekil 4.13. Sobrino 1994 ile Hesaplanan LST (Yer Yüzey Sıcaklığı) Haritası	60
Şekil 4.14. ARFF Dosya Yapısı.....	62
Şekil 4.15. Eğitim Veri Setinin Yükleme Ekranı.....	65
Şekil 4.16. Test Veri Setinin Yükleme Ekranı.....	66
Şekil 4.17. Çok Katmanlı Algılayıcı Seçim Ekranı	67
Şekil 4.18. Çok Katmanlı Algılayıcı Ayar Ekranı	68
Şekil 4.19. Destek Vektör Regresyonu Seçim Ekranı.....	69
Şekil 4.20. Destek Vektör Regresyonu Ayar Ekranı.....	70
Şekil 4.21. Regresyon Analizi Seçim Ekranı	71
Şekil 4.22. Regresyon Analizi Ayar Ekranı	72

SİMGELER VE KISALTMALAR

ANN	: Artificial Neural Networks / Yapay Sinir Ağları
ARMA	: Autoregressive Moving Average
ARIMA	: Autoregressive Integrated Moving Average
AVHRR	: Advanced Very High Resolution Radiometer
CART	: Classification and Regression Tree
CNN	: Convolutional Neural Network / Evrişimli Sinir Ağı
CONUS	: Amerika Birleşik Devletleri Kıtası
ConvLSTM	: Convolutional Long Short-Term Memory / Evrişimsel Vadeli Bellek
CPU	: Central Processing Unit / Merkezi İşlem Birimi
ÇKA	: Çok Katmanlı Algılayıcı
DSR	: Downward Solar Radiation / Küresel Aşağı Güneş Radyasyonunu
DVM	: Destek Vektör Makinesi
DVR	: Destek Vektör Regresyonu
ELM	: Extreme Learning Machine
GPU	: Graphics Processing Unit / Grafik İşlemci Birimi
LST	: Land Surface Temperature /Yer Yüzey Sıcaklığı
MAX	: Maximum Display
MGM	: Meteoroloji Genel Müdürlüğü
MLP	: Multi Layer Perceptron / Çok Katmanlı Algılayıcı
MSE	: Mean Squared Error / Ortalama Karesel Hata
NASA	: National Aeronautics and Space Agency
NDVI	: Normalized Difference Vegetation Index
NOAA	: National Oceanic and Atmospheric Administration
NWP	: Numeric Weather Prediction / Sayısal Hava Tahmin
RF	: Random Forest
RMSE	: Root Mean Squared Error / Kök Ortalama Kare Hatası

RNN	: Yenilenen-Tekrarlanan Sinir Ağları / Recurrent Neural Network
RS	: Remote Sensing / Uzaktan algılama
SHT	: Sayısal Hava Tahmin
SSIM	: Yapısal Benzerlik Endeksi / Structural Similarity
SVM	: Support Vector Machine / Destek Vektör Makinesi
SVR	: Support Vector Regression / Destek Vektör Regresyonu
UA	: Uzaktan Algılama
WMO	: World Meteorology Organization / Dünya Meteoroloji Teşkilatı
YSA	: Yapay Sinir Ağları

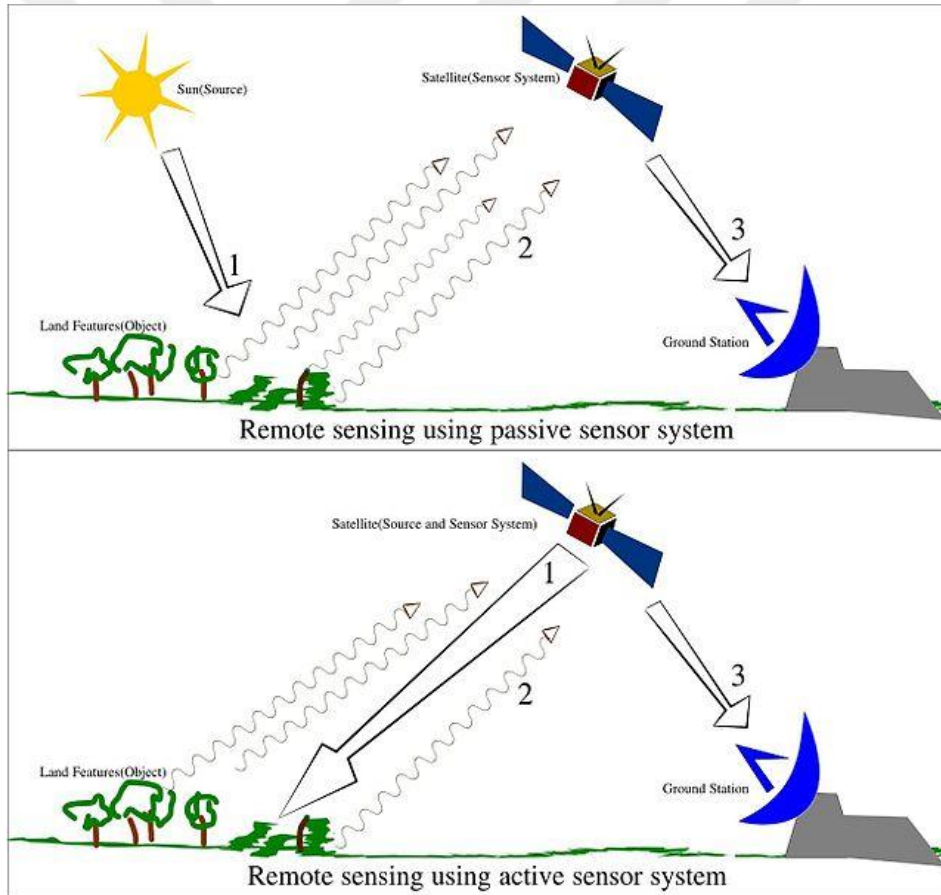
1. GİRİŞ

İnsanlık, tarih boyunca enerjiye ihtiyaç duymuştur. Önceleri ısınma, pişirme gibi ihtiyaçlar, sanayi devrimiyle birlikte yaşamın tüm süreçlerinde enerjiden bağımsız kalınmayacağını ortaya koymuştur. Sanayinin ilk dönemlerinde kömürle başlayan organik-fosil enerji kaynağı bağımlılığı sonraları petrol ürünleri ile katlanarak devam etmiştir. Bu artışta bu yakıtların kullanım kolaylığı önemli rol oynamıştır. Benzinli ve dizel motorların icadı, elektriğin depolanma sorunu yüzünden elektrik motorunun yaygınlaşmasını özellikle ulaşım araçlarında engellemiş hâkimiyeti petrol tabanlı enerji çevrimlerine teslim etmiştir. Yıllar içinde, fosil yakıt kullanımı artmış çok kısa sürede ciddi bir çevre kirlenmesi yaratmaya başlamıştır. Bu süreçte karbon emisyonu küresel ısınma oluşturacak boyutlara ulaşmış, karbon parmakızı kavramını hayatımıza sokmuştur. Zorunluluk insanlığı temiz enerji kaynakları aramaya yöneltmiştir (Raha ve Pal, 2010). Bu arayış özellikle de yenilenebilir olması sebebi ile başta güneş enerjisi (solar radiation) olmak üzere rüzgâr, dalga ve hidroelektrik gibi kaynaklara yönelmeye sebep olmuştur. Bu kaynakların elektrik enerjisine rahat dönüştürülmesi de ayrıca büyük avantajdı. Bu kez de insanlığın karşısına bu enerjilerin istendiğinde, istenildiği kadar üretilmesi sorunu çıkmıştır. Rüzgârın değişken olması eski çağlarda yelkenli denizcilerin de ciddi problemi olmuş ve buhar gücüne geçişte bu enerjinin yönetilebilir olmaması önemli etken olmuştur. Rüzgar türbinlerinden sürekli, aynı miktarda enerji üretimi mümkün değildir. (Dincer ve Acar, 2015) Nitekim güneş enerjisinde de benzer şekilde coğrafi konuma, mevsime özellikle de gece gündüz döngüsüne bağlı olarak ciddi üretim planlama sorunları söz konusudur. Bir bölgede ne miktarda ve zamana bağlı üretim değişimini tespit etmek ve planlamak yatırım kararları açısından son derece hayatidir. Ekonomik, sosyal, coğrafi ve mühendislik temelli birçok faktör bu kararları etkiler. Çünkü enerji santrali kurulacak yerin konumu, iletim hatları hatta enerji kullanıcısının profili, tüketim miktarını dolayısı ile üretim kapasitesini belirleyecektir.

1.1. Uzaktan Algılama Uyduları ve Meteoroloji

Günümüzde güneş enerjisinden genellikle fotovoltaik sistemler yardımıyla elektrik üretilmektedir. Güneş enerjisi dünyaya elektromanyetik dalgalar şeklinde ulaşır ve ulaşan enerji belli bir alan için spektrometre denilen cihazlar ile ölçülür. Bunlar yer meteoroloji istasyonlarında bulunur. Sürekli olarak, her konumda zaman ve mekândan bağımsız ölçümler yapmak mümkün olamamaktadır. Hava ve uzay gözlem platformlarında benzer görevleri gören cihazlar kullanılır. Atmosfer koşulları ve değişimi yapılan ölçümleri etkiler sıcaklık, nem, basınç, rüzgâr ve benzeri parametreler yeryüzüne düşen enerji miktarını etkiler. Meteoroloji istasyonlarını her yere kurmak maliyetlidir ve yapılan ölçümler coğrafi konuma bağlı olarak değişkenlik gösterir. Adana merkezde yapılan ölçümler eş zamanlı olarak yakındaki Toros Dağlarında yapıldığında aynı sonuçları vermeyecektir. Bu yüzden uzaktan algılama denilen yöntemle yeryüzünü, uydularda bulunan algılayıcılar yardımı ile gözlemlemek ve ölçümler yapmak hem daha pratik hem daha az maliyetlidir. Uzaktan algılama (remote sensing), nesneler ile fiziki temas olmaksızın herhangi bir uzaklıktan yapılan ölçümler ile nesneler hakkında bilgi edinme yöntemi şeklinde tanımlanmaktadır. Ölçüm değerleri nesnelerin elektromanyetik spektrum içindeki davranışlarına, lokasyon özelliklerine ve zaman içinde güneşin erişimlerine bağlıdır (Curran,1985). Meteoroloji uyduları ile bulutluluk, yüzey sıcaklığı vb. birçok meteorolojik gözlem yapılmaktadır. Bu sayede ulaşılabilen noktalarda, denizlerde, çöllerde, dağlarda meteorolojik ölçümler yapılabilmektedir. Dünyaya gelen solar radyasyonun % 50'si yeryüzüne ulaşır, bunun da %30'u yansıtılarak gider. %70 70'lik kısım yer koşullarına bağlı olarak soğurulur. Yeryüzüne dik olarak ulaşan ortalama güneş enerjisine, güneş sabiti denir, yaklaşık değeri 1368 W/m^2 dir. Bu sabit yıllık süreçte en çok %0,1 değişim gösterir. Suyu, kayaya gelen radyasyonun emilimi ve yansıtması farklı olacağından meteorologlar yüzey sıcaklığı denilen gölgede çim ekili yüzeyden 1,5 m yükseklikte ölçülen sıcaklığı temel alırlar (Ackerman ve Knox, 2015).

1972 yılında ilk kez LANDSAT-1'in gönderilmesi, ardından LANDSAT-TM serisi ve NOAA, SPOT vb uyduların geliştirilmesi ile uzaktan algılama yöntemleri yeryüzünün incelenmesinde yeni olanaklar sağlamıştır. 2022 yılında NOAA (Ulusal Okyanus Atmosferik Teşkilatı) GOES-T uydusunu, üzerinde gelişmiş birçok cihazıyla göndermeye hazırlanmaktadır. Uydu verilerinin artık meteoroloji, jeoloji, ziraat, atmosfer ve iklim gibi birçok farklı alanda kullanımı kaçınılmaz olmuştur. Uydular gövdelerinde bulunan çeşitli algılayıcılar yardımıyla yeryüzünü algılar ve çeşitli formatlarda bu bilgileri yer istasyonlarına iletirler. Bu çalışmada uydu ve meteoroloji verileri birlikte kullanılmıştır.



Şekil 1.1. Aktif ve Pasif Uzaktan Algılama ile Gözlem Uydusunun Veri Toplaması (Wikipaedia, 2021)

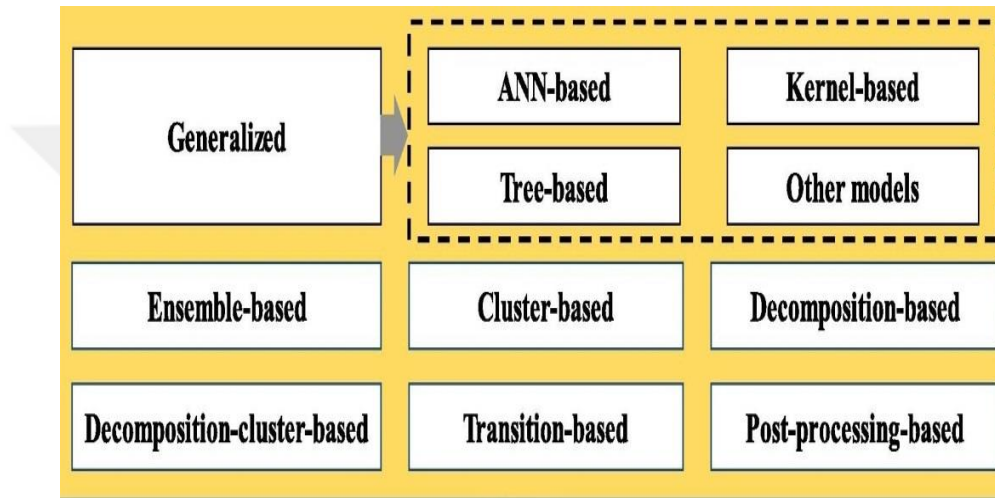
1.2. Makine Öğrenmesi ve Güneş Enerjisi Tahmini

Meteorolojik tahminleri yapmak için sürekli atmosferik parametrelerin ölçümü gereklidir. Günümüzde Sayısal Hava Tahmin (Numeric Weather Prediction) NWP modelleri de kullanılmaktadır. Atmosferin kaotik doğası ölçülen veri ile tahmini zorlaştırmakta doğrusal çıkarımlar yapmak çoğu zaman mümkün olmamaktadır. Doğrusal olmayan çözümlere ihtiyaç duyulduğunda ise hesaplama yeteneği ve zaman kısıtları ile karşı karşıya gelinmektedir. Bu sorunların üstesinden gelecek tahminlerin yapılabilmesi için farklı çözüm algoritmalarına ihtiyaç duyulmaktadır. Ampirik, İstatistiksel modeller yanında makine öğrenmesi denen, problemlerin çözümünü eldeki verilerden hareketle, eğitilerek tahmin süresinin kısalığını ve doğruluğunu sağlayacak algoritmalar geliştirilmiş bilim ve teknolojide birçok problemin çözümünde kullanılmaktadır.

Çoğu zaman meteorolojik veriler her nokta için olmadığından, uydu görüntülerinden elde edilen verilerden sonuçlar çıkaracak yöntemlere ihtiyaç duyulur. Günümüzde güneş enerjisi tahmini için kullanılan bazı yaklaşımlar vardır. Ampirik modellemede genelde gün ışığına, bulutluluk, sıcaklık vb meteorolojik değişkene dayalı ölçümlerden yapılan hesaplamalar kullanılır. (Besarat ve ark., 2013) istatistiksel modellemede, temelde regresyon ve zaman serilerine dayalı veri analizi söz konusudur. Yapay zekânın (Artificial Intelligence) AI alt alanı olan makine öğrenmesi (Machine Learning) ML teknikleri güneş enerjisi tahmini için son dönemlerde çok kullanılmaktadır. ML yöntemlerinden biri olan yapay sinir ağları (Artificial Neural Network) ANN doğrusal olmayan bir yaklaşımla, belirli bir bölge için tahminlerde bulunabilen bir yöntemdir. ANN kaotik bir süreç olan güneş enerjisi tahmini, doğru tahmin etme yeteneğine sahiptir (Fente, 2018). Bu yöntemleri bir arada kullanan hibrit yöntemlere de rastlanmaktadır.

Güneş enerjisi tahmini için ML yöntemleri literatürde yapay sinir ağları tabanlı, çekirdek tabanlı, ağaç tabanlı ve diğer modeller diye sınıflandırılmaktadır. ANN tabanlı yöntemlere örnek olarak; Multi Layer Perceptron (MLP), Convolutional Neural Network (CNN), vb verilebilir. Çekirdek tabanlı yöntemlere

örnek olarak, Support-Vector Machine (SVM), Support-Vector Regression (SVR) vb verilebilir. Ağaç tabanlı yöntemlere örnek olarak; Classification and Regression Tree (CART), Random Forest (RF) vb verilebilir. Diğer modellere örnek olarak; Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA) vb verilebilir (Zhou ve ark., 2021).



Şekil 1.2. Güneş Enerjisi Tahmininde Kullanılan Makine Öğrenmesi Yöntemlerinin Sınıflandırması (Zhou ve ark., 2021).

Bu çalışmada; ML yöntemlerinden literatürde en sık kullanılan ve başarılı sonuçlar veren ANN tabanlı gruptan, MLP (Çok Katmanlı Algılayıcı) ile çekirdek tabanlı gruptan SVR (Destek Vektör Regresyonu) ve istatistiksel yöntemlerden ise LR (Lineer Regresyon) ile yer yüzeyine gelen güneş enerjisinin tahmini yapılmıştır. Algoritmaların girdi verisi olarak NOA-AVHRR uydusu ve meteorolojik veriler kullanılmıştır. Çalışmanın amacı uydu verileri ve meteorolojik verilerden oluşan girdi setlerinin güneş enerjisi miktarını belirlemede ağırlıklarını araştırmaktır. Algoritmaların çalıştırılması için ücretsiz bir ML programı olan WEKA kullanılmıştır.



2. ÖNCEKİ ÇALIŞMALAR

Kaba ve ark. 2018 yılında yaptıkları çalışmada Türkiye'de bulunan 30 istasyon üzerinden güneş radyasyonunu tahmin etmek için derin öğrenme yöntemi kullanılmıştır. Girdi öznetelikleri olarak astronomik faktör, dünya dışı radyasyon ve iklim değişkenleri, güneşlenme süresi, bulut örtüsü, minimum sıcaklık ve maksimum sıcaklık kullanılmış ve çıktı SR olarak elde edilmiştir. Modelin eğitimi ve test edilmesi için 2001-2007 yıllarını kapsayan 34 istasyonun veri seti sırasıyla kullanılmış ve simüle edilen değerler yer-gerçeklik değerleri ile karşılaştırılmıştır. Genel belirleme katsayısı, ortalama karekök hata ve ortalama mutlak hata sırasıyla 0.980, 0.78MJm⁻²gün⁻¹ ve 0.61MJm⁻²gün⁻¹ olarak hesaplanmıştır. Sonuç olarak, derin öğrenme modeli, günlük küresel güneş radyasyonunu tahmin etmek için çok kesin ve karşılaştırılabilir sonuçlar vermiştir.

Karaman ve ark. 2021 yılında yayınladıkları çalışmada, ekstrem öğrenme makinelerinin (ELM) ile ANN güneş radyasyonu tahmininde performanslarını araştırmışlardır. 2010-2018 yıllarında Karaman ilinden elde edilen veriler kullanılarak yapılan bu çalışmada rüzgâr hızı ortalama sıcaklık ve güneşlenme süresi girdi parametresi olarak kullanılmıştır. Sonuçlar karşılaştırıldığında ELM'nin daha yüksek bir tahmin performansı gösterdiği ortaya konmuştur. Ayrıca, en iyi tahmin yanıtını elde etmek için ANN ve ELM farklı aktivasyon fonksiyonları ile test edilmiştir. ANN için en iyi tahmin sonucu 0.9828 olarak tansig fonksiyonu ile elde edilirken, ortalama kare hata (MSE) 0.000129 olarak elde edilmiştir. ELM için en iyi tahmin sonucu sin fonksiyonu 0.991 olarak elde edilmiş ve MSE 0.000881 olarak hesaplanmıştır. Elde edilen korelasyon katsayıları için güven aralığı tahminleri yapılmış ve sonuçlar karşılaştırılmıştır. ANN tahmin performansı, %95 güven aralığı ile ELM'den daha iyi olduğu belirlenmiştir.

Fallahi ve ark. 2018 yılında yaptıkları çalışmada, İrandaki çalışma bölgelerinde aylık ortalama güneş radyasyonunu tahmin etmek için ANN ve çoklu doğrusal regresyon (MLR) modellerinin doğruluğunu test etmişlerdir. Modellerin

girdi verileri boylam, enlem, yükseklik ve ay ve NOAA uydusunun AVHRR sensöründen elde edilen veriler ilk veri serisini oluşturmaktadır. Yer istasyonlarında ölçülen veriler ikinci veri serisi (DS2) diye adlandırılmıştır. AVHRR sensöründen arazi yüzey sıcaklığını (LST) hesaplamak için alanın yayınlılık değerleri dikkate alınmış ve bu amaçla AVHRR sensörünün 1 ve 2 kanallarından hesaplanan normalize edilmiş bitki örtüsü farkı indeksi (NDVI) kullanılmıştır. Elde edilen sonuçlar, eğitim aşamasında $R^2 = 0.96$, RMSE = 1.04, MAE = 1.1 ve test aşamasında $R^2 = 0.96$, RMSE = 1.06, MAE = 1.15 ile DS1 veri setinin kullanıldığı ANN modelinin, ANN modeline kıyasla daha tatmin edici bir performans elde ettiğini göstermiştir. Uzaktan algılama verilerine sahip ANN modelinin, yer ölçüm istasyonu olmayan yerlerde güneş radyasyonunu tahmin etme potansiyeline sahip olduğu sonucuna varılmıştır.

Şahin, M. (2012), ANN ve ELM (Extreme Learning Machine) kullanılarak değişen iklim koşullarına sahip 61 lokasyonda güneş radyasyonu tahmin etmiştir. ANN ve ELM yöntemleri 2002 ve 2003 yıllarına ait verilerle eğitilirken, bu yöntemlerin doğruluğu 2004 yılı verileriyle test edilmiştir. NOAA-AVHRR (National Oceanic and Atmospheric Administration Advanced Very High Resolution Radiometer) uydusunun verileri, ANN ve ELM modellerinin geliştirilmesinde girdi olarak güneş radyasyonu çıktı olarak hesaplanmıştır. Sonuçlar daha sonra istatistiksel yöntemlerle meteorolojik değerlerle karşılaştırılmıştır. Güneş radyasyonu hesaplamasında istatistiksel olarak anlamlı sonuç elde etmek için EML yönteminin ANN yöntemine göre daha başarılı olduğu ortaya konulmuştur.

Şenkal ve ark. (2019), meteorolojik bir uydu olan NOAA-AVHRR verilerinden elde edilen aylık küresel kara yüzey sıcaklık değerlerini kullanarak Becker ve Li yöntemleri ile güneş radyasyonunun tahmininde bulunmuştur. Çalışmada genelleştirilmiş regresyon sinir ağını kullanmışlardır. Sinir ağını eğitmek için Türkiye'de 2002 yılından kısa dönem (Adana) ve 1998-2002 yıllarından uzun dönemli (İzmir) meteorolojik uydu ve coğrafi veriler

kullanılmıştır. Sistemin giriş katmanında meteorolojik uydu ve coğrafi veriler (enlem, boylam, yükseklik, ay ve ortalama yer yüzey sıcaklığı) kullanılmıştır. Güneş radyasyonu ise çıktı parametresidir. ANN ile tahmin edilen ve temel değerler arasındaki ortalama kareler ve korelasyon katsayısı verileri elde edilmiştir. Meteorolojik uydu verilerinden sağlanan kara yüzey sıcaklık verileri ile yapay sinir ağları yöntemini kullanan güneş radyasyonunun tahmini ile ilgili öncü bir çalışmadır.

Ravinesh ve ark. (2019), Avustralya üzerinden uzun vadeli güneş radyasyonunu tahmin etmek için bölgesel olarak uyarlanabilir ve tahmine dayalı olarak verimli bir ELM modeli tasarlayan bir çalışma yapmıştır. MODIS (Moderate Resolution Imaging Spectroradiometer) algılayıcısından alınan ilgili uydu tabanlı giriş verileri (normalize edilmiş fark bitki indeksi, yer yüzey sıcaklığı, bulut tepe basıncı, bulut tepe sıcaklığı, bulut etki emisyonu, bulut yüksekliği, ozon ve yakın kızılötesi netliğinde su buharı), zamansal girdi değişkenleri (periyodiklik, enlem, boylam ve yükseklik) ile yaklaşık olarak eşit dağıtılmış ve yer tabanlı uzun vadeli güneş radyasyonunu (hedef çıktı) ile eşleştirilmiş toplam 41 çalışma alanı için uygulanmıştır. 41 çalışma alanından 26'sı evrensel bir modelin tasarımı için bir ELM algoritmasına dâhil edilmiştir ve geri kalanı model çapraz doğrulama için kullanılmıştır. Evrensel olarak eğitilmiş bir ELM oluşturulup ve uzamsal olarak konuşlandırılabilir model 15 test çalışma alanında uygulanmıştır. Optimum ELM modelinden katmanların özelliğini elde etmek için gizli katman etkinleştirme işlevlerini optimize etmek adına deneme yanılma yoluyla elde edilip ve rekabetçi yapay zekâ algoritmaları olan RF (Random Forest), M5 Tree ve MARS (Multivariate Adaptive Regression Spline) ile kıyaslanmıştır. İstatistiksel ölçümler, evrensel olarak eğitilmiş ELM modelinin çok iyi bir doğruluğa sahip olduğunu ve RF, M5 Tree ve MARS modellerinden daha iyi performans gösterdiğini ortaya koymuştur.

Şahin ve ark.(2012), Türkiye üzerinden aylık ortalama günlük güneş radyasyonunu tahmin etmek için MLR (Multiple Linear Regression) ve ANN'nin

tahmin kapasiteleri araştırılmıştır. Uydu verileri Türkiye üzerinde 73 farklı lokasyon için kullanılmaktadır. Yer yüzey sıcaklığı, yükseklik, enlem, boylam ve ay bilgileri ANN ve MLR yöntemlerini kullanarak güneş radyasyonunu modellemek için girdi değişkenleri olarak kullanılmıştır. Güneş radyasyonu tahminleri, istatistiksel temeller kullanılarak meteorolojik değerlerle değerlendirilmiştir. Elde edilen sonuçlar, ANN modelinin MLR modeli ile karşılaştırıldığında daha iyi bir performans elde edebileceğini göstermiştir. Ayrıca, meteorolojik istasyon verilerinin kullanılması yerine uydu verilerinin kullanılmasında güneş radyasyonu tahmininde daha doğru sonuçların elde edildiği belirtilmiştir. Son olarak, yerleşik ANN modeli, Türkiye üzerinden günlük güneş radyasyonunun yıllık ortalamasını tahmin etmek için kullanılmıştır. Sonuç olarak, Türkiye için uydu tabanlı bir güneş radyasyonu haritası oluşturulmuştur.

Şahin ve ark. (2014), Türkiye'de güneş radyasyonu tahmini için ileri beslemeli sinir ağına dayalı ELM (Extreme Learning Machine) yöntemini kullanmıştır. Çalışmada ELM modeli tasarlanırken 2000-2001 için Türkiye'de bulunan 20 lokasyondan alınan çok yüksek radyometrik çözünürlüklü NOAA (National Oceanic and Atmospheric Administration) uydu verileri kullanılmıştır. Uydu tabanlı aylık yer yüzey sıcaklığı, yükseklik, enlem, boylam ve şehir bilgileri ELM'ye girdi olarak uygulanmış ve çıktı değişkeni olan güneş radyasyonu tahmin edilmiştir. ELM modelinin uygulanabilirliğini göstermek için, ELM modeli ile geri yayımlı geleneksel ANN (Artificial Neural Network) modeli arasında tahmin yeteneği ve öğrenme hızı açısından bir performans karşılaştırması yapılmıştır. ELM modelinin performansını test etmek için 2002 yılı verileri kullanılmıştır. Karşılaştırma sonuçları, ELM modelinin genel test konumları için ANN modelinden daha iyi tahmin verdiğini göstermiştir. Ayrıca, ELM modelinin ANN modelinden yaklaşık 23,5 kat daha hızlı olduğu ortaya konmuştur.

Bu çalışmada Şenkal O. (2018), güneş radyasyonu, yükseklik, hava basıncı ve sıcaklığı, su buharı basıncı, kuru hava yoğunluğu, su buharı yoğunluğu ve karışım oranları hakkında bir veri tabanı oluşturmuştur. Güneş radyasyonu için

ANN metodlarından RP'yi kullanmıştır. Tahmin için gereken basınç, nem, sıcaklık, ay ve yükseklik bilgileri Türkiye Devlet Meteoroloji Teşkilatı (TSMS) kaynaklarından almıştır. 2000'den 2002'ye kadar zamanı kapsayan verilerin 2000 ve 2001 yıllarını eğitim için 2002 yılı verilerini ise doğrulama amaçlı kullanmıştır. Çalışmasını önceki çalışmalarda kullanılan ANN metodlarıyla karşılaştırmış ve 0.9464 ile MLP (0.8570) ve LM (0.9120) yöntemlerinden daha iyi sonuç elde etmiştir.

Şenkal O. (2015), 2000-2002 yılları arası Türkiye içinde belirli yer ve tarihte (ay), belirli yükseklik, sıcaklık, basınç ve nemde, ortalama yağışlı su(PW) ve güneş radyasyonunu modellemek ve tahmin etmek için ANN metodu kullanmıştır. RP metodu ve lojistik sigmoid transfer fonksiyonu kullanılan ağda, Türkiye Devlet Meteoroloji Teşkilatı (TSMS) ve Wyoming Üniversitesi tarafından 2000-2002 yılları arasında Türkiye'de dağıtılan beş istasyondan alınan meteorolojik ölçümler eğitim verisi olarak kullanmıştır. Modelin eğitimi için 2000 ve 2001, testi içinse 2002 yılı verilerini kullanmıştır. Kullanılan istasyonlarda istasyonlarda bulunan ve ölçülen değerler arasında, PW için aylık ortalama günlük toplam ve güneş ışınlama değerleri için ulaşılan RMSE değerlerini sırasıyla 0.0062 gr/cm² ve 0.0603 MJ/m² (Eğitim şehirleri), 0.5652 gr/cm² ve 3.2810 MJ/m² (Test şehirleri) olarak bulmuştur.

Şenkal O. (2010), GRNN metodu ile güneş radyasyonunu tahmin etmek için yaptıkları çalışmada, Türkiye genelindeki 19 istasyondan Ocak ve Aralık 2002 dönemindeki coğrafi ve uydu tahmini verilerini, istasyonların 10'u eğitim, 9'u doğrulama için olacak şekilde kullanmıştır. Enlem, boylam, yükseklik, ϵ_4 ve ϵ_5 için yüzey yansırılık ve yer yüzey sıcaklığını ANN'nin giriş katmanında kullanmış ve güneş radyasyonu çıktı olarak almıştır. Elde ettiği aylık ortalama günlük toplam için tahmin edilen ve ölçülen değerler ile ANN değerleri arasındaki Ortalama Karekök Hatası (RMSE) ve korelasyon katsayısı (R^2) 0.1630 MJ/m² ve %95.34 (Eğitim istasyonları), 0.3200 MJ/m² ve %93.41 (Test istasyonları) değerlerinin yeterince iyi olduğunu belirterek geliştirilen GRNN'nin Türkiye'de güneş

radasyyonunu tahmin etmede kullanılabileceği sonucuna varmıştır.

Şenkal O. ve Kuleli T. (2009), güneş radyasyonu tahmini için bir ANN modeli öne sürmüştür. RP, SCG öğrenme algoritmaları ve lojistik sigmoid transfer fonksiyonunun kullanıldığı ağlarında 12 ilin (Antalya, Artvin, Edirne, Kayseri, Kütahya, Van, Adana, Ankara, İstanbul, Samsun, İzmir, Diyarbakır) 1997 yılı için Ağustosan Aralık ayına kadar olan meteorolojik verileri Türkiye geneline yayılarak kullanılmıştır. Girdi verisi olarak meteorolojik ve coğrafi verileri (enlem, boylam, yükseklik, ay, ortalama yayılan radyasyon ve ortalama ışın radyasyonu) kullanarak çıktı olarak güneş radyasyonunu elde etmişlerdir. Güneş radyasyonu, Meteosat-6 uydusu C3D verileri ile aylık ortalama günlük toplam olarak tahmin etmişlerdir. Uydu verilerinden öncelikle albedoları hesaplamışlar ve bunlardan bulutluluk indeksi elde etmişlerdir. Açık gökyüzü indeksi ve bulutluluk indeksi ile fiziksel yöntem için piksellerin güneş ışımasını hesaplamışlardır.

Chen ve ark.(2021), yaptığı çalışmada, Amerika Birleşik Devletleri Kıtası (CONUS) üzerinden yarım saatlik ve günlük küresel Aşağı Güneş Radyasyonunu (DSR) tahmin etmek için GOES-16 uydu verileri kullanılmıştır. GOES-16 uydusunun 16 bandındaki tüm veriler, 1 km uzamsal çözünürlükle DSR'yi tahmin etmek için MLR ve RF modellerinde girdi olarak kullanılmıştır. Rasgele orman (RF) modelleri, doğrusal olmayan regresyon ve sınıflandırma problemleri için güçlü topluluk algoritmasıdır. MLR ve geleneksel RF modelleri karşılaştırıldığında, RF modeli DSR'nin tahmin edilmesinde $R^2 = 0.95$, $RMSE = 66.92$, $MBE = 0.06$ analiz oranlarını vermiştir. Ayrıca aynı çalışmada yarım saatlik DSR'a göre saatlik ve günlük DSR tahminlerinde daha yüksek doğruluk değerlerine ulaşılmıştır. Bu değerler saatlik tahminler için $R^2 = 0.96$, $RMSE = 60.44$, $MBE = -0.23$, ve günlük tahmin değerleri için ise $R^2 = 0.97$, $RMSE = 17.64$ ve $MBE = 0.86$ olarak belirlenmiştir. CONUS üzerinden yarım saatlik ve günlük küresel DSR'nin RF modeli ile tahmin edildi modelin yardımcı verilere (örneğin bulutluluk bilgisi gibi) ihtiyacı olmadığı tespit edilmiş, bu da yöntemin basit ve verimli çalışmasını sağlamıştır. Fotovoltaik sistemler (PV) kurmaya ve güneş

enerjisi tahmin etmeye yardımcı olmak için etkili sonuçlar sunmuştur.

Cornejo-Bueno ve ark. (2019), Meteosat ikinci nesil uydularında bulunan (Dönen Gelişmiş Görünür ve Kızılötesi Görüntü), SEVIRI radyometresi 12 spektral (3 görünür, 8 kızılötesi ve 1 yüksek çözünürlüklü görünür) kanaldan gelen veriler ile fiziksel modellerden küresel güneş radyasyonu tahmin probleminde Makine Öğrenimi (Çok Katmanlı Algılayıcılar, Aşırı Öğrenme Makineleri, Destek Vektör Regresyonu ve Gauss Süreçleri) performansını karşılaştırmıştır. İspanya'nın Toledo şehrinde bulunan radyometrik istasyondan alınan veriler üzerinde açık gökyüzü modeli, güneş radyasyonu tahmini, bulut indeksi ve yansıtıcılık verileri ile Meteosat görüntüleri kullanılmıştır. Fiziksel model olarak bilinen Heliosat-2 yöntemi, Copernicus Atmosfer İzleme Hizmetine (CAMS) ve SolarGIS modeline dayanan bir model kurulmuştur. Küresel güneş radyasyonu ve yansımının bir yıllık veriler uydu verileri (2013-2014) değerlendirilmiştir. Fiziksel modeller girdi değişkeni olarak dâhil edildiğinde, ELM'nin, Rmse:62.81 R²: 95,7% değerleri ile en iyi performansı sergilediğini görülmüştür. Makine Öğrenmesi SVR ise çok sayıda girdi değişkeninden etkilenmiştir. SVR değişkenler azaltılıp 5 girrdi veri kullanımı ile iyileşmiş sonuçlara ulaşılabilmektedir. Bu durumda ise, ELM algoritması ile Rmse=60.60 ve R²=96 ile en iyi güneş radyasyonu tahminini elde edilmiştir.

Linares-Rodriges ve ark. (2013), İspanyanın Katalonya bölgesindeki geniş çalışma bölgeleri için yaptıkları bu çalışmada günlük küresel güneş radyasyonunu tahmin etmek için yapay sinir ağı modeli oluşturmuşlardır. Bu modelde girdi değişkenleri olarak gökyüzü açıklığını ifade eden bir değişkeni ve Meteosat uydusunun 11 kanalına ait olan uydu görüntülerini kullanmışlardır. Kurulan ANN modeli MLP algoritması olup bir gizli katmanlı 25 nöronludur. Bir çok çalışmada kullanılan görünür band verilerinin aksine bu çalışmada kızılötesi kanal verileride girdi olarak algoritmaya dahil edilmiştir. Meteosat 9 uydusunun kızılötesi kanalları. Ocak 2008'den Aralık 2008 zamansal aralığına sahip olan çalışmada 83 adet istasyonun 65 eğitimde 18 ise test amaçlı kullanılmıştır. Çalışmada elde edilen hata

RMSE hata 4% ile 9% arasında bulunmuştur.

Rahimikhoob ve ark.(2013), İran için yaptıkları bu çalışmada NOAA-AVHRR uydusu bantlarını girdi parametresi olarak kullanan bir ANN modeli kurmuşlardır. MLP algoritması üzerinden kurulan bu modelde tek gizli katman kullanılmış nöron sayısı ise yapılan denemelerde en iyi sonucu veren değer olan 18 olarak belirlenmiştir. Çalışmada 2005-2008 yıllarını kapsayan 661 adet uydu görüntüsünün %70 eğitim için geri kalan %30 test için kullanılmıştır. Ham bant verileri değişik kombinasyonlar ile ANN modelinde kullanılarak yüzeye gelen güneş ışıınımı değerleri ile mukayese edilmiştir. Bu karşılaştırmada AVHRR 'ın 5 bandının girdi olduğu 18 nöronlu MLP algoritmasında en iyi sonuç $R^2=0,70$ ve RMSE ise 100.89 W/m^2 olarak elde edilmiştir.

3. MATERYAL ve METOT

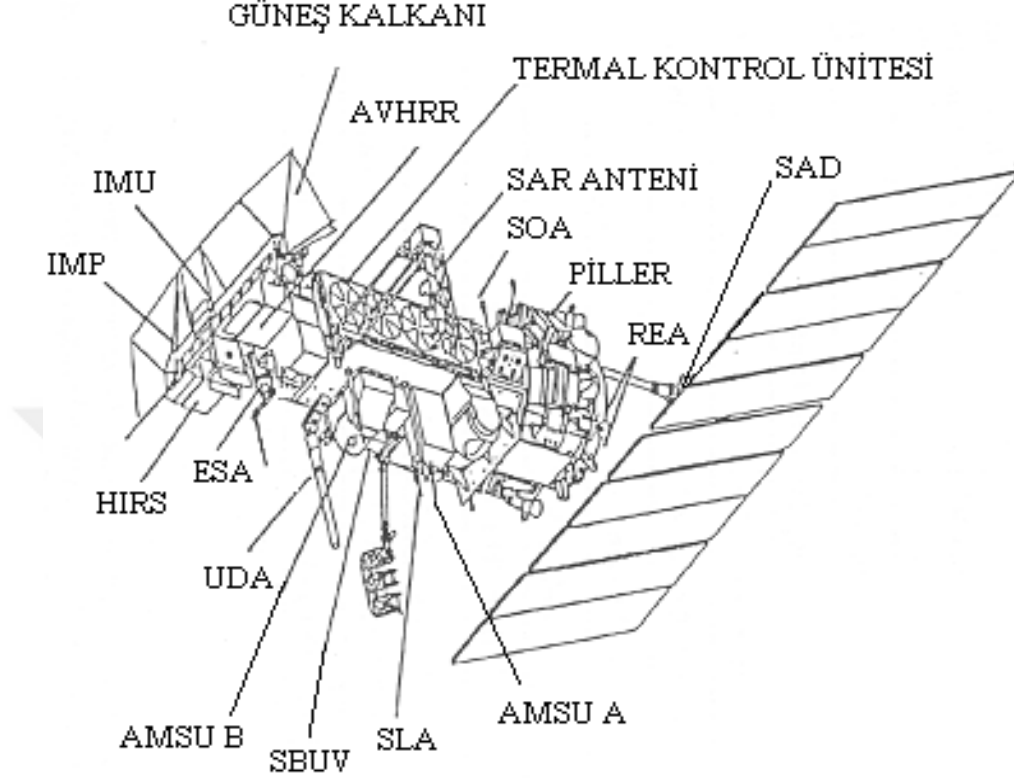
3.1. NOAA Uyduları

Ulusal okyanus atmosferik teşkilatı (NOAA) küresel okyanuslar, atmosfer, uzay ve güneş ile ilgili bütün çalışmaları yapan, verileri toplayan ve bu elde etmiş olduğu verileri bilime uygulayan bir kuruluştur. NOAA şiddetli hava hadiseleri için uyarılar yapmakta, denizler ve gökyüzü için haritalar çıkarmakta, okyanus ve sahil sel kaynakların kullanımı ve korunması için gerekli önlemleri almakta ve doğayı anlamamız konusunda geliştirici araştırmalar yapmaktadır.

NOAA uyduları yerden yaklaşık 830-870 km. yükseklikteki yörüngelerinde yer almaktadır. Bu Uydular kutupsal yörüngeli uydulardır. Yani kutuptan kutuba dünya etrafında tur atmakta ve taradığı bölgelerin görüntülerini dünya üzerindeki alıcı merkezlerine ulaştırmaktadır. Dünyanın da kendi etrafında dönmesi sebebiyle her turda değişik alanı tarayabilirler.

1998 baharında NOAA-K adlı uydunun başarıyla yörüngeye oturtulmasıyla yeni bir uydu serisi NOAA (POES) atmosferde görevine başlamıştır. NOAA-K' nin başarılı operasyonundan sonra NOAA-L(6-12) ve NOAA-M(13-17) ve NOAA-N(18-19) adlı TİROS-N (1978) serisinin en son gelişmiş serisi yörüngede yerini almıştır. Bu NOAA-POES serisi yeni uydular eski serilere göre birçok değişiklik içerirler. Bunun nedeni bu uyduların yapılarına eklenen yeni algılayıcılar ve sistemlerdir.

Şu an aktif olarak görev yapan, yeni dizayn edilmiş, Çok Yüksek Çözünürlüklü Gelişmiş Radyometre (AVHRR/3) algılayıcısına sahip olan NOAA-18, NOAA-19 uyduları bulunmaktadır. Bu çalışmada Mayıs 1998' de atılıp görüntü almaya başlayan NOAA-15 ve Ekim 2002' de görüntü almaya başlayan olan NOAA-16 ile NOAA-17 ve yeni dizayn edilmiş AVHRR/3 algılayıcısına sahip olan NOAA-18, NOAA-19 uydularının verileri kullanılmıştır. Şekil 3.1' de NOAA uydusu ve verilerinde yer alan algılayıcı birimler şematik olarak gösterilmiştir.



Şekil 3.1. NOAA uydusu ve üzerinde yer alan algılayıcı sistemler (NOAA KLM Guide)

Amerika Birleşik Devletlerince planlanıp uzaya gönderilen ve faaliyetlerine halen devam eden bu uyduların günümüz bilim adamlarının ihtiyaçlarını ve veri gereksinimlerini karşılayabilmesi için sürekli olarak yeni geliştirilen algılayıcı platformlarla değiştirilmeleri gerekmektedir. Bu çalışmada NOAA Uydusu üzerinde yer alan AVHRR/2 ve AVHRR/3 algılayıcısından elde edilen veriler kullanılmıştır.

3.1.1. Çok Yüksek Çözünürlüklü Gelişmiş Radyometre (AVHRR)

Bu dedektör görünür bölgede algılama yapan kanalların düşük enerjili ışıkları algılamasında ve spektral alanda büyük kazanç sağlar. Ayrıca 3A diye adlandırılan 1.6 μm de kar ve buz dağılımını inceleyen bir ek kanal içerir.

Yağlama ve motor aksamında yapılan değişimlerle tarama mekanizmasının ömrü uzatılmıştır, ayrıca yeni eklenen bir güneş kalkanı ile AVHRR' in tarama detektöründe güneş ışınlarından dolayı ortaya çıkacak kalibrasyon ve gürültü olaylarının önüne geçilmiştir.

Altı kanallı tarama özelliğine sahip AVHRR/2-3 dedektörünün her kanalı dünya üzerindeki tarama alanını aynı anda görür. Her tarayıcıdan elde edilen bilgiler kullanıcılar tarafından yapılan işlemler sonrasında hidrolojik, okyanus bilimleri, meteorolojik arazi kullanımı üzerine çok bantlı analizler yapılabilecek hale gelir. 1, 2 ve 3A bandında elde edilen veriler elektromanyetik spektrumun görünür ve yakın kızılötesi bölgesinde yer yüzeyinden yansıyan enerjinin algılanmasıyla elde edilir. 3B, 4 ve 5 bantlardan elde edilen veri yer yüzeyinin, suyun, deniz yüzeyinin ve bunların üzerinde bulunan bulutların sıcaklıklarından dolayı yaydıkları enerjinin belirlenmesi ile elde edilir. Bu 6 bantta elde edilen verilerden yalnız 5 tanesi aynı anda dünyaya gönderilir. 3A ve 3B kanallarından elde edilen veriler görüntünün gündüz veya gece alınmasına bağlı olarak kendi aralarında değiştirilerek gönderilir. 3B verileri daha ziyade gündüz verilerinde gönderilir. Çizelge 3.1'de AVHRR' ın bantları ve özellikleri verilmiştir.

Çizelge 3.1. AVHRR/2-3 algılayıcısının bantları ve bant özellikleri (NOAA KLM Guide)

BANT ÖZELLİKLERİ		NOAA 9-11-12-14	NOAA 15-16-17-18-19
Görünür Bölge	Bant - 1	0,55 - 0,90	0,58 - 0,68
Yakın Kızılötesi	Bant - 2	0,72 - 1,00	0,82 - 0,87
Yakın Kızılötesi	Bant - 3	3,55 - 3,93	-
Yakın Kızılötesi	Bant - 3a	-	1,57-1,78
Yakın Kızılötesi	Bant - 3b	-	3,55 - 3,93
Kızılötesi	Bant - 4	10,5-11,5	10,3 - 11 ,3
Kızılötesi	Bant - 5	11,5 - 12,4	11,5 - 12,4

AVHRR genelde meteorolojik görüntüleme sistemi olarak dizayn edilip NOAA-KLM serisinde yerini almıştır. Bu tarama sistemi çok küçük bir görüş açısı (1.3 miliradyan) ile dünyayı bir ufkundan diğer ufkuna düzgün bir tarayıcı ayna ile tarayan bir sistemdir.

Taranan alanların yönelimi uydu yörüngesine dik ve tarama aynasının dönme hızı arka arkaya gelen tarama satırları sürekli uydunun nadir pozisyonuna bakacak şekilde ayarlanmıştır. Uydu 893 km yukarıda ve 55.4 derecelik bir nadir açısına sahiptir. AVHRR' nin 6 kanalı aynı zamanda dünyanın aynı yerini tarayacak şekilde dizayn edilmiştir. Bu altı kanaldaki verinin 5' i aynı anda istasyona iletilir. AVHRR' de bulunan radyometreler uydu yörüngesinde 3 sene çalışacak şekilde hesaplanmıştır.

AVHRR/2-3 6 kanallı tarama radyometresidir. Bu 6 banttan 3 tanesi görünür ve yakın kızılötesi bölgede diğer 3 tanesi termal kızılötesi bölgede tarama yapan kanallardır. AVHRR/2, 10.3µm ve 12.5µm aralığında 2 tane, 1µm genişliğinde tarama yapan bant bulunur. Alet 20.32 cm çapında toplayıcı bir teleskop çeşidi olan kırıcı Cassegrain teleskopunu kullanır. 3B, 4 ve 5. kanallarda kullanılan termal kızılötesi detektörler 2 basamaklı pasif radyant soğutucular ile soğutulur.

Tarayıcıdan alınan sinyaller uyduya bakan MIRP(Manipulated Information Rate Processor) ünitesine gönderilir. Sonra örnek veri sinyaller sırayla tekrar AVHRR'a gönderilir. Altı banttan elde edilen verileri aynı anda 40 kHz de örneklenir ve 10 bitlik 2 tabanlı kodlar haline getirilir.

AVHRR verileri sadece belirli zaman aralıklarında dünya yüzeyini tararlar, tarama işlemi sürekli değildir. Aktif 5 kanaldan elde edilen veriler radyometrelerde 10 bitlik veri halinde sayısallaştırılır ve uydudaki MIRP adlı ünitenin çıkışına 10 bitlik sayısal veri olarak gönderilir.

AVHRR tek bir ünite içerisinde bir araya getirilmiş 5 modülden oluşur. Bu modüller;

- Tarama modülü
- Elektronik modül
- Radyant soğutma modülü
- Optik Sistem
- Ana Gövde

Ana optik sistem 2 kısımdan oluşmuştur; bir toplayıcı teleskop ve düzenleyici optik ünite. Teleskop 20.3 cm çaplı yansıtıcı Cassegrain tipidir. Düzenleyici optik ünite teleskopun topladığı ışığı 6 spektral banda göre böler ve bunları kendi kanallarına odaklar.

Bu bantlar şunlardır :

Kanal 1.....	0.58-0.68 μm
Kanal 2.....	0.725-1.00 μm
Kanal 3A.....	1.58-1.64 μm
Kanal 3B.....	3.55-3.93 μm
Kanal 4.....	10.3-11.3 μm
Kanal 5.....	11.5-12.5 μm

1, 2 ve 3a kanalları için polarizasyon etkileri optik ışık demetinin aynı tip 2 ışık bölücünden geçirilmesi sonucu minimize edilmiştir.

AVHRR'nin aldığı veriler MIRP ünitesinde anlamlandırılır. AVHRR'nin dijital çıkış zamanlayıcı devresine bir veri örneği atımı gönderir ki bu zamanlayıcı devre her radyometrik kanal için gelen ardışık verileri sonlandırır. Bu MIRP ünitesi verileri şu 4 halde işler:

- Otomatik Veri Transferi (APT).
- Global Alan Tabakası (GAC).

- Bölgesel Alan Tabakası (LAC).
- Yüksek Çözünürlüklü Veri Transferi (HRPT).

Hem dedektörden alınan analog bilgi-10 bit sayısal örneğe bir analog sayısal çevirici yardımıyla, kontrollü olarak çevrilir. Bu çeviriciye MIRP denir. MIRP tarafından işlenen veriler uydu tarafından yer istasyonuna gönderilir.

Global Alan Tabakası (GAC); MIRP ünitesinde GAC veri tipi işlenmiş AVHRR/2 ile TIP verilerinin birleştirilmesiyle oluşturulur. Her 1 saniyede 2 tane GAC veri penceresi oluşturulur ki AVHRR veri penceresinin 3 te biridir. Her 3 AVHRR tarama çerçevesinin 1 tanesinin kullanılması ile GAC ve AVHRR verileri birbirlerine uyumlu hale gelir. Bunun için MIRP ünitesi AVHRR'ın her kanalının her tarama sırasında komşu 4 örneğinden 1'ini ihmal etmesiyle uyum sorununu bozar. 10 bitlik veri oluşturulurken 8 bitlik TIP verileri 2 bitlik ekleme ile 10 bit haline getirilir. Bu veriler direkt kullanıma elverişli değildir.

MIRP ünitesinde HRPT verileri ve LAC çıktı verileri işlenmemiş AVHRR ve TIP verilerinin birleştirilmesi ile oluşur. HRPT verilerinin pencere ve veri oranı AVHRR ile aynıdır. Fakat TIP verileri HRPT verilerinin 3' de biri olduğundan TIP verileri 3 defa tekrarlanmak zorundadır. GAC verilerinde olduğu gibi TIP verilerine de 2 bitlik ekleme yapılır. LAC; Kaydedilmiş HRPC verisidir. Uydunun DTR ünitesine GAC gibi kaydedilir.

3.1.2. NOAA-AVHRR Termal Kanal Verilerinin Kalibrasyonu

Termal kanalların her ölçüm için skala edilmiş eğim ve kesişme değerlerinin birimi $mW/(m^2-sr-cm^{-1})$ 'dir. Sensörün herhangi bir i kanalı için ölçülmüş radyans değeri giriş verilerini bir lineer fonksiyonu olarak şöyle ifade edilir;

$$E_i = S_i C + I_i \quad (3.1)$$

Burada E_i i. Kanalin $\text{mW}/(\text{m}^2\text{-sr-cm}^{-1})$ biriminde hesaplanan radyans değeri, C giriş verisi (0 ile 1023 arasında değişir) ve S_i ile I_i 'lerde i. kanala ait eğim ve kesişim değerleridir. NOAA-15 AVHRR algılayıcısıyla başlayan, AVHRR' ın 4. ve 5. kanalının radyan enerji lineersizliğini ve 3. kanalın kalibrasyonundan kaynaklanan hataları düzelten yeni bir veri kalibrasyon metodu NESDIS tarafından oluşturulmuştur. Gelecekteki araçlarla birlikte, her bir kanal için artış ve eğim kalibrasyonun katsayıları, kullanıcıya lineer kalibrasyonun yapılması için imkan veren level 1b veri akımının içine koyulur. Ancak, bu katsayılar, her bir kanal için bir sabit olan ve iç kalibrasyonun hedef sıcaklığından bağımsız parlaklık düzeltmeleri yapan sıfır olmayan uzay parlaklığı R_{sp} kullanarak elde edilir.

$$S = \frac{R_T - R_{SP}}{C_T - C_{SP}} \quad (3.2)$$

$$I = R_{SP} - S \times C_{SP} \quad (3.3)$$

burada S eğim (slope), I kesişim (intersept), R_T hedef parlaklık , C_T hedef sayısı ve C_{sp} uzay sayısıdır. Bu katsayılar aşağıda tanımlanan, lineer bir parlaklık sağlar. R_{LIN} :

$$R_{LIN} = S_i \times C_i + I_i \quad (3.4)$$

Bu lineer kalibrasyonun prosedürü, NOAA Technical Memorandum NESS 107' de daha detaylı bir şekilde anlatılmıştır. Aşağıdaki denklem, AVHRR kanalları 3, 4 ve 5 için, eşitlik 3.4'ten elde edilen lineer parlaklığın bir fonksiyonu olan düzeltilmiş parlaklık radians enerjisi sağlar.

$$E = A \times R_{LIN} + B \times R_{LIN}^2 + C \quad (3.5)$$

Burada A, B ve C kanallara ait katsayılarıdır değerleri NOAA kullanım klavuzunda yer almaktadır. Radyans enerji değerlerinden parlaklık sıcaklığı değerlerine geçiş ters Planck ışıma eşitliği diye adlandırılan bir eşitlik yardımıyla gerçekleşir.

$$T(E) = \frac{C_2 V}{\ln \left(1 + \frac{C_1 V^3}{E} \right)} \quad (3.6)$$

Burada T, E radyans değerine karşılık gelen ⁰K cinsinden sıcaklık, V ölçüm yapılan sensör kanalının merkezi dalgaboyudur (cm⁻¹). Çizelge 3.6'da merkezi dalgaboyu değerleri yer almaktadır. C₁ ve C₂ sabitler olup C₁=1.1910659x10⁻⁵ mW/(m²-sr-cm⁻¹) ve C₂=1.438833 cm-K değerlerine sahiptir. Elde edilen T(E) değerine ek sıcaklık düzeltmesi uygulanmaz. NOAA-AVHRR algılayıcısının 3., 4. ve 5. Bantları için bu denkleme uygulanan katsayılar Çizelge 3.6'e dahil edilmiştir. NOAA-15 ve sonrası uydu verilerinden sıcaklık hesaplaması yapabilmek için eşitlik (3.6)'dan elde edilen T(E) değeri eşitlik (3.7) yardımıyla düzeltilmelidir. Çizelge 3.2 bu düzeltme ve eşitlik (3.6) yardımı ile hesaplama yapabilmek için gerekli olan sabitleri içermektedir.

$$T = [T(E) - A] / B \quad (3.7)$$

Çizelge 3.2. NOAA-15 ve Sonrası AVHRR uydusu için sıcaklık düzeltme ve merkezi dalgaboyu katsayıları

Kanal	Merkezi dalgaboyu	A	B
Kanal 3	2695.9743	1.621256	0.998015
Kanal 4	925.4075	0.337810	0.998719
Kanal 5	839.8979	0.304558	0.999024

3.1.3. Bitki Örtüsü İndeksinin Hesaplanması

Yer kabuğu genel çizgilerle ayırt edildiğinde aşağıdaki ana örtü tipleri ile kaplı bulunmaktadır (Van De Griend ve Owe, 1993).

- Bitki örtüsü (Orman ve ağaçlar, otlak ve kültür bitkileri),
- Kaya toprak kompleksi,
- Su yüzeyleri,
- İnsan yapısı materyallerdir.

Yeryüzü objeleri içinde uzaktan algılama yöntemi ile kolayca tanınıp incelenen objeler bitki örtüsüdür. Bunun nedeni bitkinin yakın kızılötesi bölgede ışığı çok fazla yansıtmasıdır (Dinç ve ark., 2001).

Spektrumun görünür bant bölgesinde, gönderilen elektromanyetik radyasyonun yeşil bitkilerden olan yansıma miktarı %20'yi aşmamakta, fakat kızılötesi bant bölgesindeki yansıma miktarı %50-60 civarında olmaktadır. Buna karşın, su yüzeyleri, bulutlar ve kıraç arazi, her iki bant bölgesinde aynı yanıtı vermektedir. Bu nedenle NOAA uydularının 1. ve 2. kanallarındaki elektromanyetik dalga yayılımı ile yapılan ölçümler arasındaki fark bitki örtüsü ile kaplı bölgeler için bir gösterge olmaktadır. İnceleme bölgesinin bulutlarla kaplı olması, bu çalışmalarda problem oluşturmaktadır. Rayleigh ve Mie saçılmaları ve bulutlar, özellikle kızılötesi bant bölgesinde elektromanyetik radyasyonun azalmasına ve bitki örtüsü indeksinin düşük hesaplanmasına neden olmaktadır.

NOAA uydusu verileri için NDVI aşağıdaki bantlar cinsinden formülize edilir.

$$NDVI = (Ch\ 2 - Ch\ 1) / (Ch\ 2 + Ch\ 1) * 255 \quad (3.8)$$

veya dalgaboyu bölgeleri cinsinden,

$$NDVI = (NIR - VIS) / (NIR + VIS) * 255 \quad (3.9)$$

Burada NIR ve VIS değerleri yakın kızılötesi ve görünür bant bölgelerinde yüzeyden olan yansıtma miktarlarıdır. Bulunan NDVI değerleri -1 ile 1 arasından 0 ile 1 arasına çekilir burada (-) değerler ve 0 bitki örtüsü olmayan yüzeylerdir bu nedenle herhangi bir yanlışlığa imkan tanınmaz zira yayınlık hesaplanırken logaritma alınması gerekmektedir ve (-) değerlerin logaritması yoktur.

3.1.4. Temel Yer Yüzeyi Objelerinin Yayınlığı

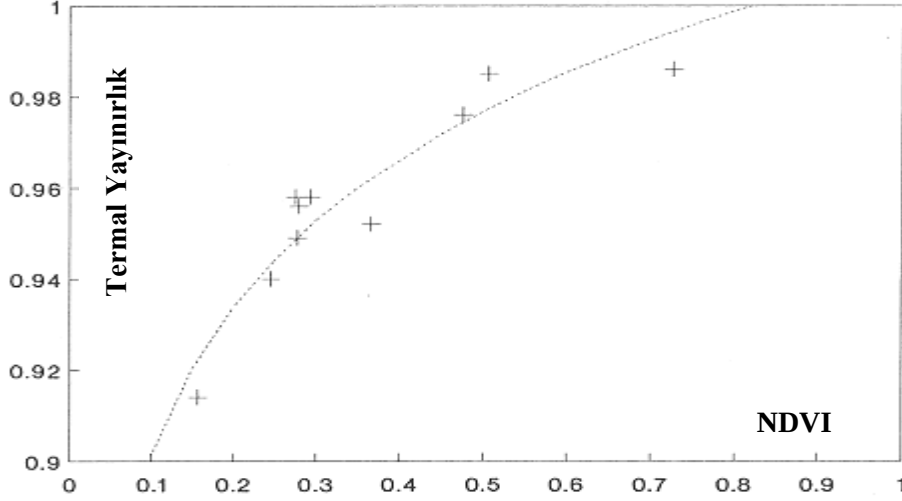
Yer yüzeyi yayınlığının belirlenmesi ve ölçülmesi için birçok yöntem vardır. Bu yöntemlerin en başında NDVI yardımı ile yayınlığın belirlenmesi gelir. Yüzey yayınlığı bir cismin soğurduğu enerjiyi cisimden dışarı yayma yeteneğidir ve 0 ile 1 arasında değer alır.

Yayınlık hesaplamaları Becker (1987) tarafından dizayn edilmiş 8-14 µm aralığında ölçüm yapan portatif bir yayınlık kutusu yardımıyla yapılmıştır. Bu kutunun iç yüzeyleri yüksek yansıtma kapasitesine sahip olan alüminyum ile kaplanmıştır. Aynı zamanda kutunun üst kısmına bir Kara cisim kaynağı alt kısmına ise termal kızılötesi bölgede çalışan bir radyometre yerleştirilerek düzenek bilinmeyen yüzeylerin yayınlıklarını belirlemek için kullanışlı hale getirilmiştir.

Van de Griend ve Owe (1993), bu düzeneği kullanılarak 10 değişik yüzey tipinin NDVI değerleri ve termal yayınlıkları ölçülmüştür. Bu ölçümler sonunda elde edilen standart sapma değerlerinin birçok durumda %1 den küçük olduğu görülmüştür. Elde edilen verilerden NDVI ve yayınlık arasındaki bağıntı aşağıdaki logaritmik fonksiyon yardımıyla açıklanmıştır.

$$\epsilon = a + b \log (\text{NDVI}) \quad (3.10)$$

a ve b bağıntı sabitleri 8 – 14 µm aralığındaki spektral bölge için belirlenmiştir ve değerleri sırasıyla 1.0094 ve 0.047 dir. Şekil 3.26' da yayınlık ile NDVI arasındaki ilişki görülmektedir.



Şekil 3.2. Yayınırılık ile NDVI Değerleri Arasındaki Bağntı (Griend ve Owe, 1993)

NOAA-AVHRR verilerinde yayınırılık tespitinin zorluğu, verideki bir pikselin dünya yüzeyinde 1.1 x 1.1 km' lik bir alanı göstermesi ve bu alandaki yüzey bitki örtüsünün ve doğal olarak bitki örtüsüne bağlı olarak yayınırılık değerlerinin farklı olmasıdır. Yersel çözünürlüğün düşürülmesi ile pikselde bulunan bitki grupları arasındaki yayınırılık faktöründen doğan spektral kontrast ortadan kalkar. Bu problemi çözmek için yapılabilecek yaklaşımlardan biri pikselin ortalama yayınırılık değeri olduğunu kabul ederek radyometrik sıcaklığın belirlenmesini sağlamaktır (Wan, 1999).

Bitki örtüsü olmayan çıplak yüzeylerin yayınırılıkları tatmin edici derecede belirlenebilir. Bu tip yüzeylerde yayınırılık yüzeyi oluşturan madde ve onun yapısına, suyu tutma derecesine bağlıdır ve genelde 0.95'in altında değerler alır. Örneğin 8 ile 14 μm dalgaboyu aralığında granit 0.815 ve asfalt 0.959 değerindedir (Griend ve Owe, 1993). AVHRR verisiyle LST' yi hesapladığımız zaman, yüzey yayınırılığının önemi ve LST' yi nasıl etkileyeceğini bilmemiz gerekir. Termal bölgede(8-14 μm) farklı materyallerin, farklı yayınırılıkları vardır. Çizelge 3.3'te 300 °K de, termal bandın 8-14 μm aralığında yayınırılıklar verilmiştir.

Çizelge 3.3. Bazı maddelerin yayınlık değeri (Qin, 1999)

Materyal	ϵ	Materyal	ϵ
Deniz suyu	0.98	Ağaç yaprağı	0.96
Buz	0.96	Odun	0.90
Kar	0.85	Granitik kaya	0.89
Nemli toprak	0.95	Silicon Kum	0.909
Kuru toprak	0.92	Dolomite	0.958

Yüzey yayınlığının başka bir özelliği ise tüm dalga boyları için sabit olmamasıdır. Yer yüzey yayınlığının, uydu algılayıcılarının çeşitli “termal kanallarında” önemsiz farklılıkları vardır. AVHRR verileri için, kanal 4 ve 5 in yüzey yayınlıklarının genellikle 0.002-0.005 gibi farklar vardır yer yüzey örtüleri ise bu bantlarda genel olarak 0.95 ile 0.98 gibi yayınlık değerlerini alır (Li ve Becker 1990).

Toprak yayınlığının, toprak çeşidine ve nemine göre oldukça değişken olmasına rağmen, yapılan çalışmalar, toprak yayınlığının değeri 0.84-0.98 örtü yayınlığının değeri 0.94-0.99 ve doymuş örtü yayınlığının değeri 0.98-1.0 arasında olduğunu göstermiştir. Birçok çalışmada termal spektral bölgeler için, yer yüzeyinin bir kara cisim gibi olduğu farz edilmiştir ve yayınlığı 0.96-0.98 değerleri aralığında alınmıştır. Bitki örtüsünün fazla olduğu tarımsal alanlar için yaprak yayınlığının değişimden dolayı 10,5-12,5 μm ’ de gözlenen radyans değerlerinin yüzey sıcaklığı hesaplamasında kullanılmasında 1,5 $^{\circ}\text{K}$ ’lik fark görülebilir (Wan, 1999).

3.1.5. Split-Window Metodu ile Yer Yüzey Sıcaklığının Hesaplanması

NOAA-AVHRR verilerinden yüzey sıcaklığı hesaplaması yapmak için birçok değişik metod bulunmaktadır. Bunlardan en yaygın olanları sırası ile; Tek kızılötesi kanal metodu, Çok-Açı metodu, Çok-Kanal metodu ve Split-Window metodudur. Bu dört metod arasında uygulanması ve sonuçlarının başarılı olması yönünden Split-Window metodu literatürde diğer uygulamalara göre öne çıkmaktadır.

İlk tip Split-window algoritmaları 1970’li yıllarda NOAA-AVHRR verilerinden deniz yüzey sıcaklığı hesaplamaları için oluşturulmuştur. Bu algoritmalar çok defa yeni adaptasyonlar yapılarak iyileştirilmiş böylece geliştirilerek başarılı sonuçlar elde edilmesi sağlanmış ve bugüne kadar gelinmiştir. Kuru atmosfer şartlarında 0.3 °K ve nemli atmosfer şartlarında 1.0 °K doğruluklarla hesaplamalar yapılmıştır (Czajkowski ve ark. 1998). Bu algoritmalar çeşitli bilim adamları tarafından yüzey sıcaklığı belirlemek için kullanılmış ve değişik doğruluk oranları ortaya çıkmıştır. Bulunan bu değerlerde hata oranları 6-10 °K arasında değişmektedir. Bunun üzerine bu metodlar üzerine daha doğru sonuçlara ulaşmak için adaptasyonlar yapılmış bu adaptasyonda yer yüzeyi ile ilgili sabitler algoritmalara eklenmiştir. Böylece yeni algoritmalar oluşturulmuştur. Bunlardan en önemlileri Becker ve Li (1990), Sobrino ve ark. (1991), Prata (1994), Ulivieri (1992) dir Split-window tekniği birbirinden farklı, fakat birbirine yakın yanyana 2 dalgaboyu aralığındaki değişik atmosferik soğurulma karakteristiğini içerir. Bu kanalların ölçülebilen ısı farkı atmosferden geçen ışınların etkisiyle yorumlanır.

Yüzey sıcaklığı belirlemede kullanılan algoritmalar, atmosferik yarı-doğrultulmuş sıcaklık verileri içeren termal kanalların lineer kombinasyonundan oluşmuştur. Yüzey sıcaklığı hesaplamalarında çalışma bölgesinden yerden ölçülen sıcaklık verileri veya değişik modeller yardımıyla hesaplanan sıcaklık verileri için değişik düzeltme algoritmaları ortaya çıkarılmıştır. Split-window algoritmalarının temel formülasyonu şudur;

$$T_{sr} = a_0 + a_1 T_{kanal\ 4} + a_2 T_{kanal\ 5} \quad (3.11)$$

T_{sr} = Yüzey sıcaklığı

$T_{kanal\ 4}$ = 4. kanal için parlaklık sıcaklığı

$T_{kanal\ 5}$ = 5. kanal için parlaklık sıcaklığı

a_0, a_1, a_2 ‘ler sabitlerdir.

Bu metodun temeli, her kanalın yayınlık değelerinin bilinir olmasıdır. Split-Window algoritmalarındaki sabitler 4. ve 5. kanalların yayınlık değelerine çok güçlü bir şekilde bağıdır (Cihlar ve ark. 1997).

Yüzey yayınlığı ve atmosferik etkilerin ikisinin de bilinmemesinden dolayı Split-window eşitliğinin tek bir çözümü yoktur bu belirsizliği çözmek için Split-window algoritmalarından aşağıda belirtilen bazı yaklaşımlar yapılır (Vogt, 1996; Wan, 1999).

- Tüm sıcaklıkların tek olması. Bu yaklaşımla parlaklık sıcaklığı, yüzey sıcaklığı, hava sıcaklığı tek olarak (aynı) kabullenilir böylece atmosferik etkiler azaltılır.
- Yüzey yayınlığının zamanla değişimi stabildir ve 1 değerine yakındır.
- Yüzey “lambertian yüzey” olarak düşünülür.
- Planck fonksiyonu lineerdir. Böylece bir kanal için ölçülen ışıma bu fonksiyonun lineer olmasının sayesinde diğer kanalları da açılabilir.
- Atmosferdeki soğurulum küçüktür ve atmosferin en alt tabakasında gerçekleşir. Bu ortamda az su buharı bulunması halinde geçerlidir.
- ϵ_4 ve ϵ_5 özdeşdir. $\epsilon_4 > \epsilon_5$
- Yüzey ısısı 305 K' i aşmaz

Bu varsayımlar deniz yüzey sıcaklığı hesaplamaları içindir. Split-window tekniği kullanarak yer yüzey sıcaklığı hesaplanırken bu varsayımlarla çalışılır. Bu varsayımlar LST hesaplamaları için pek geçerli değildir (Vogt, 1996; Becker ve Li, 1990).

Yüzey sıcaklığı hesaplamaları ile deniz yüzey sıcaklığı hesaplamalarının kabullenimleri arasında ilk olarak aşağıdaki farklılıkları sayabiliriz;

- Yüzeylerin yayınlılık değerleri zamana ve ortama bağlı olarak hızla değişebilir.
- Yayınlılık bir piksel içinde, pikselin boyutuna bağlı olarak değişir ki NOAA-AVHRR pikselleri 1.1 km x 1.1 km dir.
- Yüzey yayınlılık değerleri spektral tayfa bağlılık gösterir. Her ne kadar 4. kanalın ve 5. kanalın spektral yayınlılık değerleri arasında 0.01'lik küçük bir fark olmasına rağmen bu fark hesaplamaları etkiler.
- Yüzey sıcaklığında günün dönemlerine göre sıcaklık değerleri deniz sıcaklığına göre dalgalanmalar gösterir.
- Havanın sıcaklığı ile yer yüzeyinin sıcaklığı farklıdır. Bundan dolayı bu sıcaklık değerleri aynı kabul edilemez.
- Atmosferin etkisi ışığın aldığı yola bağlı olarak değişik yüksekliklerde değişik etkilere yol açar.
- Efektif yayınlılık, yani algılayıcıdan ölçülen yayınlılık görüş açısına ve anizotropik yüzeylerin spektral özelliklerine bağlıdır (Sobrino, 1991).

Bütün bu faktörlere rağmen, yüzey sıcaklığı hesabını ve değerini etkileyen en büyük etken ve problem bilinmeyen yüzey yayınlılık değeridir. Birçok araştırmacıya göre yayınlılık değerlerinin tahmininde yapılacak düzeltmelerin, atmosferik etkilerde yapılacak düzeltmelere göre Split-Window algoritmalarında daha büyük etkisi vardır (Li ve Becker 1993). Şu anki bilgiler ışığında yüzey sıcaklığı çalışmaları için Split-window algoritmalarının kullanımı en iyi yöntemdir.

Split-window algoritmalarını kabaca aşağıdaki 5 başlık altında toplayabiliriz;

1. Saf ampirik SW tekniği
2. Yayınlılığa bağlı SW tekniği
3. Su buharına bağlı SW tekniği
4. Görüş açısına bağlı SW tekniği

5. Yukarıdaki herhangi ikisinin kombinasyonu olan SW tekniği

Saf ampirik Split-Window tekniği en eski olanıdır. Mevcut verilerden saf katsayılar türetilmiştir ki bu katsayılar ampirik katsayılardır. Bunlar da enternasyonal olmadıklarından global geçerlilikleri yoktur.

Su buharına bağlı Split-Window algoritmalarında atmosferik profillerin iyi bilinmesi ve alınan verilerden su buharına bağlı katsayıların elde edilebilmesi için geniş bir veri arşivinin olması gerekmektedir. Ayrıca atmosferik durumların zamanla hızlı bir şekilde değişebileceği de gözönüne alındığında bu tekniğin uygulamadaki zorlukları ortaya çıkmaktadır.

Algılayıcı görüş açısına bağlı olan 4. teknikte de algılayıcı yapısı ve uydu yörüngesinden aynı noktaya tekrar bakma açısından büyük zorluklar vardır. Bütün bunlardan dolayı yayınlırlığa bağlı olan 2. teknik birçok veri ve çalışmada ön plana çıkmaktadır.

Yayınlırlığa bağlı Split-Window algoritmalarında yer alan sabitler spektral yayınlırlığın bir fonksiyonudur. İlk deneme Kerr ve ark. (1992) tarafından yüzeye bağlı yayınlırlık bilgilerinin kullanımına gerek kalmadan yüzey sıcaklığı için değişik sabitler kullanılmıştır. Yerin ve bitki kaplı ortamların değişik yayınlırlık faktörlerinin durumlarının belirlenmesinde kesinlik sağlamak için, sözü edilen iki yer yüzey tipi içinde yüzey sıcaklığı hesaplanmıştır ki bu hesaplamada bir bitki bilgisi, NDVI veya MSAVI (Modified Soil-Adjusted Vegetation Index) korelasyon amacıyla kullanılmıştır.

Sadece değişik dalgaboyları aralığında spektral yayınlırlık değerlerinin bilinmesi durumunda tam ve uygun sabitlerin bulunması mümkün kılınabilir. Bu sabitler Becker ve Li (1990) tarafından türetilmiştir.

Işıma transfer eşitliğinin çözümünde farklı yaklaşımlar kullanarak farklı katsayılarla elde etmek mümkündür, yani yeni bir metodun geliştirilmesi ve a_1 katsayılarının daha önce belirtilen genel Split-window tekniğinin temel formu olan denklem 3.34 ile benzer katsayılar ve ölçüm sonuçları vermesi olağandır.

Yayınırılık değeri büyük alanlarda güçlü ve hızlı bir şekilde bitki örtüsünün değişmesinden dolayı değiştiği için Becker-Li eşitliği bölgesel skalalar için geçerlidir.

Kerr ve Lagouarde (1989) Split-window algoritmalarının yüzey sıcaklığı hesaplamalarında lokal olarak geçerli olduğunu ve yüzey tiplerinin karakteristik özelliklerine mutlaka dikkat edilmesinin zorunluluk olduğunu iddia etmişlerdir. Becker-Li' nin çalışmaları bu iddiaların ilk kez test edildiği durumdur ve buradaki a_i katsayıları yüzey sıcaklığı hesaplamasının yapıldığı zaman dilimi içerisindeki aktif yayınırılık değerlerinden ortaya çıkarılmıştır. Bu metot Fransa' nın güney batısında yer alan bir test bölgesinde, yayınırılık değerlerinin ölçülmesi ve bu değerlerin kullanılması ile birleştirilip test edilmiştir. Sonuçların dikkate değer bir oranda yer verileri ile uyduğu görülmüştür.

Tabi ki bu formülasyonun da diğer formülasyonlar gibi diğer coğrafi bölgelere ve değişik yüzey karakteristiklerinin olduğu yerlerde denenmesi gerekmektedir. Becker-Li çalışmasındaki ana temel a_i sabitlerinin atmosferik etkilerden bağımsız olarak direkt yayınırılığı temel alarak geliştirilmiş olmasıdır. Daha sonraki bazı çalışmalarda ise atmosferik etkiler belirli ölçülerde hesaba katılarak a_i katsayıları hesaplanmıştır. Literatürde değişik yüzey tiplerinin yayınırılık değerlerinin hesaplanması için yapılmış birçok çalışma ve bulunmuş birçok değer olmasına rağmen bu metodun kullanımında yayınırılık değerlerinin çalışma bölgesinden derlenmiş olmasının büyük önemi vardır (Yıldız, 2009).

Yüzey sıcaklığı hesaplamasında birçok algoritma vardır, bunlardan bir kısmı Becker-Li algoritmasına benzerken bir kısmı ise direkt bu algoritmayı baz alarak türetilmiştir. Bazı algoritmaların sabitleri ise yer verilerinden üretilerek bulunmuştur. Bu tip algoritmaların en önemlisi Prata ve Platt (1991)' ın algoritmasıdır. Sobrino (1994) ve Ulivier (1992) algoritmaları ise daha önce türetilmiş algoritmalara bazı eklemeler yapılması ile oluşmuş olanlara iyi birer örnektir.

Becker- Li 1990

$$T_{BELi90} = 1.274 + [(T_4 + T_5)/2 \times [1 + 0.015616 (1 - \varepsilon) / \varepsilon] - 0.482 (\Delta\varepsilon / \varepsilon^2)] + (T_4 - T_5)/2 \times [6.26 + 3.98 (1 - \varepsilon) / \varepsilon + 38.33 (\Delta\varepsilon / \varepsilon^2)] \quad (3.12)$$

Prata – Platt 1991

$$T_{PP91} = 3.45 + (T_4 - T_0) / \varepsilon_4 - 2.45 (T_5 - T_0) / \varepsilon_5 + 40 (1 - \varepsilon_4) / \varepsilon_4 + T_0 \quad (3.13)$$

Ulivieri 1992

$$T_{UL92} = T_4 + 1.8 (T_4 - T_5) + 48 (1 - \varepsilon) - 75 \Delta\varepsilon \quad (3.14)$$

Sobrinho 1994

$$T_{SO94} = T_4 + 2.76 (T_4 - T_5) + 38.6 (1 - \varepsilon) - 96 \Delta\varepsilon \quad (3.15)$$

Bu çalışmada girdi parametresi olarak 3.15 de verilen Sobrinho 1994 ile hesaplanan yer yüzey sıcaklığı değerleri kullanılmıştır.

3.2. Yapay Zeka / Makine Öğrenmesi

Binet'e göre zeka, iyi akıl yürütme, iyi hüküm verme ve kendi kendini aşma kapasitesidir (Wikipedia, 2021). Akademide mantıksal dilsel beceriler ortaya koyan kişiler zeki sayılır. Çoklu zekâ kuramı Howard Gardner tarafından 80'lerde ortaya konduğunda zekâyı tek bir yetenek olarak görmekten öte, çok boyutlu bir model olduğu yaklaşımı yaygınlaşmaya başladı. Zeka artık IQ testlerinin tekeline çıktı. Beyin araştırmalarından esinlenen yapay sinir ağları yaklaşımı ile bilgisayarlar üzerinden bir çeşit zeka üretme ile başlayan yapay zeka kavramı da süreç içinde çok farklı anlamlar kazandı. Yapay sinir ağı kavramı ilk kez McCulloch ve Pitts tarafından 1943'te önerilmiş ve Donald Hebb tarafından 1949'da sinir ağları için Hebbian öğrenimini geliştirilmiştir. 1951'de Marvin Minsky ve Dean Edmonds, ilk sinir ağı bilgisayarı olan SNARC'ı icat etti. Bu başarının ve Turing'in Turing testini geliştirmesinin ardından araştırmacılar sinir ağları ve akıllı sistemlerle ilgili

çalışmalara giderek daha fazla ilgi duymaya başladı. John McCarthy'de 1956'da Dartmouth Üniversitesi'nde düzenlenen bir çalıştayda Yapay Zeka terimini ortaya attı. İnsan zekasını taklit etmek için bir makinenin yapılabileceği varsayımı üzerine kurulmuştur Katılımcılar arasında Minsky, Claude Shannon gibi bilgi teorisinin geliştiricileri de vardı. Amacı karmaşık hesaplar yapabilen, değişken bir ortamda öğrenebilen, tepki ve kararlar verebilen makineler, bilgisayar sistemleri oluşturmak olan yeni bir disiplin olarak ortaya çıktı (Mueller ve Massaron, 2018). Yarım asırda bilişsel bilimler, mantık, bilgi tabanlı sistemler, olasılık, doğadan esinlenen çözüm algoritmaları, bulanık yaklaşımlar, akıllı ajanlar gibi alanlar gelişti. Günümüzde yapay zeka kavramının bir alt kümesi olarak makine öğrenmesi geliştirildi. Makine öğrenmesi, bilgisayarların verilerden öğrenmek, yeni tahmin ve sınıflamalar için programlanması şeklinde tanımlanabilir. Burada algoritma denen, kendine has çözüm yöntemleri önem kazanıyor. Günümüzde öneri sistemleri, müşteriler hakkındaki bilgilerden hareketle, onlara tavsiyelerde bulunabilen, elektronik ticarete şirketlerin sıkça kullandığı bir teknoloji olup makine öğrenmesine dayanmaktadır.

3.2.1. Makine Öğrenmesi Türleri

Makine öğrenmesini gerçekleştirebilecek birçok algoritma vardır. Bu algoritmalar genel olarak üç ana başlıkta toplanır. Bunlar; denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli öğrenme algoritmalarıdır

Makine öğrenmesi algoritması verilen verinin çıktısının olduğu bir ilişkisel veri setini eğitim amaçlı kullanır sonrasında verilen verilerin çıktılarını tahmin etmeye çalışıyorsa bu bir denetimli öğrenme örneğidir (Jiang, 2020).

Örneğin bu tez çalışmasında uydu görüntülerinden ve meteorolojiden elde edilen verileri kullanarak güneş enerjisi tahmini yapılacaktır. Sonuç verisi olarak yer meteoroloji istasyonlarında ölçülen güneş enerjisi verisi eğitim verisetinde bulunur ve uydu verisi, o veriye ait güneş enerjisi ölçüm verisi ile eşleştirilir. Denetimli öğrenme yapılırken makine öğrenmesi algoritması eşleştirilmiş verinin

uygun bir fonksiyonla, ölçülmüş bir çıktı verisine bağlanması üzerinden bu fonksiyonu tahmin eder. Bu fonksiyon katsayı değerlerini test verisi üzerinde uygular gerekirse düzeltme yapar. Veri setinde bulunduğu ağırlık katsayılarından çıkarım yapmaya çalışır ve uygun tahminler ile modeli üretir.

Denetimli öğrenme; modelin ölçülmüş veriler ile eğitildiği, bilinenden bilinmeyene gidilen regresyon benzeri algoritmaların kullanıldığı makine öğrenmesi yöntemidir, literatürde en sık kullanılan yöntem olduğu görülür.

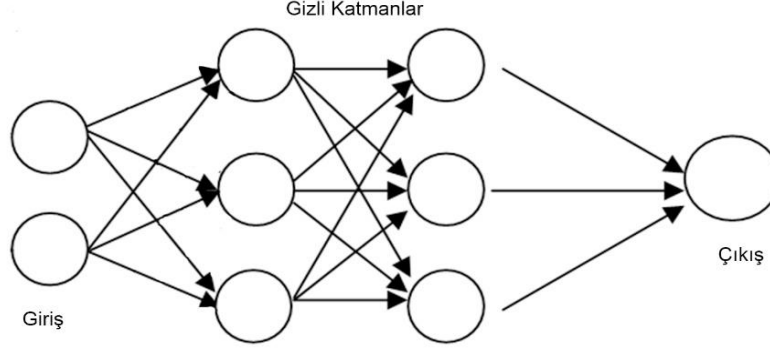
Denetimsiz öğrenmede veriler aranan sonuç verileri ile eşleştirilmemiştir. Algoritma veri içerisinde bağlantıları, örüntüleri istatistiksel ilişkiler kurarak tespit eder ve sonuçta ortaya kurallar belirler. Denetimsiz öğrenme algoritmalarının çoğu sınıflamadan ziyade kümeleme de kullanılırlar (Alpaydın, 2012).

Pekiştirmeli öğrenmede, algoritma deneme yanılma yolu ile hareket ederek en fazla ödül topladığı eylemleri yapma eğilimde olur. Yanlış seçimler ise cezalandırılır. Pekiştirmeli öğrenme, veri etiketlenmediğinden ve uygun eylemlerin dış müdahale ile düzeltilmemesi ile denetimli öğrenmeden ayrılır. Pekiştirmeli öğrenme; içeriği ya da özellikleri tam bilinmeyen bir veri kümesinde keşif ile var olan bilgidan yararlanma arasında bir denge kurma modelidir (Kaelbling, 1996).

3.2.2. Yapay Sinir Ağları / Çok Katmanlı Algılayıcı

Yapay sinir ağları; bir bilgi işleme yöntemi olup, beynin öğrenme şeklini matematiksel olarak modellenmesini sağlayan yöntemler bütünüdür (Joselin, 2018). ANN öğrenme, öğrendiklerini kaydetme, veriden yeni bilgiler üretme ve veriler arasında ilişkiler oluşturabilme yeteneklerine sahiptirler. ANN günümüzde birçok alanda kullanılsa da en yaygın kullanıldığı yerler; görüntü tanıma, veri analizi, kontrol, kümeleme ve sınıflandırma problemlerinin çözümünde kullanılır.

İleri beslemeli yapay sinir ağları (Feedforward Neural Network), en basit ANN olup veri ileri doğru gizli katmanlardan geçip çıkışa aktarılır (LeCun, 2015). En yalın hali ile şekil 3.3'te gösterilmiştir.

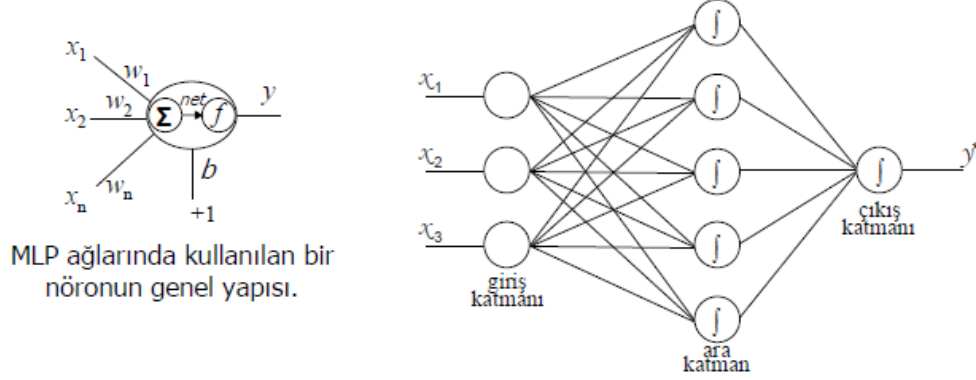


Şekil 3.3. Feedforward Neural Network yapısı

Birçok derin öğrenme uygulaması, örneğin bir görüntüyü sabit boyutlu bir giriş değeri olarak alır, veriyi ileri doğru işleyerek, sabit boyutlu bir çıktıya eşlemeyi öğrenen ileri beslemeli katmanlı sinir ağı mimarilerini kullanır. Bir katmandan diğerine geçmek için, bir dizi birim, önceki katmandan gelen girdilerinin ağırlıklı toplamını hesaplar ve sonucu doğrusal olmayan bir işlevden geçirir.

Doğrusal olmayan fonksiyon, katmanın herhangi bir çıkışından gelen değeri bir aktivasyon fonksiyonu ile bir sonraki çıkışa verir. Problem türüne ya da elde edilmek istenen çıkışa göre aktivasyon fonksiyonu kullanılır.

Literatürde Yapay Sinir Ağları'nın (ANN), Algılayıcı ve Adaline yöntemleri doğrusal olmayan çözümler üretmediği için hem mimari hem de eğitim algoritması açısından iyileştirilmiş Çok Katmanlı Algılayıcı (multi-layer perceptron MLP) ağı önerilmiştir. Mimarisi doğrusal olmayan aktivasyon fonksiyonuna dayanır, birçok nöronun birbirine hiyerarşik şekilde bağlandığı MLP, geri-yayılım adındaki öğrenme yöntemini kullanmaktadır. Shiruru,(2016).



Şekil 3.4. Çok Katmanlı Algılayıcı Nöron Yapı ve Bağlantıları (Orhan,2021)

Aktivasyon Fonksiyonu MLP'nin temel özelliğidir. Aktivasyon fonksiyonu, bir değişkenin değerini doğrusal yada doğrusal olmayan bir ilişki yaratan fonksiyondur. Türevinin kolay alınabilir olması eğitilme hızını artırır. Genellikle kullanılan üç aktivasyon fonksiyonu vardır: Sigmoid, Hiperbolik Tanjant ve Adım-Basamak fonksiyonlarıdır.

Sigmoid

$$y = \frac{1}{1 + e^{-net}}$$

Hiperbolik tanjant

$$y = \frac{1 - e^{-2net}}{1 + e^{-2net}}$$

Adım basamak

$$y = \begin{cases} 1 & net \geq 0 \\ -1 & net < 0 \end{cases} \quad (3.16)$$

Geri-yayılım (Backpropagation) yönteminde bütün ağırlık değerleri, eğim azaltma (gradient descent) denilen bir hata fonksiyonunu minimize etmesiyle bulunur. Hesaplanan hata değeri, nöronun çıkışından girişine doğru bir geri beslemeye maruz kalarak aktivasyon fonksiyonunun türevinden etkilenir. Fonksiyonun türevine uygulanan hata değeri tüm ağırlıklara, giriş değerleri oranında dağıtılır. (Jiang, 2020).

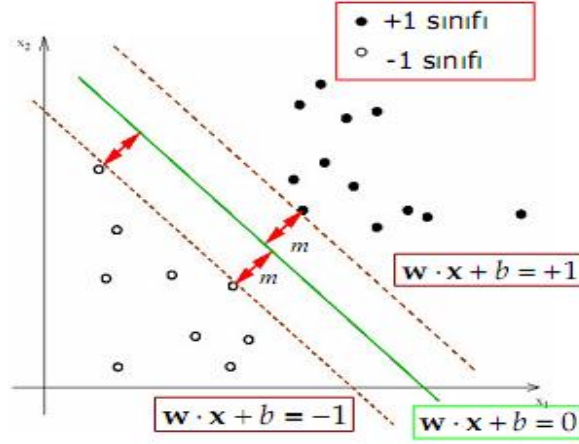
$$\begin{aligned}
\Delta w_{ij} &= -\eta \frac{\partial E}{\partial w_{ij}} = -\eta \left[\frac{\partial E}{\partial y_j} \right] \left[\frac{\partial y_j}{\partial net_i} \right] \left[\frac{\partial net_i}{\partial w_{ij}} \right] \\
\Delta w_{ij}(t+1) &= -\eta \frac{\partial E}{\partial w_{ij}} + \alpha \Delta w_{ij}(t) \\
&= -\eta \left[-e_j \right] \left[\frac{\partial y_j}{\partial net_i} \right] [x_i]
\end{aligned} \tag{3.17}$$

Eğim azaltma ile algoritma kademeli olarak sonuca ulaşırken bazen yerel minimum değerleri yüzünden adım miktarı sıfırlanır ve sonuca ulaşamayabilir. Bu yüzden eğim azaltma yönteminde ilk adımlarda değişim değerlerini belli bir (α) katsayısı ile momentum etkisi kullanılır.

MLP ağında eğitiminin durması gereken zamanı belirleyecek bir kriter olmazsa aşırı eğitim (overfittig) oluşur. Çapraz geçerlilik gereği verinin altkümesi olan eğitim kümesi devamlı ikiye bölünür. Bu parçalardan biri eğitimde ağırlıkların güncellenmesi için kullanılır, diğeri ise (validation set) üzerinden eğitim başarısı hesaplanır ve başarımda düşüş gözlemlendiğinde eğitim durdurulmalıdır.

3.2.3. Destek Vektör Makinesi / Regresyonu

Destek Vektör Makinesi (Support Vector Machine SVM) yöntemi, Vapnik ve arkadaşları tarafından geliştirilmiş etkin bir makine öğrenmesi yöntemidir (Cortes ve Vapnik, 1995). Bu yaklaşım, yapısal risk minimizasyonu prensibine dayanır. SVM veriyi sınıflandırırken sınıfların birbirlerine en yakın durumlarını bularak bu örneklerin ayırıcı yüzeye dik uzaklıklarını maksimize etmeyi amaçlar. Ayırıcı yüzeyin, veri kümesi üzerindeki başarısı değişmeden birçok farklı alternatifi olabilir. SVM sayesinde ayırıcı yüzey her iki sınıfa da aynı mesafede ve maksimum uzaklıktadır. SVM ayırıcı yüzeyi ise her iki sınıfa da aynı maksimum uzaklıkta bulunmaktadır.



Şekil 3.5. Destek Vektör Makinesi Vektör Sınıflaması (Orhan,2021)

Doğrusal olmayan bir veri kümesinde uygulanan çekirdek (kernel) fonksiyonu algoritmanın farklı boyutta, çeşitli düzlemlerde ikili doğrusal SVM çözümleri uygulayabilmesini sağlar. Sık kullanılan fonksiyonlardan bir kaç, doğrusal (linear), polinom (polynomial) ve RBF (radial basis function) fonksiyonlarıdır. Ayhan, Erdoğan, (2014).

$$L(\alpha) = \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j d_i d_j K(x_i, x_j) \quad (3.18)$$

SVM, eğitim verisiyle modeli eğittikten sonra girdinin farklı kümelere ikili sınıflandırılması için kullanılan denetimli bir makine öğrenmesi tekniğidir. Farklı değişkenlerle, farklı düzlemlerde çalışabilir. SVM, özellikle iki farklı veri kümenin sınıflandırılmasında kullanılmıştır. SVM sınıflandırma için kullanılan bir model oluşturarak çalışır. Çekirdek fonksiyonu yöntemine dayanır. Boyutlandırma veya kümelemeyi azaltmak için sınıflandırma/regresyon kullanılır Boser ve ark. (1992). En küçük kareler gibi hatayı minimuma indirmek için basit doğrusal regresyon ile, girdi verileri için bir fonksiyon modeli seçer, tüm veri seti için toplam karesel hatayı minimuma indiren bir şekilde modelin ağırlıklarını değiştirir. Hedef iki küme arasındaki ortalama uzaklığı, her iki kümenin en yakın veri noktasından

maksimum uzaklığı olan bir düzlem tanımlamaya çalışır. Verilen eşitlikle tüm eğitim verileri için bir Lagrange çarpanı hesaplanır ve $\alpha_i > 0$ olan x_i örnekleri destek vektörlerini tanımlamaktadır. Bu destek vektörler ile aşırı düzlemler hesaplanır. SVM algoritması bulduğu bu değerlere göre optimum aşırı düzlemi tespit eder (Vapnik ve ark.1996). SVR sınıfları değil olası nümerik değerleri tahmin etmede özelleşmiş bir SVM'dir.

3.3. Klasik Regresyon Analizi

Regresyon analizi, iki ya da daha çok değişken arasındaki doğrusal yada doğrusal olmayan ilişkiyi ölçmek için kullanılan bir metottur. Eğer tek bir değişken kullanılarak analiz yapılıyorsa buna tek değişkenli regresyon, birden çok değişken kullanılıyorsa çok değişkenli regresyon analizi olarak isimlendirilir. Regresyonda, değişkenlerden biri bağımlı diğerleri bağımsız değişken olmalıdır. Gujarati, (2008).

$$y_i = \beta_0 + \sum_{j=1}^p X_{ij}\beta_j + \varepsilon_i \quad (3.19)$$

Bir regresyon denkleminin kestirmede kullanılacak yöntem adi en küçük kareler (ordinary least squares) olarak adlandırılır. Bu yöntemde göre parametre kestirimleri artıkların kare toplamının minimum (en küçük) değerini bularak elde edilir. iki parametrelili doğrusal regresyon için en küçük kareler parametre tahminleri için formüller şöyledir, Montgomery ve Hines, (1990):

$$\hat{\beta}_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (3.20)$$

Doğrusal regresyon yöntemini kullanmak için temelde şu varsayımların bulunduğu kabul edilmektedir:

Çıkarımsal yöntem olduğu için kullanılan iki değişkenli örneklemin ya istatistiksel rastgele örneklem olduğu ya da anakütleyi çok iyi temsil ettiği bilinmektedir. Bağımlı değişken içinde hata bulunmaktadır. Bu hatanın bir rassal değişken olduğu ve ortalama hatanın sıfır olduğudur. Sistematiik hata da bulunması mümkündür ama bu hatanın incelemeye alınması regresyon analizi kapsamı dışındadır.

Bağımsız değişken hatasızdır. Eğer bağımsız değişkende hata bulunduğu varsayılırsa özel bir yöntem şekli olan değişkenler-içinde-hata modeli teknikler kullanılarak model kurulmalıdır.

Hatalar zaman içinde ve kendi aralarında birbirine bağımlı değildir. Buna otokorelasyon veya serisel korelasyon bulunmaması varsayımı adı verilir. Hata varyansı sabittir ve veriler arasında hiç değişmediği varsayılır. Bu eşvaryanslılık veya homoskedastisite varsayımı adı ile anılır. Eğer bu varsayım uygun değilse ağırlıklı en küçük kareler yöntemi uygulanabilir. Hataların varyans-kovaryans matrisinin çapraz elamanları sabit hata varyansı olur ve matrisin diğer çapraz dışı elemanları 0 olur.

Eğer çoklu regresyon analizi yapılıyor ve uc veya daha çok parametre için kestirim isteniyorsa, bağımsız değişkenlerin birbirleri ile bağlantısının olmaması gereklidir. Buna çoklu doğrusallık (multicollinearity) olmaması varsayımı adı verilir. Hatalar bir normal dağılım gösterir. Eğer bu hataların normalliği varsayımı uygun değilse genelleştirilmiş doğrusal model uygulanabilir (wikipedia, 2021).

3.4. Değerlendirme Kriterleri

Bu çalışmada tahmin edilen güneş enerjisi değerlerinin tahmin doğruluğu için Hata Karelerinin Ortalamasının Kare Kökü (RMSE) ile Korelasyon Katsayısı (r) ölçütleri baz alınmıştır. Weka işlemler sonunda beş kritere ait değerleri rapor etmektedir. Aşağıdaki tabloda verilen bu hata kriterleri görülmektedir.

Çok Katmanlı Algılayıcı, Destek Vektör Regresyonu ve Regresyon Analizi algoritmalarının ortaya koyduğu model sonuçları ile ölçüm sonuçlarının

karşılaştırılmasında, değerlendirme kriterlerinden r ve RMSE'nin kullanılma sebebi regresyon tabanlı tahmin problemlerinde daha iyi sonuçlar ortaya koyduğu bilinmesidir, Montgomery ve ark., (2012).

Çizelge 3.4. Weka Sonuçlarında Verilen Değerlendirme Kriterleri

Correlation coefficient	$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$	Korelasyon Sabiti
Mean absolute error	$MAE = \frac{1}{N} \sum_{i=1}^N \hat{\theta}_i - \theta_i $	Ortalama Mutlak Hata
Root mean squared error	$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{\theta}_i - \theta_i)^2}$	Hata Karelerinin Ortalamasının Kare Kökü
Relative absolute error	$RAE = \frac{\sum_{i=1}^N \hat{\theta}_i - \theta_i }{\sum_{i=1}^N \bar{\theta} - \theta_i }$	Bağıl Mutlak Hata
Root relative squared error	$RRSE = \sqrt{\frac{\sum_{i=1}^N (\hat{\theta}_i - \theta_i)^2}{\sum_{i=1}^N (\bar{\theta} - \theta_i)^2}}$	Bağıl Kök Hatası

Regresyon tabanlı tahmin problemlerinde r korelasyon değeri ile RMSE kök ortalama kare hatası modeli değerlendirmek açısından en önemli kriterdir. r değeri 0-1 arasında, 0.70 üstü olması tahmin sonuçları ile gerçek sonuçlar arasında güçlü bir ilişki olduğunu ortaya koyar. RMSE sıfırdan, sonsuza gidebilen bir

değerdir küçük olması tercih edilir.

Korelasyon Sabiti (r): Korelasyon aralarında açıklayıcı ilişki bulunan bağımsız ve bağımlı değişken arasındaki ilişkinin gücünü ve yönünü açıklar. Yön pozitif yada negatif olabilir İki değişken arasındaki ilişkiyi ortaya koyan katsayıya ‘basit korelasyon katsayısı’ denir.

Ortalama Mutlak Hata (MAE): İstatistikte aynı olguyu ifade eden eşleştirilmiş gözlemler arasındaki hataların bir ölçüsüdür. Gerçek verilerle aynı birime sahiptir ve hataları aynı birimlerde ölçülen modelleri karşılaştırılabilir. Genellikle büyüklük olarak RMSE'ye benzer, ancak biraz daha küçük sonuçlar verir.

Hata Karelerinin Ortalamasının Kare Kökü (RMSE): Bir modelin uygunluğunun kalitesini ölçmenin yaygın bir yoludur. Öngörülen sonuçlar gerçek sonuçlara çok yakınsa, RMSE küçük olacaktır. Tahmin edilen ve gerçek yanıtlar önemli ölçüde farklılık gösteriyorsa, RMSE büyük olacaktır. Hata karelerinin toplamının ortalamasıdır.

Bağıl Mutlak Hata (RAE): Bir modelin Ortalama hatasını, tahmin modeli tarafından üretilen hatalarla karşılaştıran bir oran olarak ifade edilir. İyi bir tahmin modeli sıfıra yakın bir oran üretecektir. Modelin toplam mutlak hatasını, diğerinin toplam mutlak hatasına bölerek normalleştirir.

Bağıl Kök Hatası (RRSE): Basit bir tahmin edicidir. Tahmin hata değerlerinin karesinin toplamını, gerçek hata değerlerinin karesinin toplamına bölerek normalleştirir. Hataları farklı birimlerde ölçülen modeller arasında karşılaştıma yapılabilir (Hiregoudar, 2020).

3.5. Uygulama Yazılımı (WEKA)

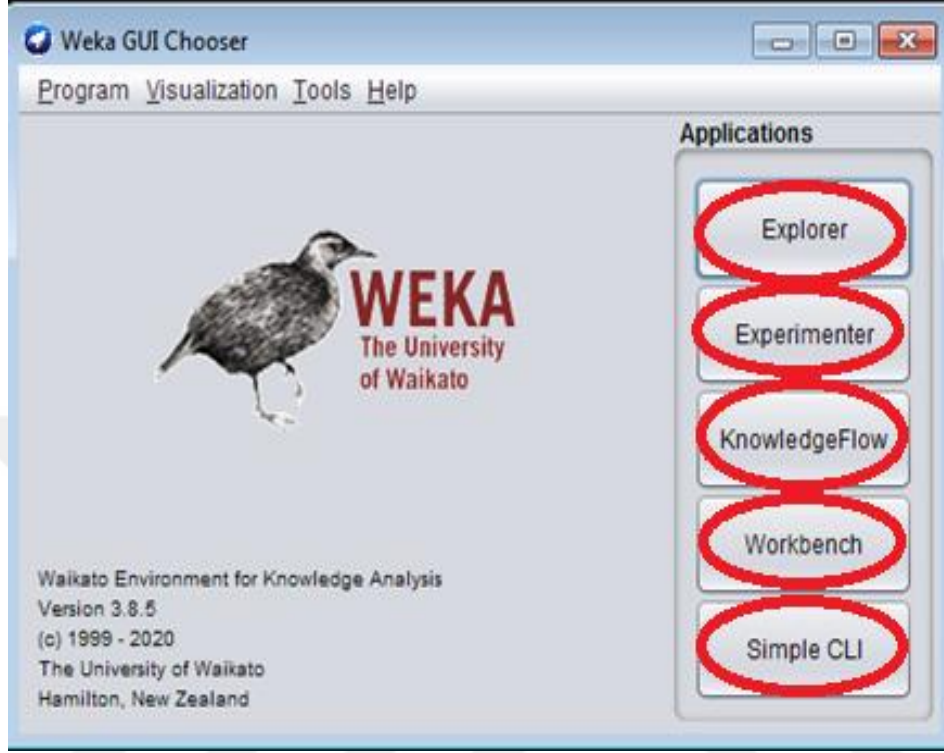
WEKA veri madenciliği ve analizinde kullanılan ücretsiz bir yazılımdır. Yeni Zellanda’da 1993’ten itibaren Waikato Üniversitesi tarafından geliştirilmeye başlanan, "Waikato Environment for Knowledge Analysis" sözcüklerinin baş harflerinden oluşan adı, Yeni Zellanda’da yaşayan bir kuş türünün de adıdır ve logo resmi olarak kullanılmaktadır. Java dilinde geliştirilmiş olup kod yazılarak ta

kullanılabilmektedir. Yazılım ara yüzünden kullanımının kolay olması kodlamaya gerek duymadan profesyonel makine öğrenme yöntemlerinin kullanılabilmesi yaygınlaşmasında etkili olmuştur. Birçok makine öğrenmesi algoritmasını içermekte, bir kısmı da harici olarak yüklenebilmektedir (Şeker, 2013)

Resmî sitesi <http://www.cs.waikato.ac.nz/~ml/weka/> olup, indirilerek Linux, Mac OS ve Windows işletim sistemlerine kurulumu yapılabilmektedir. Ücretsiz kamu lisansı ile dağıtılmaktadır. Çalıştırıldığında aşağıdaki yazılım ara yüzü ile karşılaşılır. Bu çalışmada yazılımın 3.8.5 versiyonu kullanılmıştır.

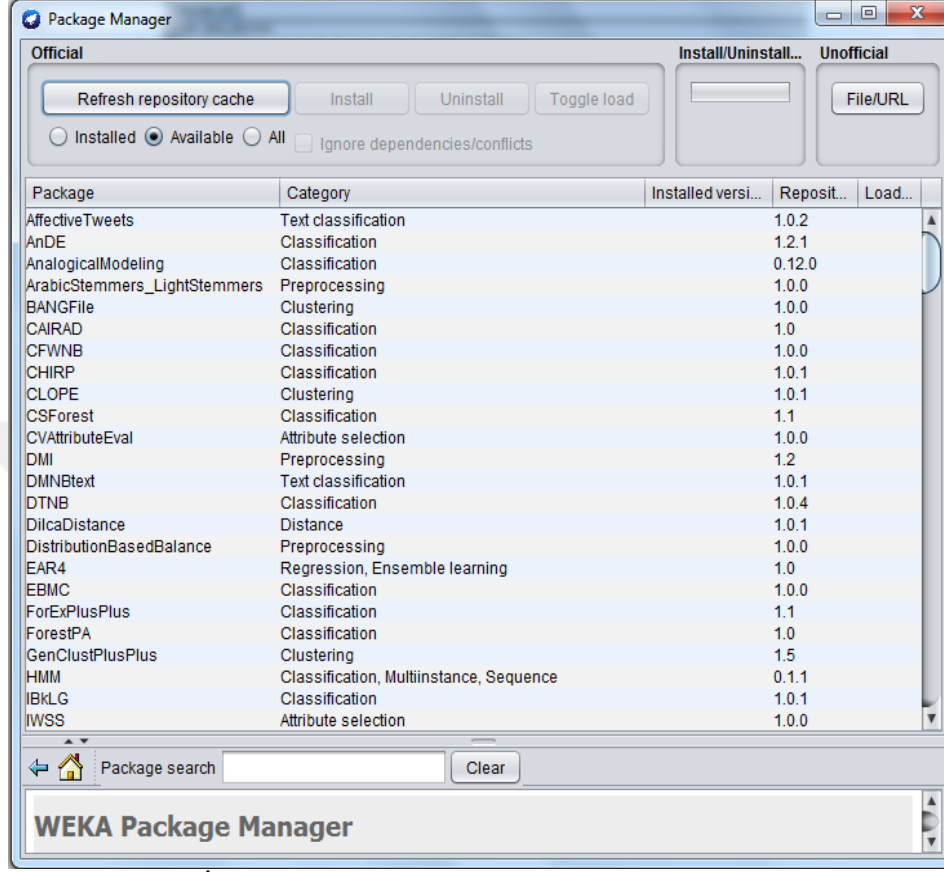
3.5.1. WEKA Arayüzleri ve İşlevleri

Yatay menüdeki Program sekmesinde yapılan işlemlerin kayıtları görülebilir. Visualization sekmesinde veri setindeki veriler X-Y düzleminde görselleştirilebilir. ArffViewer ve SqlViewer ile dosya yada veritabanından çekilebilen veri setleri düzenlenebilmektedir. Help sekmesinde yazılım ile ilgili dokümantasyonlara ulaşılabilir. Aydemir, (2018).



Şekil 3.6. WEKA Yazılım Ara Yüzü

Tools menüsünde, Paket yöneticisi (Package Manager) ile yeni kütüphaneler kurulabilir. Weka’da bulunan, hem kendine ait hem de üçüncü parti algoritmaların bulunduğu-kurulduğu birimdir. Kullanılan tüm araçlar buradan kontrol edilebilir.



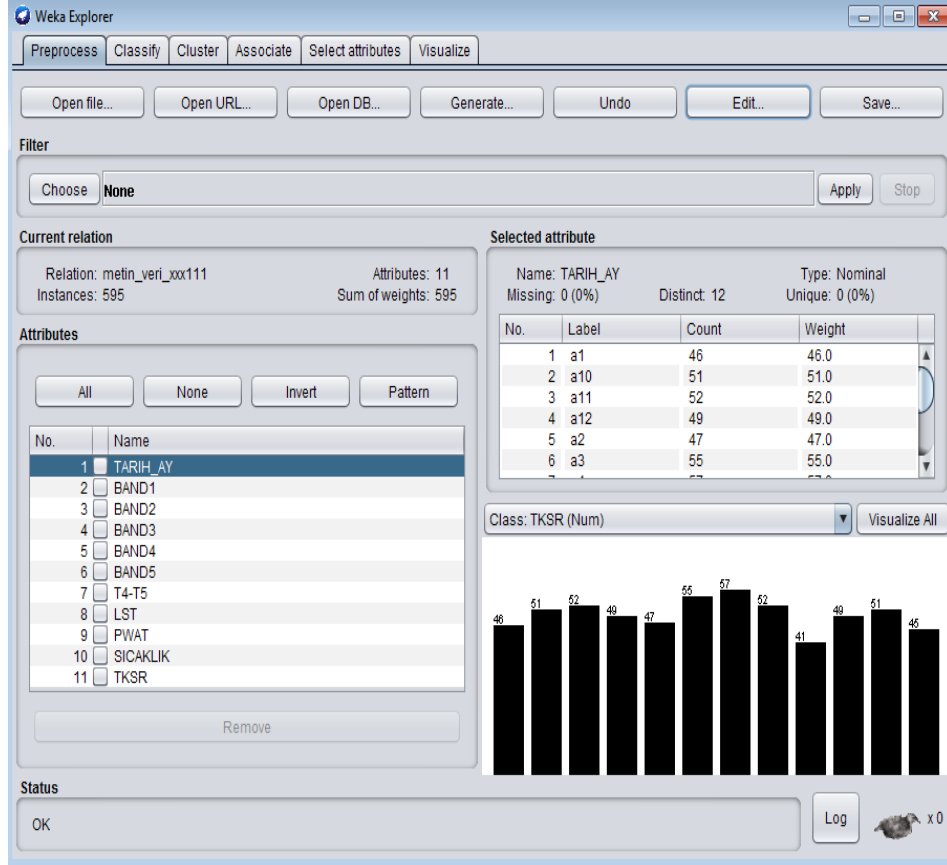
Şekil 3.7. WEKA İçeriğine Erişim Ara Yüzü, Paket yöneticisi

Ara yüzde Applications başlığı altında 5 düğme yer almaktadır. Her biri ayrı görevleri olan ara yüzlere erişim sağlamaktadır. Bunlardan en çok kullanılan Explorer ara yüzüdür.

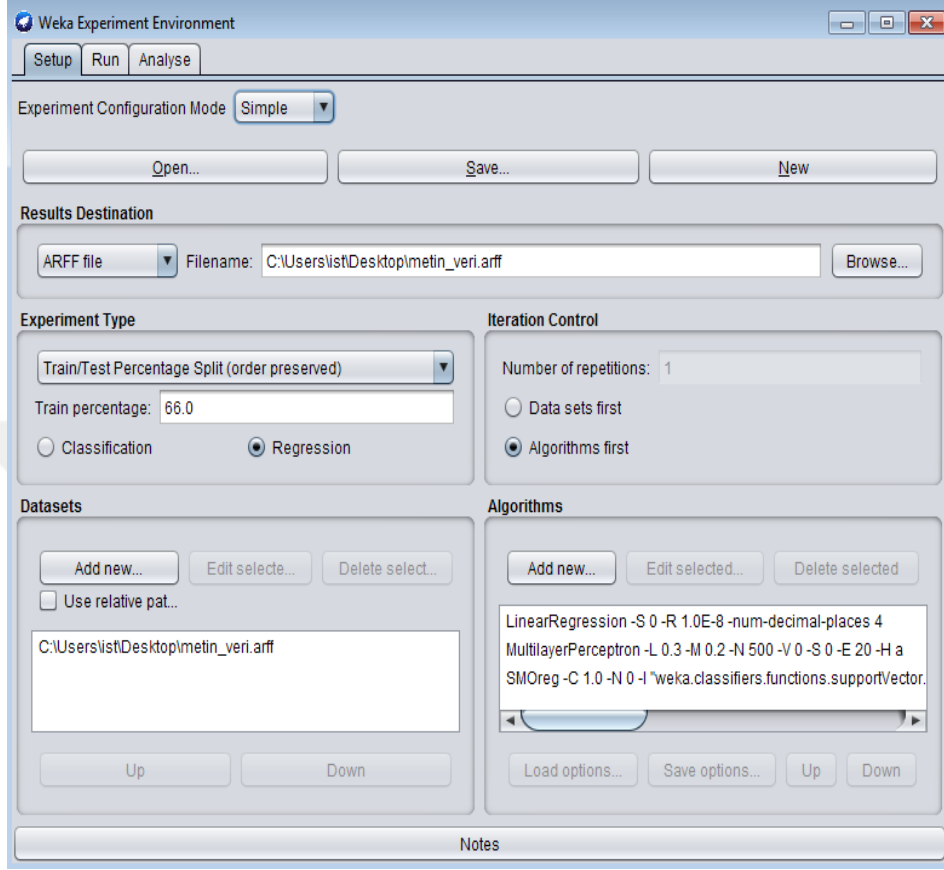
Explorer ara yüzü; Veri ön işleme (Pre-Process), sınıflandırma (Classify), kümeleme / bölütleme (Cluster), ilişkilendirme (Associate), özellik / filtre seçim (Select attributes) ve görselleştirme (Visualize) işlemlerine erişim sağlamaktadır.

Veri ön işleme veri giriş ve ayarlarının yapıldığı kısımdır. En çok kullanılan arayüzdür. Sınıflandırma sekmesi bütün algoritma ve filtrelerin seçilip ayarlandığı sekmedir. Denetimli öğrenme algoritmalarının yer aldığı sınıflama, regresyon tabanlı işlemlerin ayarlandığı kısımdır. Kümeleme sekmesinde denetimsiz

öğrenme algoritmalarının uygulamaları yapılabilmektedir. İlişkilendirme ve özellik seçim sekmelerinde veri alanlarının bazı filtre işlemlerinin yapıldığı, veriler arası ilişkilerin tespit edildiği kısımdır.

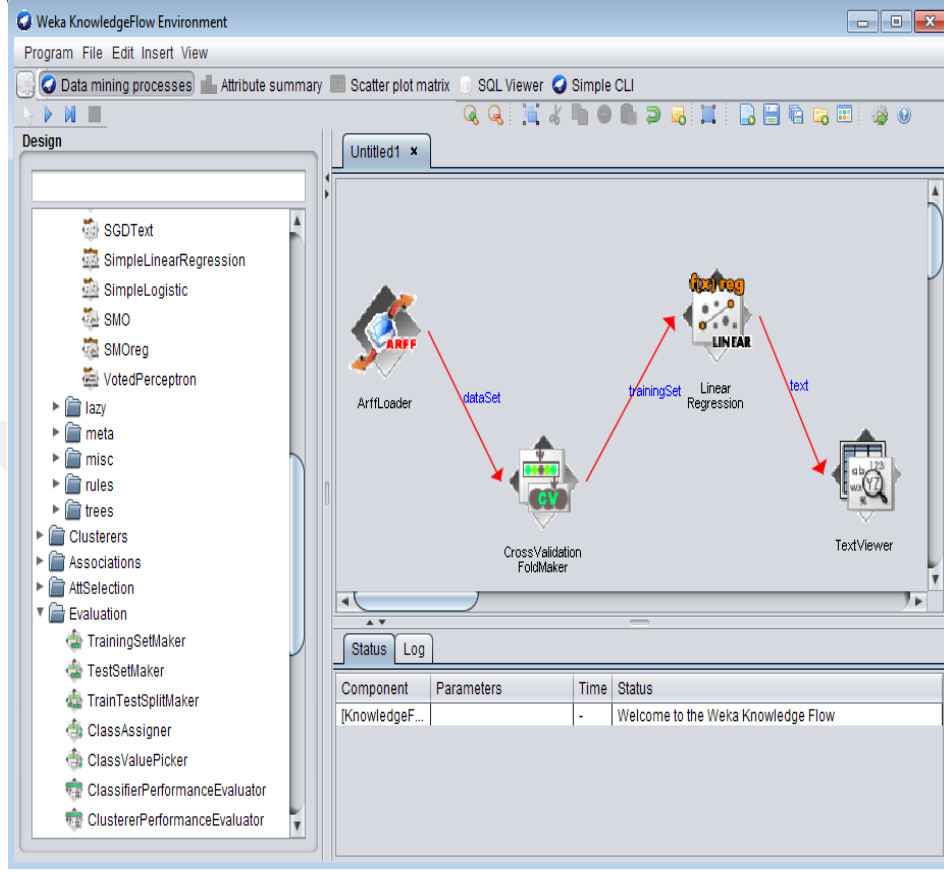


Şekil 3.8. Explorer Ara Yüzü



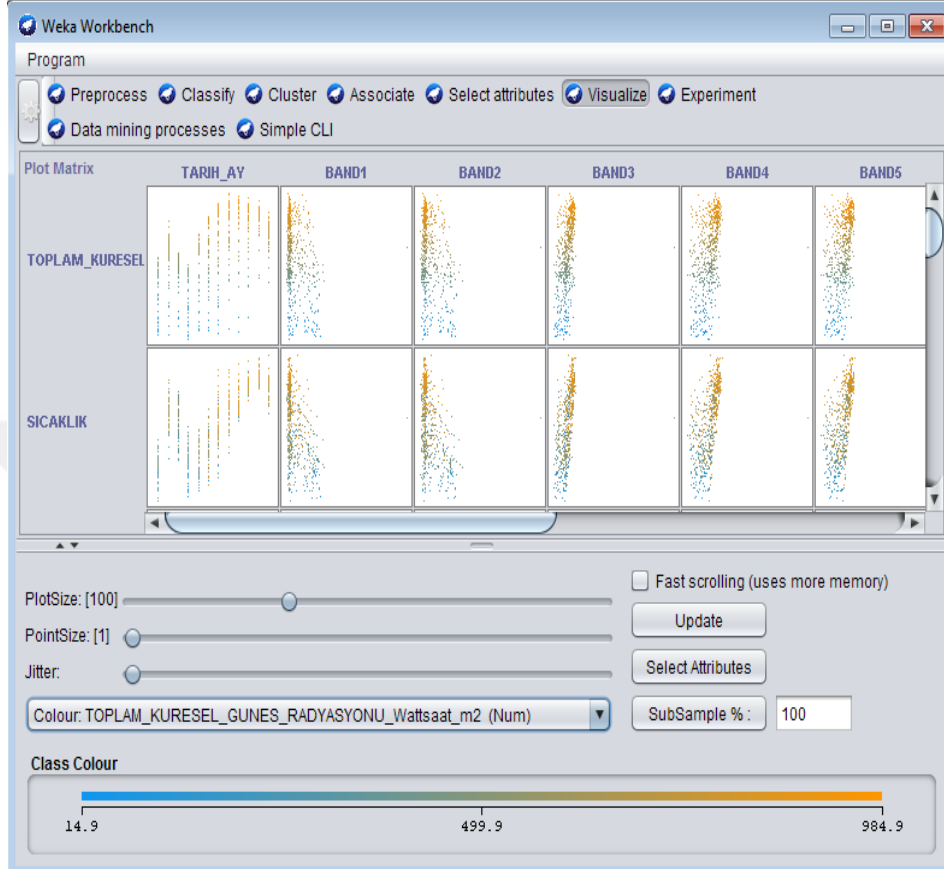
Şekil 3.9. Experimenter Ara Yüzü

Experimenter ara yüzü; Explorer ile yapılabilen işlemlerin, karşılaştırmalı deneylerin yapılmasına imkan sağlamaktadır. Örneğin farklı algoritmaların aynı yada farklı veri setlerinde nasıl sonuç vereceği gözlemlenmektedir. Ayüzü 3 farklı sekmeye ayrılmıştır, Setup: Deneylerin tasarlanması için veri ve algoritma seçimi yapılan sekmedir. Run: Deneylerin çalıştırılması ve hata bilgilendirmesi için kullanılan sekmedir. Analyse: Deneylerin sonuçlarını analiz etmeye yarayan sekmedir. Sonuçlar kayıt edilebilir, özünde istatistiksel t testini algoritmaların veri setindeki başarımlarını karşılaştırır.



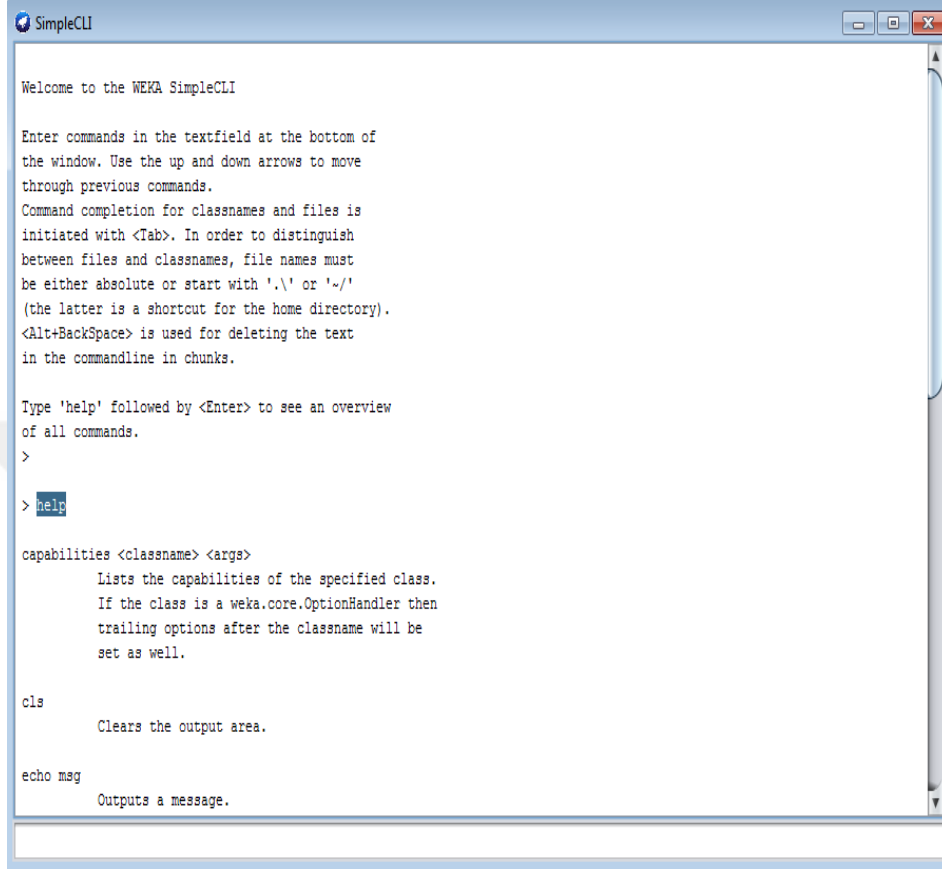
Şekil 3.10. Knowledge Flow Ara Yüzü

Knowledge Flow ara yüzü; Explorer ile yapılabilen işlemlerin görsel nesneler halinde iş süreçlerinin bağlantı mantığına uygun biçimde ilişkilendirilmesi prensibine göre çalışır. Arka planda çeşitli kodların yapacağı işlemleri, somut nesneler arası ilişkilerin görselleştirilmiş olarak yapılabildiği kısımdır.



Şekil 3.11. Workbench Ara Yüzü

Workbench ara yüzü; Diğer tüm ara yüzleri tek çatı altında birleştiren, birlikte kullanımına imkan sağlayan ara yüzüdür. Profesyonel işlemlerin tek birimden yapılmasını sağlar. Verilerin betimsel istatistik yöntemleri ile normalize, standardize edilmesinden, görselleştirilip sunulmasına kadar içinde birçok araç ihtiva eder.



Şekil 3.12. Simple CLI Ara Yüzü

Simple CLI ara yüzü; Komut satırından kod yazarak işlem yapan bir ara yüzüdür. Help komutu ile diğer komutlar hakkında bilgi edinilebilir. Weka geliştirilmesinde Java programlama dili kullanılmış bir yazılım olup, görsel arabirimler ile yapılan tüm işlemlerin profesyonel kullanıcılar için kod yazılarak ta yapılmasına imkan sağlar. Java Api yazılım nesnesi sayesinde kendi programlarında Weka programının olanaklarını yazılımcılar kullanabilmektedir. Son yıllarda R istatistik yazılımının da birlikte kullanımını sağlayan geliştirmelere gidilmiştir. Veri analizi ve madenciliği için kullanışlı bir işbirliği ortaya konmuştur.

3.5.2. WEKA Algoritmaları

Weka'da beklenti maksimizasyonu, kendini organize eden harita algoritmaları vb kümeleme algoritmaları dışında sınıflandırma için de birçok algoritma mevcuttur. Sınıflandırma, topluluk ve regresyon tabanlı algoritmalar bayes, functions, lazy, meta, misc, mlr, rules ve trees başlıkları altında toplanmıştır. Weka'da çalışabilen bazı yaygın kullanımı olan algoritmalar tanıtılacaktır.

Lojistik regresyon: Lojistik regresyon, ikili bir sınıflandırma algoritmasıdır. Giriş değişkenlerinin sayısal olduğunu varsayar ve bir Gauss (çan eğrisi) dağılımına sahiptir. Burada sonuç doğrusal olmak zorunda değildir, regresyon fonksiyonu ve lojistik fonksiyon kullanılarak dönüştürülür. İkili sınıflandırma problemlerini destekler.

Naif Bayes: Bayes bir sınıflandırma algoritmasıdır. Geleneksel olarak, giriş değerlerinin olduğunu varsayar, nominal girdiler bir dağılıma uygun olduğu varsayılarak desteklenir. Bayes Teoreminin basit bir uygulamasını kullanır. Her sınıf, eğitim verilerinden hesaplanır ve birbirinden bağımsız olduğu varsayılır. Teknik olarak şartlı bağımsızlık olarak adlandırılır. Naive Bayes'in çok etkili bir sınıflandırma algoritması olduğu gösterilmiştir. Naif Bayes, her sınıf için sonsal olasılığı hesaplar bu nedenle, hem ikili sınıflamayı hem de çoklu sınıflamayı destekler.

Karar Ağacı: Karar ağaçları, sınıflandırma ve regresyon problemlerini destekleyebilen bu algoritma Sınıflandırma ve Regresyon Ağaçları (CART) diye de adlandırılır. Veriyi ifade etmek için ters çevrilmiş ağaç yapısı kullanır. Kökten başlayarak dallar tahmin yapılana kadar yapraklara doğru hareket eder. Karar ağacı, tahmin yapmak için içinde açgözlü greedy algoritmasını çalıştırır.

K-En Yakın Komşu: Bu algoritma hem sınıflandırmayı hem de regresyonu destekler. KNN diye bilinir. Tüm eğitim veri setini depolayarak ve k'yi bulmak için sorgularla çalışır. Bir tahmin yaparken benzer eğitim kalıplarını kullanır. Bu nedenle, başka bir model yoktur. Veri örnekleri arasındaki mesafe tahminlerde anlamlıdır, genellikle çok iyi performans gösterir.

Torbalama: Bootstrap, istatistiksel bir tahmin tekniğidir. Sınırlı miktarda veri için yararlı bir tekniktir. Bu örnekleme yaklaşımı, makine öğrenmesi modelleriyle kullanılabilir. Eğitim verilerinin bir kısmı değiştirilerek birden fazla farklı makine öğrenmesi modelleri geliştirilir. Her model bir tahmin yapmak için kullanılır ve sonuçların ortalaması alınır. Torbalama ile en çok kullanılan algoritma yüksek varyansa gereksinim duyan karar ağaçlarıdır.

Rastgele Orman: Sınıflandırma için kullanılabilecek karar ağaçları için kullanılan torbalamanın bir uzantısıdır. Torbalı karar ağaçları algoritmasının dezavantajı karar ağaçlarının, ağaç oluşturma sürecindeki her adımda en iyi bölünme noktasını seçen açgözlü bir algoritma kullanmasıdır. Bu, ortaya çıkan ağaç yapılarının benzer ve tahminlerin varyansını azaltıyor olmasını getirmektedir. Rastgele Orman, açgözlü mekanizmayı bozan torbalanmış karar ağaçlarında farklı bir yöntem kullanır. Ağaç oluşturma sırasında bölme noktaları rastgele şekilde seçilebilir.

AdaBoost: Sınıflandırma problemleri için bir topluluk makine öğrenimi algoritmasıdır. Üretilen modellerde, önceki modeller tarafından yapılan tahmin hatalarını düzeltmeye çalışan bir mekanizma kullanır. Her biri tek bir karar noktasına sahip kısa karar ağacı modelleri kullanır. Eğitim veri kümesindeki her örnek ağırlıklandırılır ve ağırlıklar modelde doğru veya yanlış sınıflandırılıp sonraki modelleri eğitir. Daha fazla iyileştirme mümkün olmayana kadar ağırlıklar yeni modele eklenir (Brownlee, 2016).

4. BULGULAR ve TARTIŞMA

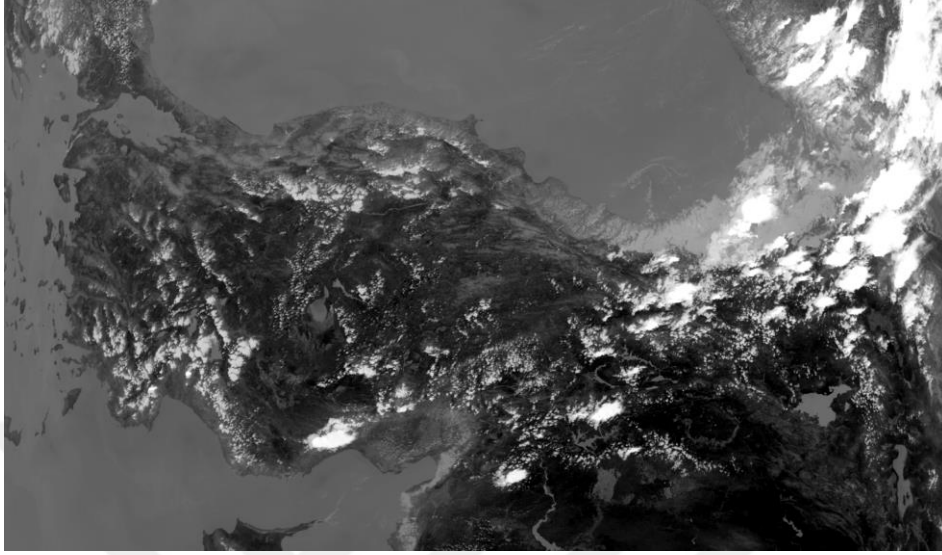
4.1. Uydu Görüntü İşlemleri

Bu çalışmada makine öğrenmesi modellerine girdi parametresi olarak NOAA uydusu üzerinde bulunan 5 bantlı bir algılayıcı olan AVHRR detektöründen elde edilen veriler kullanılmıştır.

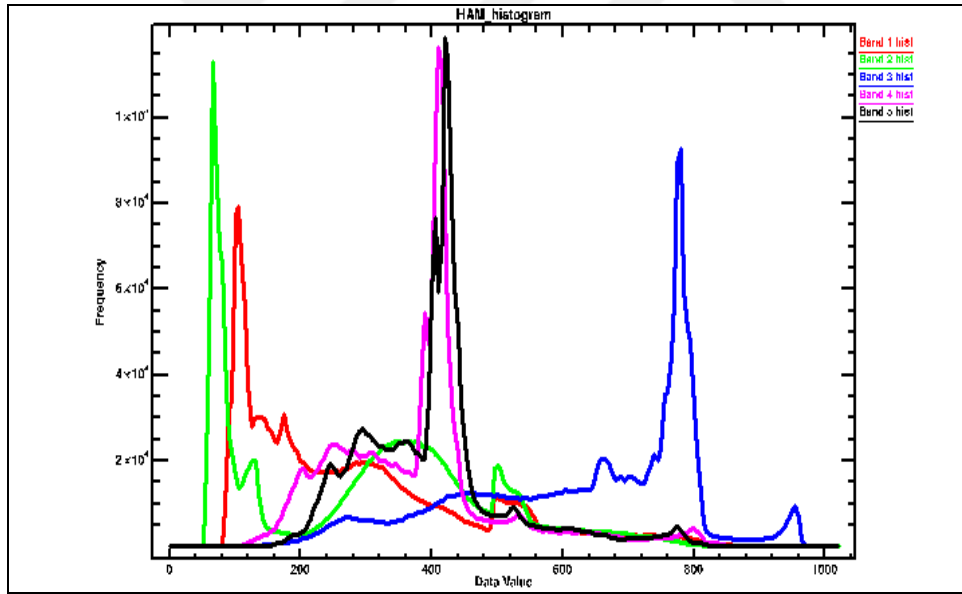
NOAA'nın resmi sitesinden (<https://www.ncdc.noaa.gov/cdo-web/>) indirilen verilerde piksel değerleri uzaktan algılamada Digital Numbers (DN) olarak adlandırılan hiç işlem görmemiş ham veri formatındadır. Bu DN formatındaki veriler post calibration diye tanımlanan ön işlemler uygulanarak anlamlandırılmalı ve herhangi bir uzaktan algılama çalışmasında kullanılabilir hale getirilmelidir. Literatürde Level1b (L1b) diye adlandırılan bu verilere radyometrik kalibrasyon ve coğrafik kalibrasyon adlı iki aşamalı kalibrasyon uygulanmalıdır.

Kalibrasyon işlemleri uzaktan algılama ve veri işleme yazılımı olan ENVI ve TerrSet-İdrisi programı kullanılarak gerçekleştirilmiştir. Kalibrasyon işlemleri ile DN değerindeki piksel değerleri sırası 3.17-3.18 ve 3.19 eşitliklerinin kullanımı ile ENVI yazılımında AVHRR 1.bant ve AVHRR 2.bant için albedo (yansıma), AVHRR 3.bant için yakın kızılötesi bölgesi sıcaklık, AVHRR 4.bant ve AVHRR 5.bant için kızılötesi bölge sıcaklık değerlerine dönüştürülmüştür. Ayrıca coğrafik düzeltme kalibrasyon işleme ile her pikselin WGS-84 projeksiyon sistemine göre enlem ve boylam cinsinden konum değerleri oluşturulmuştur.

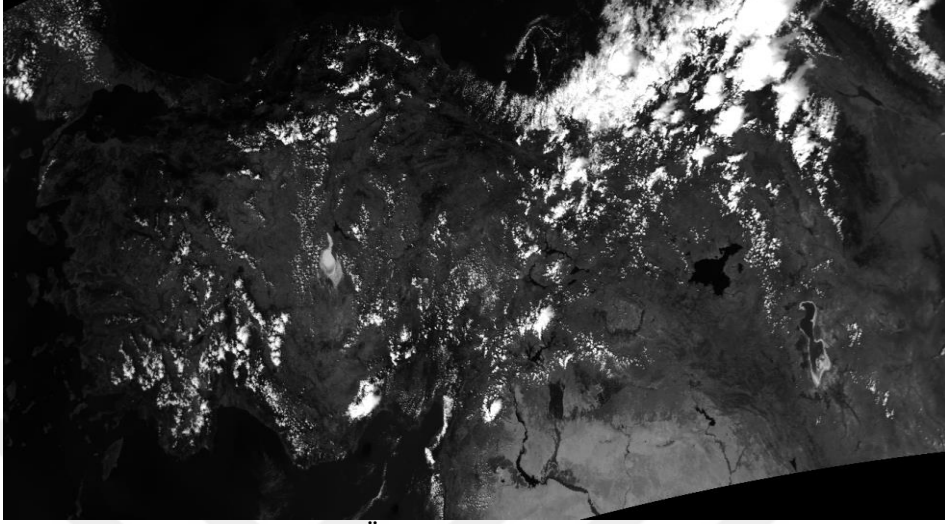
Level1b formatında indirilen ve görüntü işleme programları kullanılarak kalibrasyonu ve coğrafik düzeltmesi yapıp ilgili çalışma bölgesi kesilmiş AVHRR verisinin bant görüntüleri ve bu bantların histogramları sırası ile aşağıdaki şekillerde yer almaktadır.



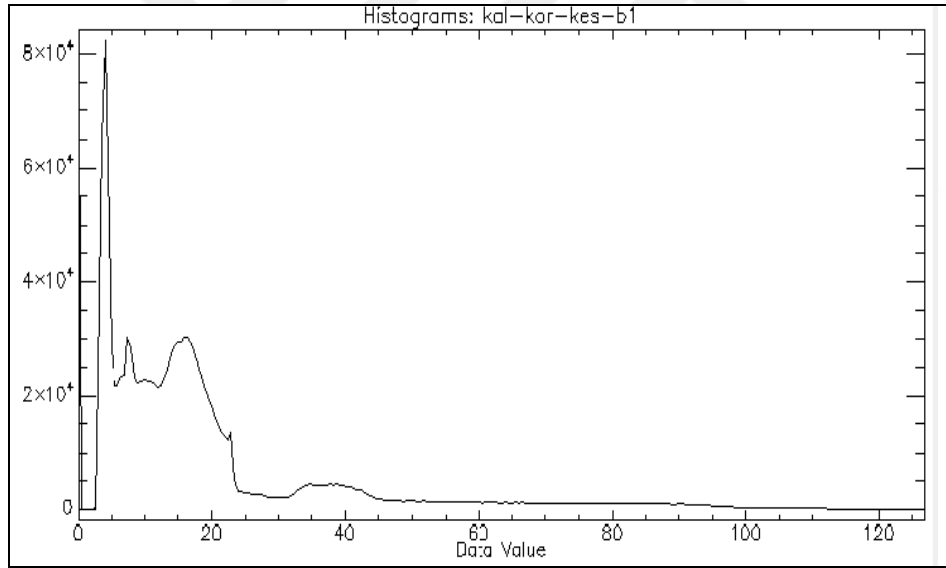
Şekil 4.1. 4 Eylül 2011 Tarihli Ham NOAA –AVHRR Görüntüsü



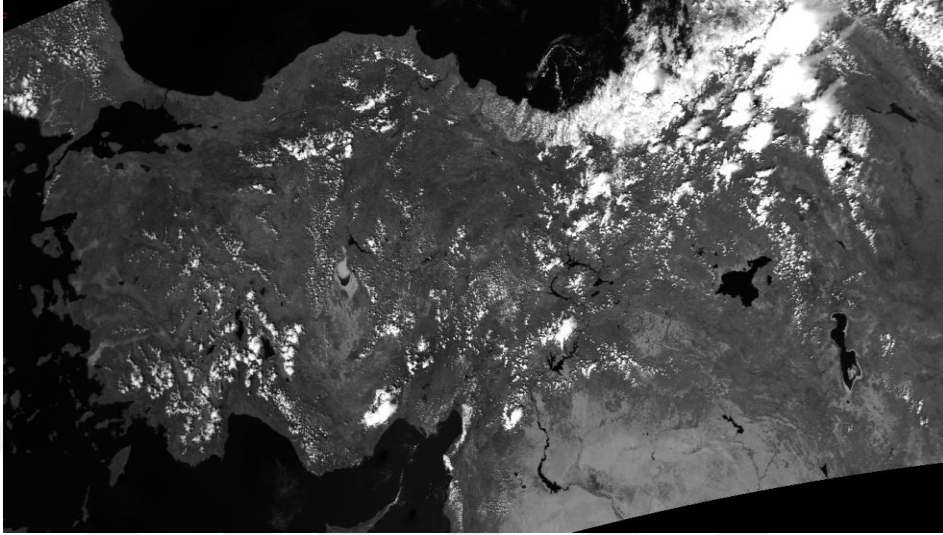
Şekil 4.2. 4 Eylül 2011 Tarihli Ham NOAA –AVHRR Görüntüsüne ait 5 Bant Histogram Değerleri



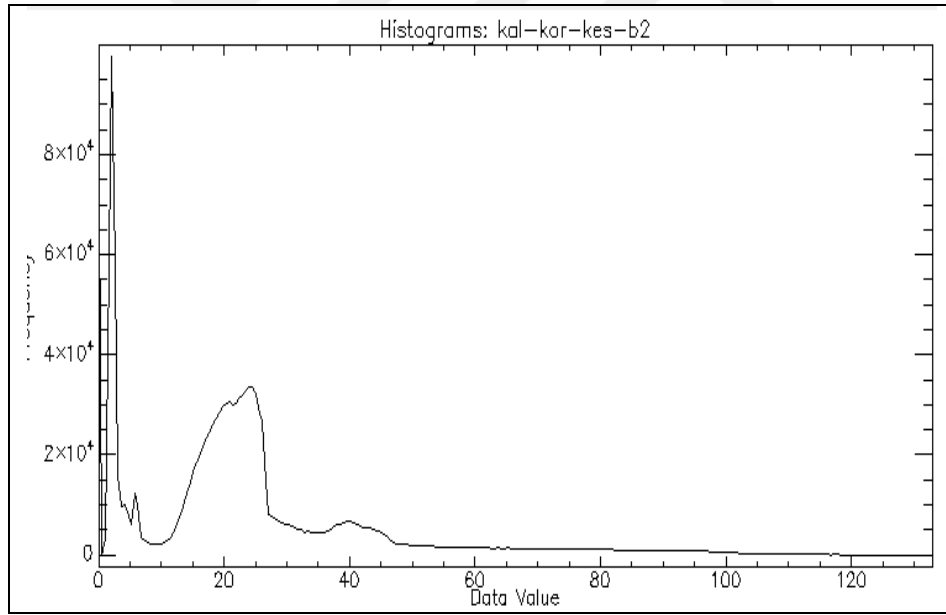
Şekil 4.3. 4 Eylül 2011 Tarihli Ön-İşlemleri Yapılmış AVHRR 1.Bant Görüntüsü



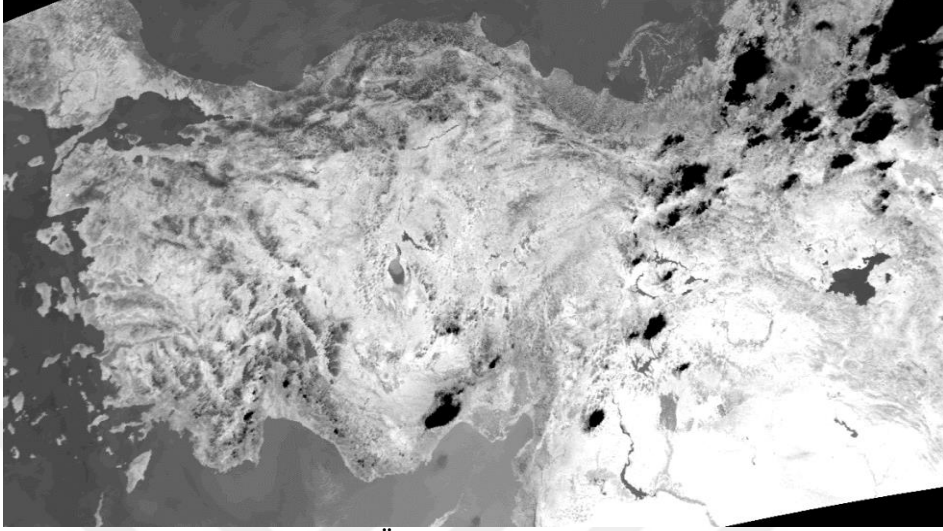
Şekil 4.4. 4 Eylül 2011 Tarihli AVHRR 1.Banta Ait histogram



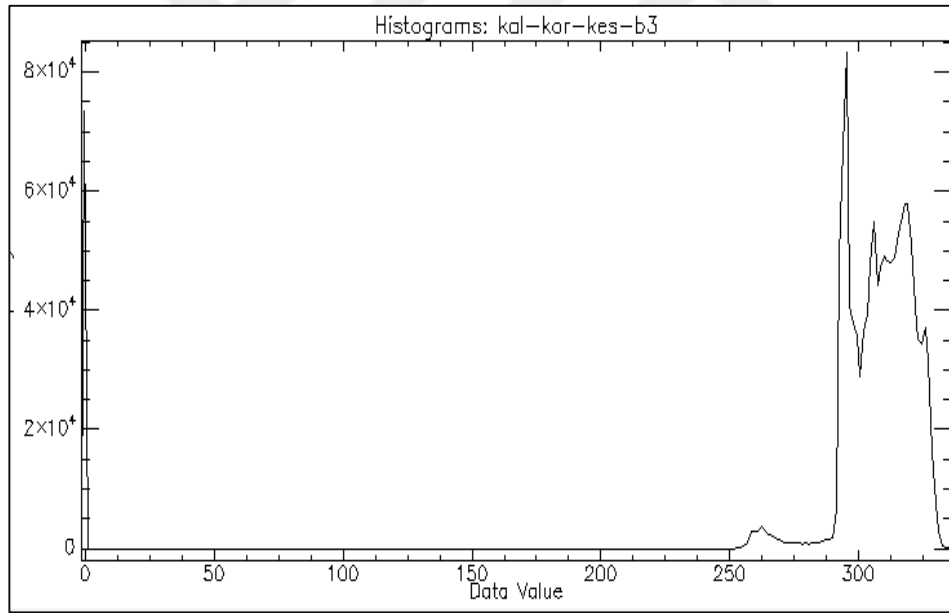
Şekil 4.5. 4 Eylül 2011 Tarihli Ön-İşlemleri Yapılmış AVHRR 2.Bant Görüntüsü



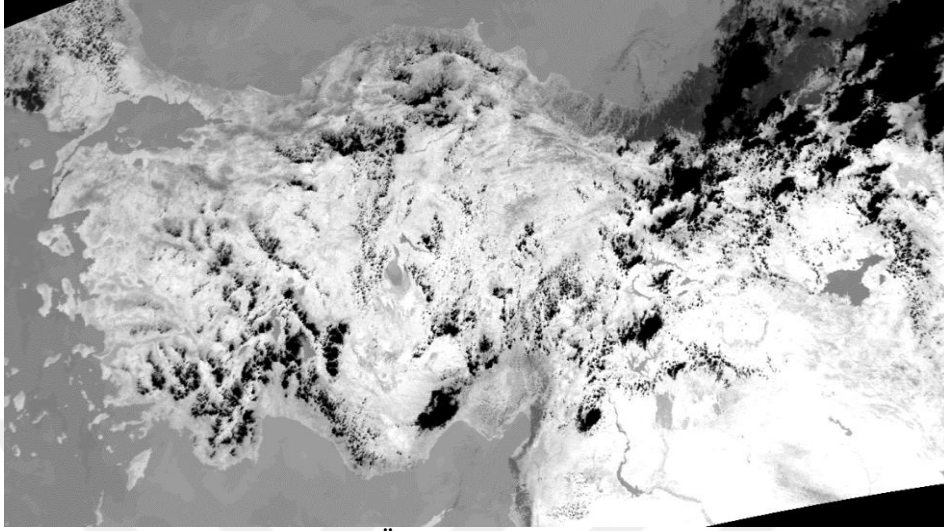
Şekil 4.6. 4 Eylül 2011 Tarihli AVHRR 2.Banta Ait Histogram



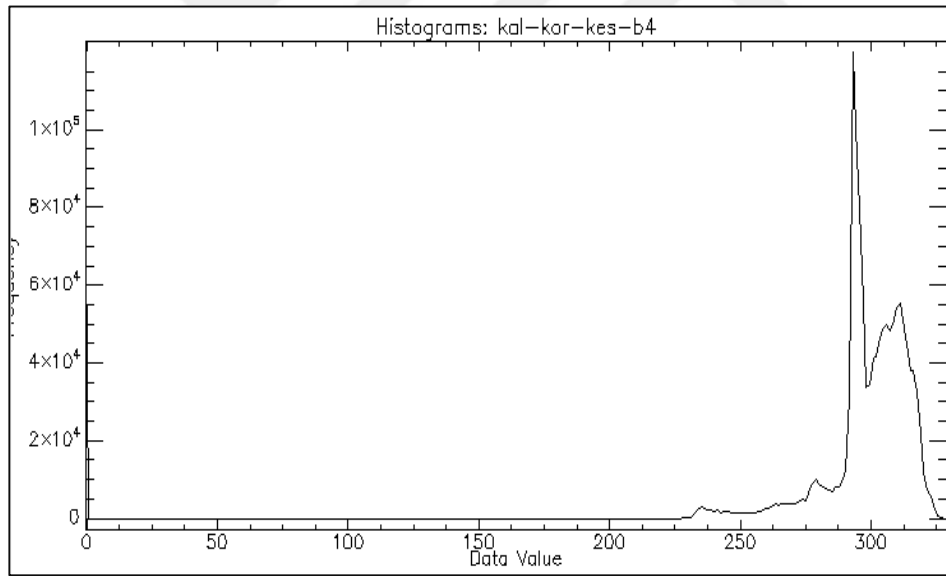
řekil 4.7. 4 Eylül 2011 Tarihli Ön-Iřlemleri Yapılmıř AVHRR 3.Bant Görüntüsü



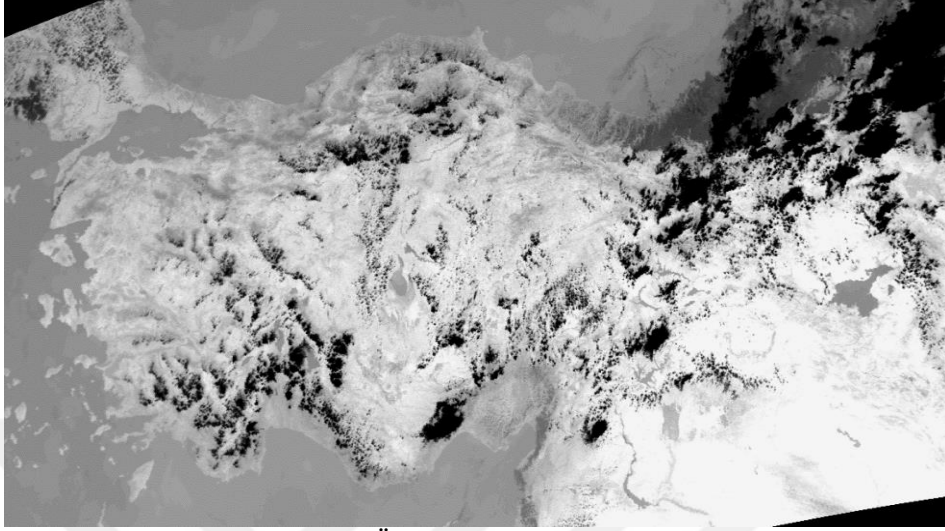
řekil 4.8. 4 Eylül 2011 Tarihli AVHRR 3.Banta Ait Histogram



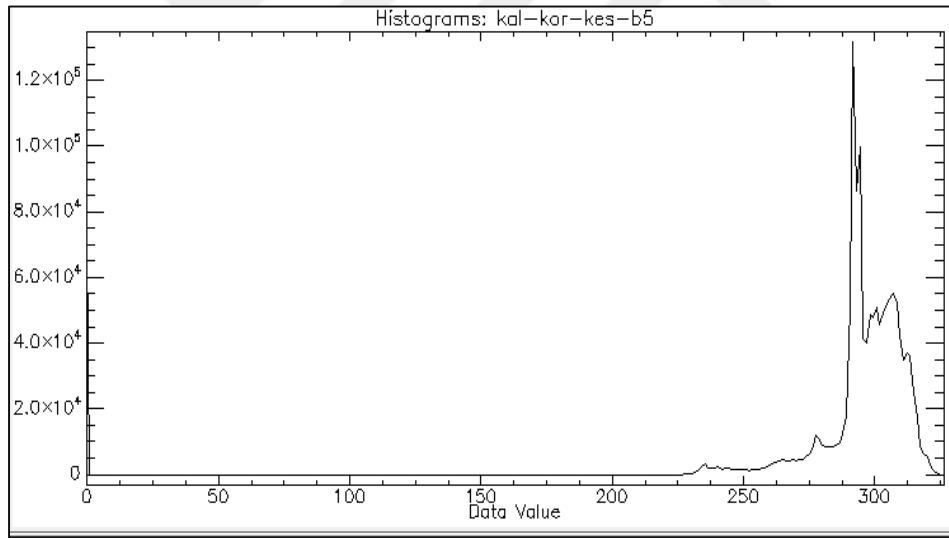
Şekil 4.9. 4 Eylül 2011 Tarihli Ön-İřlemleri Yapılmış AVHRR 4.Bant Görüntüsü



Şekil 4.10. 4 Eylül 2011 Tarihli AVHRR 4.Banta Ait Histogram

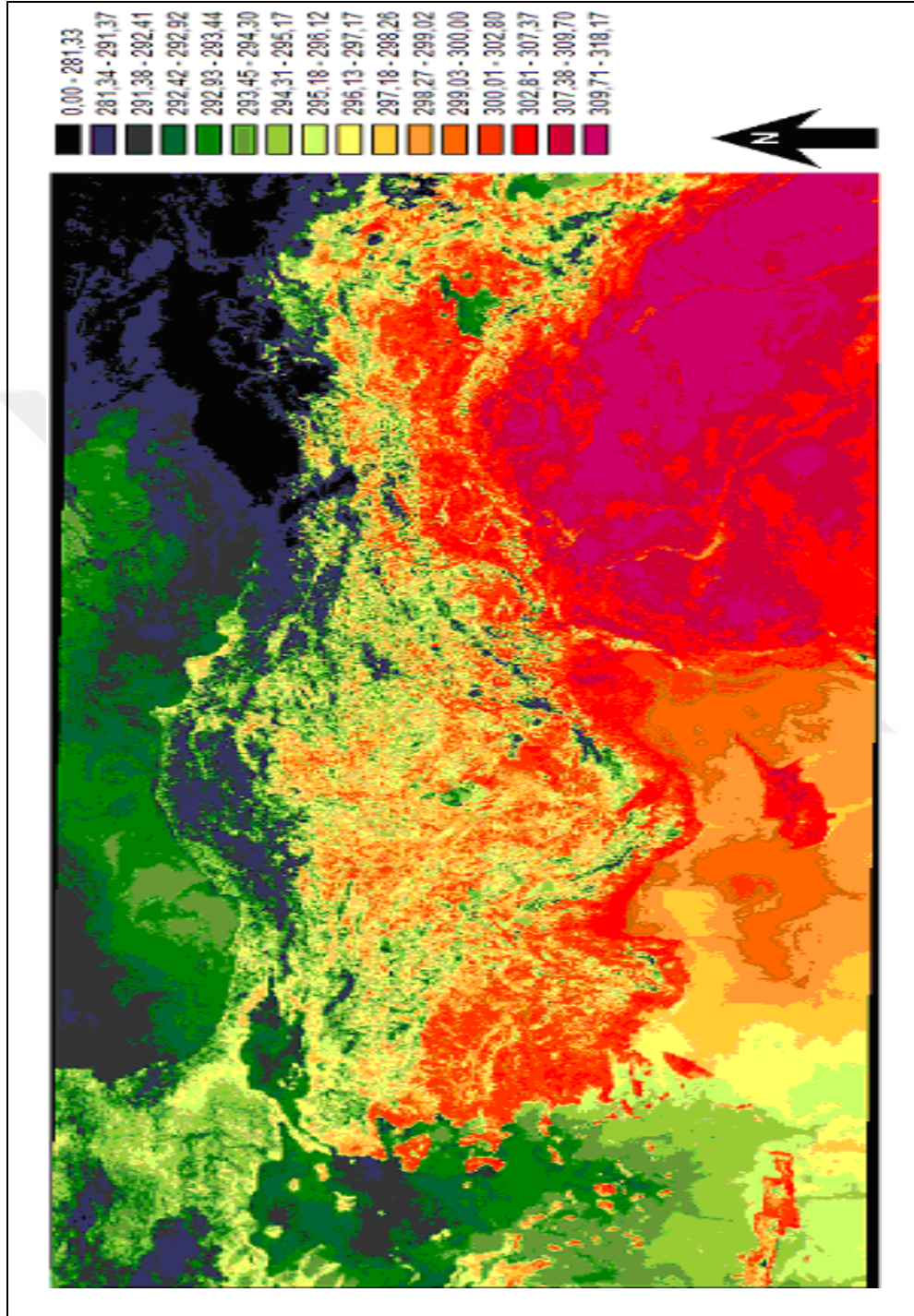


Şekil 4.11. 4 Eylül 2011 Tarihli Ön-İşlemleri Yapılmış AVHRR 5.Bant Görüntüsü



Şekil 4.12. 4 Eylül 2011 Tarihli AVHRR 5.Banta Ait Histogram

Ön-işlemleri yapılarak hazır hale getirilen uydu verilerinden LST değerlerini elde etmek için eşitlik 3.15'te yer alan Sobrino LST algoritması kullanılmıştır. Şekil 4.13'te LST haritası Türkiye ölçeğinde görülmektedir.



Şekil 4.13. Sobrino 1994 ile Hesaplanan LST (Yer Yüzey Sıcaklığı) Haritası

Bu işlemlerden sonra çalışma bölgesi kesilmiş ve IDL programı kullanılarak uydu verilerinden Adana meteoroloji istasyonunun enlem ve boylam bilgilerine ait piksel verileri çekilmiştir. Adana meteoroloji istasyonuna ait çalışmada kullanılacak meteorolojik parametrelerden Precipitable Water (PW) ve yer yüzey sıcaklık verileri Devlet Meteoroloji İşleri Genel Müdürlüğünden alınmıştır. Çizelge 5.5'te Çalışmada makine öğrenmesi metodlarının eğitim ve test bölümlerinde kullanılacak girdi ve çıktı parametrelerinin betimsel istatistik özellikleri yer almaktadır. Çizelgeden görüldüğü üzere verilerin genel standart sapmaları büyük olup veri açıklığı geniş bir skalada yer almaktadır.

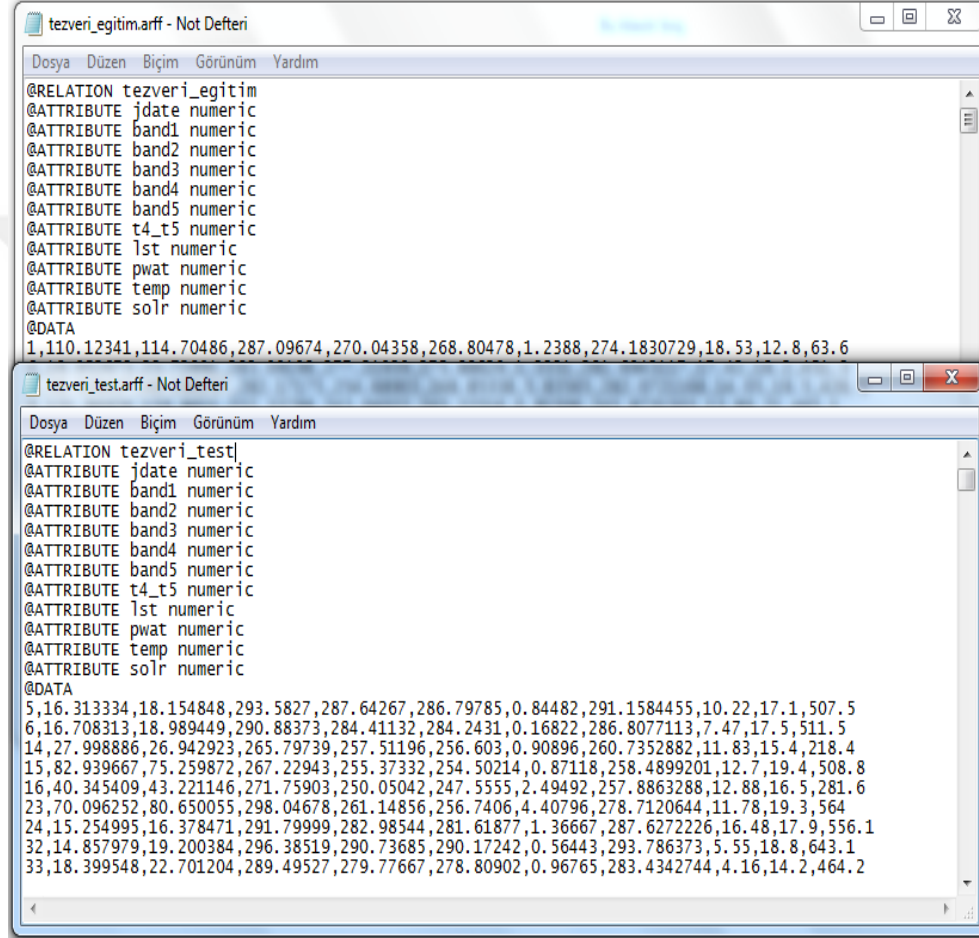
Çizelge 4.1. Verilerin Betimsel İstatistikleri

Veri Alanı	En Küçük Değer		En Büyük Değer		Aritmetik Ortalama		Standart Sapma	
	Eğitim	Test	Eğitim	Test	Eğitim	Test	Eğitim	Test
JDATE	1,0	3,0	366	364	182,2	185,2	105,7	108,9
BAND1	2.9	5,9	497	684	35,9	46,4	39,7	54,6
BAND2	1.6	4,3	500	645	40,9	51.5	40,5	53.6
BAND3	249,6	250,2	682	960	297,9	297,4	23,7	40,7
BAND4	211,7	217,3	595	776	281,2	281,5	28,1	36,4
BAND5	208,7	216,8	597	754	279,0	279,4	27,7	35,3
T4-T5	-2,0	-0,25	8,1	22,0	2,2	2.1	1,6	1,9
LST	219,6	219,3	601,9	1087,8	289,6	290,3	31,0	51,7
PWAT	4,0	1,9	48,5	62,8	22,4	22,3	9,9	10,6
TEMP	7,4	7	40,0	39,5	24,9	24,6	7,5	7.5
SOLR	15,0	7,6	985,0	987,4	600,3	599,9	240,4	246,3

4.2. Veri Setinin Düzenlenmesi

WEKA veri tabanlarından veri çekebilse de csv gibi metin dosyalarını okuyabilse de kendi asıl veri dosyası olan arff tipindeki dosyalar ile çalışmak sağlıklıdır. Özünde bir metin dosyası olan arff dosya tipinde; dosya adının

@relation kodlaması ile, her bir veri alanı @attribute kodlaması ile, değerlerin ise @data kodlaması ile aralarında virgül ile ayrılarak sütun sayısı kadar yan yana yazılarak ifade edilmektedir. Bu şekilde oluşturulan bir satır belli bir güne tekabül eden aşağıda tabloda verilen özelliklere karşılık gelen bir veri kayıdır.



Şekil 4.14. ARFF Dosya Yapısı

Veri setinde, 11'i de sayısal (nümerik) özellikte olan veri alanı yer almaktadır. Yapılan hesaplamalardan ve hazır olarak ölçüm verilerinden elde edilen veri alanları aynı güne gelenleri bir Excel tablosunda toplanmıştır. 2010 yılından 2016 yılı da dahil olmak üzere 7 yıllık, günlük veriye karşılık gelmektedir.

Ancak yılın her gününe ait bütün veri başlıklarında veri bulunmamakta olup bazı kayıtlarda eksik alanlar olduğu için ilgili günler veri setinden çıkarılmıştır. Sağlıklı şekilde 932 güne karşılık gelen veri kaydı üzerinden işlemler yapılmıştır. 2010, 2011, 2012, 2013 yıllarına ait 596 günlük veri kaydı eğitim veri seti olarak kullanılmış, 2014, 2015, 2016 yıllarına ait 336 günlük veri kaydı ise test veri seti olarak kullanılmıştır. Oran %36'ya %64 olup, makine eğitimi uygulamalarında kullanılan bir orandır. Weka'da varsayılan olarak girilen tek veri setini %66'lık eğitim, %34 test verisi olarak ayırmaktadır. Eğitim veri miktarının yüksek oranda tutulması genelde daha iyi sonuçların elde edilmesini sağlamaktadır. %2'lik dezavantaja rağmen yıllara ait günlük veriler julian gün değeri de (yılın kaçınıcı günü olduğu) bir veri alanı olarak kullanıldığı için yıl verileri kesintisiz olarak iki gruba ayrılmıştır. Eğitim grubu 4 yıllık, test grubu ise 3 yıllık veriyi kapsamıştır.

Çizelge 4.2. Veri Setinin İçeriği

Alan Adı	Veri Türü	İçerdiği Bilgi
JDATE	Nümerik	Yılın Kaçınıcı Günü Olduğu Verisi (Julian Günü)
BAND1	Nümerik	Görünür Alan Yansıma Uydu Verisi
BAND2	Nümerik	Görünür Alan Yansıma Uydu Verisi
BAND3	Nümerik	Yakın Kızıl Ötesi Alan Işıması Uydu Verisi
BAND4	Nümerik	Kızıl Ötesi Alan Işıması Uydu Verisi
BAND5	Nümerik	Kızıl Ötesi Alan Işıması Uydu Verisi
T4-T5	Nümerik	Hesaplanan Değer
LST	Nümerik	Hesaplanan Değer
PWAT	Nümerik	Yoğuşabilir Su Buharı (Meteoroloji Verisi)
TEMP	Nümerik	Hava Sıcaklığı (Meteoroloji Verisi)

SOLR	Nümerik	Yeryüzüne Gelen Güneş Enerjisi (Meteoroloji Verisi)
------	---------	---

Veri seti, uydu verisi, uydu verilerinden hesaplanmış veri, meteoroloji verisi ve Julian gün verisini içerdiği için makine öğrenmesi modellerini elde etmede en iyi veri kombinasyonunu sağlamağa yönelik literatürde henüz bir uzlaşma tespit edilemediğinden, elde edildiği kaynakla ilişkili olarak, verilerden kombinasyonlar oluşturulmuştur (Zhou ve ark., 2021). Çeşitli çalışmalarda çalışmanın amacına uygun çeşitli veri kombinasyonları kullanılmaktadır. Genellikle meteoroloji verileri ilk makine öğrenmesi araştırmalarında sıklıkla kullanılmış, uydu verileri son yıllarda kullanılmakta, olup çeşitli kaynaklardan gelen hibrit veri setleri yaygınlaşmıştır. Araştırmada kullanılan 8 adet veri seti kombinasyonu aşağıdaki tabloda verilmiştir.

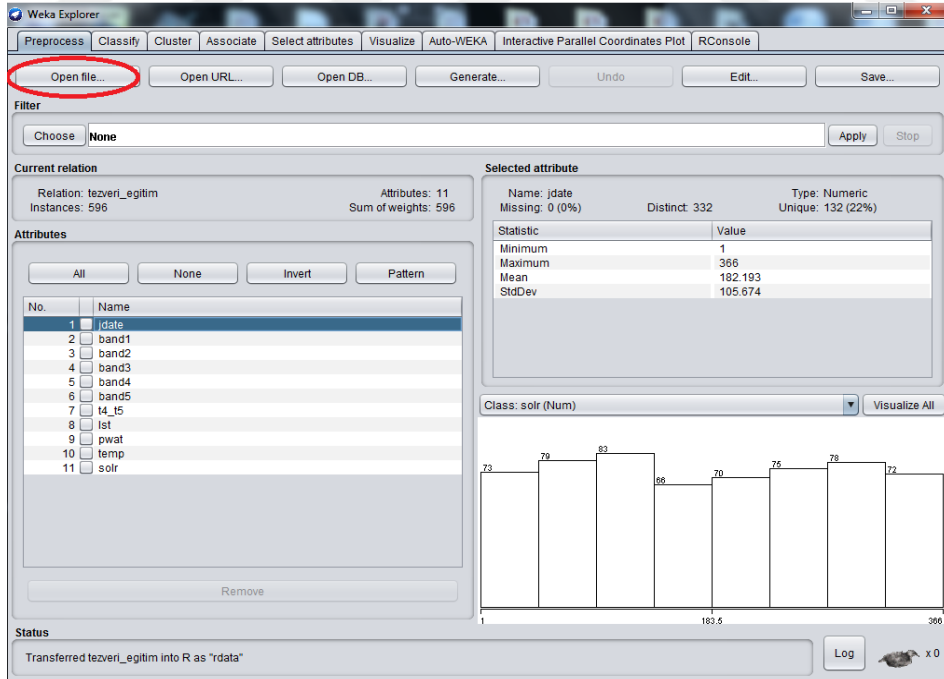
Çizelge 4.3. Veri Seti Kombinasyonları

A Grubu	B Grubu	C Grubu	D Grubu	E Grubu	F Grubu	G Grubu	H Grubu
JDATE	JDATE	JDATE	JDATE	JDATE	JDATE	JDATE	JDATE
BAND1	BAND1	BAND1		BAND1	BAND1		
BAND2	BAND2	BAND2		BAND2	BAND2		
BAND3	BAND3	BAND3		BAND3		BAND3	
BAND4	BAND4	BAND4		BAND4		BAND4	
BAND5	BAND5	BAND5		BAND5		BAND5	
T4-T5	T4-T5		T4-T5		T4-T5	T4-T5	
LST	LST		LST		LST	LST	
PWAT			PWAT	PWAT	PWAT	PWAT	PWAT
TEMP			TEMP	TEMP	TEMP	TEMP	TEMP

SOLR	SOLR	SOLR	SOLR	SOLR	SOLR	SOLR	SOLR
------	------	------	------	------	------	------	------

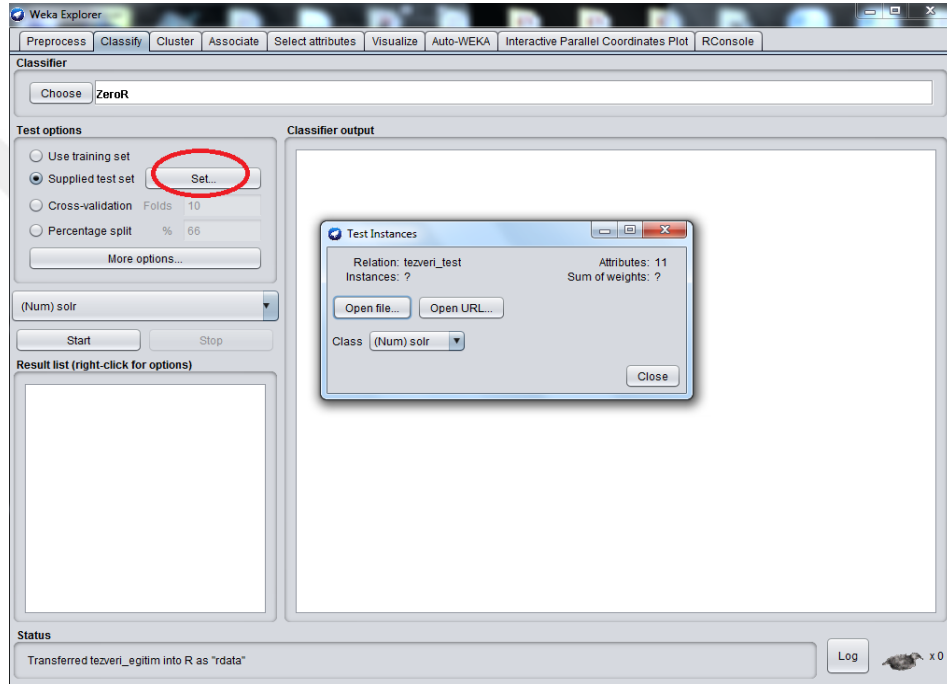
Tabloda A Grubu, tüm veri alanlarını kapsamaktadır. B Grubunda sadece uydu verileri ve hesaplanmış veriler bulunmaktadır. C Grubunda sadece uydu verileri kullanılmıştır. D Grubunda hesaplanmış veriler ile meteoroloji verileri kullanılmıştır. E Grubunda hesaplanmış veriler dışında uydu ve meteoroloji verileri bulunmaktadır. F Grubunda uydu verilerlerinden sadece görünür bant verileri bulunmaktadır. G Grubunda uydu verilerlerinden kızılötesi bant verileri bulunmaktadır. H Grubunda ise sadece meteoroloji verileri bulunmaktadır.

Eğitim ve test amaçlı olarak kullanılacak olan veri setleri WEKA yazılımına ayrı ayrı yüklenmiştir. Bu veri seti girişi kullanıcıya bağlı olarak yazılım tarafından otomatik olarak seçilebildiği gibi kullanıcıda veri miktarını kendisi belirleyebilir. Weka explorer arayüzünden preprocess sekmesi altında eğitim veri seti yükleme şekil 4,15 de gösterilmiştir.



Şekil 4.15. Eğitim Veri Setinin Yükleme Ekranı

Test veri seti ise Classify sekmesi altında yer alan Supplied test set isimli secenek üzerinden yüklenerek makine öğrenmesi metodları çalıştırmaya uygun hale getirilmiştir.Şekil 4.16’da test veri seti yükleme penceresi yer almaktadır.



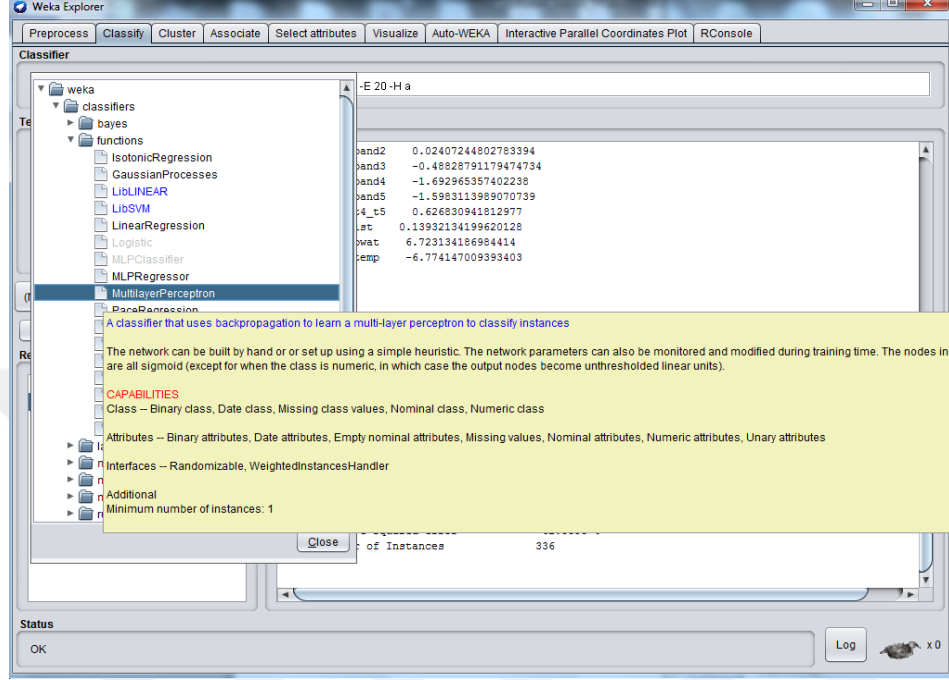
Şekil 4.16. Test Veri Setinin Yükleme Ekranı

4.3. Çok Katmanlı Algılayıcı Uygulaması

ANN tabanlı algoritmalar içinde en yağın kullanımı olan MLP hem sınıflandırma hem de tahmin problemlerinde başarılı sonuçlar ortaya koyduğu için Güneş enerjisi tahmininde araştırmacıların sıklıkla başvurduğu bir yöntemdir (Şenkal,2019).

Bu çalışmada kullanılan ANN algoritmalarından biri olan MLP metodu kullanılmıştır. Metodunun seçimi için Explorer ara yüzü altında, Preprocess sekmesinden, veri seti yüklenip, Classify sekmesinden, functions grubundan MLP

algoritması seçilip, çalıştırılır. Şekil 4.17 MLP seçim ekranını göstermektedir.

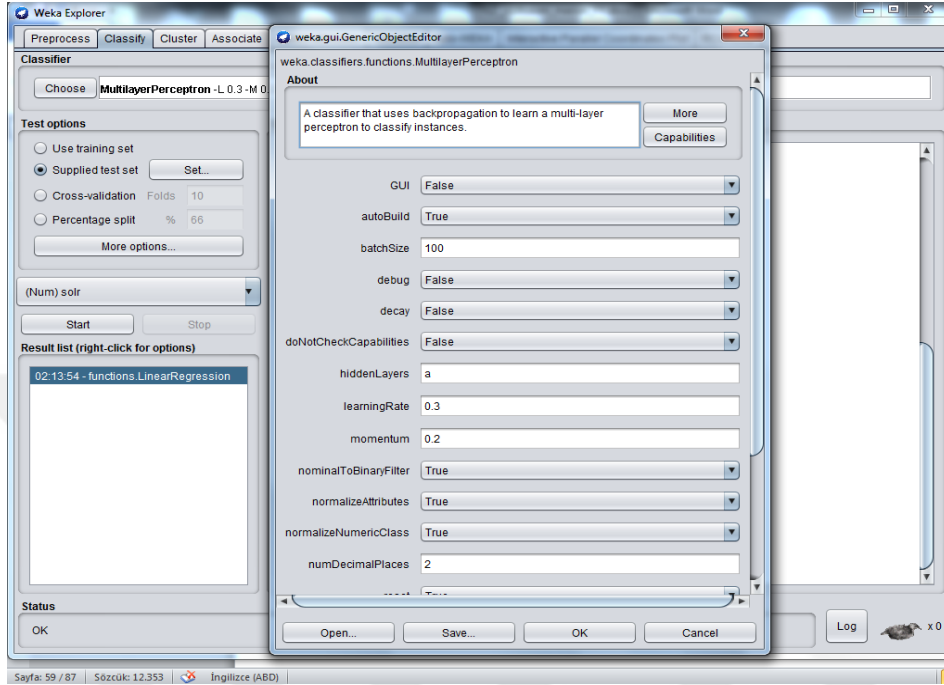


Şekil 4.17. Çok Katmanlı Algılayıcı Seçim Ekranı

MLP algoritmasının çalışması için değişik ayarlar mevcuttur. Kullanıcılar bu tip ANN çalışmalarında olduğu gibi en iyi sonuçlara ulaşabilmek açısından birçok ayar denemesi yapmaktadır (Hochreiter, 1998).

Bu çalışmada yapılan birçok deneme neticesinde WEKA ayarlarının (weka.classifiers.misc.InputMappedClassifier -I -trim -W weka.classifiers.functions.MultilayerPerceptron -- -L 0.234 -M 0.123 -N 500 -V 0 -S 0 -E 20 -H 21 -output-debug-info) olarak seçilip, algoritma özelliklerinden Debug aktif hale getirilip, gizli katmandaki nöron sayısını otomatik yerine 21 adet seçilmesi, öğrenme oranını 0.234'e ve momentum değeri 0.123'a ayarlanması durumunda en iyi MLP sonuçlarına ulaşılmıştır. Şekil 4.18 MLP ayar penceresi görülmektedir.

Eğitim sonrası uygulanan test verilerini MLP algoritması ile koşturulması ile tüm veri setleri için hesaplanan tahmin değerleri çizelge 4.6'da verilmiştir.



Şekil 4.18. Çok Katmanlı Algılayıcı Ayar Ekranı

Çizelgeden görüldü üzere E Grubu verilerle en yüksek tahmin değerini olan $R=0,88.42$ 'lik bir tahmin başarısı elde edilirken B Grubu verilerle ise en düşük tahmin başarısı olan $R=0,62.81$ 'lik değer elde edilmiştir.

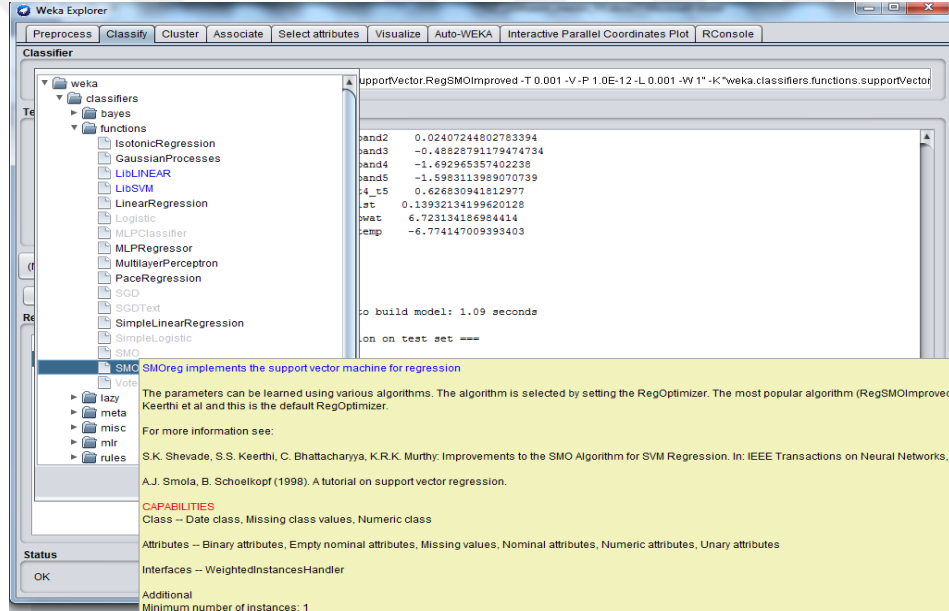
Çizelge 4.4. Çok Katmanlı Algılayıcı Sonuçları

	R	RMSE
A Grubu	0.8602	153.9469
B Grubu	0.6281	226.1079
C Grubu	0.7329	273.8941
D Grubu	0.8170	187.5498
E Grubu	0.8842	116.2783
F Grubu	0.8365	136.0534
G Grubu	0.8579	138.1388

H Grubu	0.8693	140.1118
---------	--------	----------

4.4. Destek Vektör Regresyonu Uygulaması

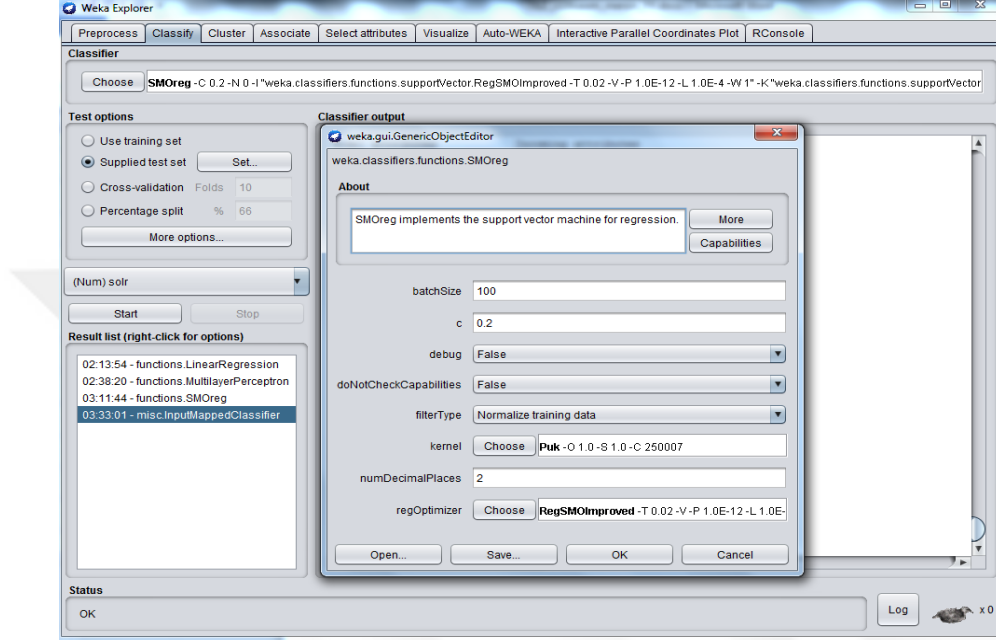
SVR algoritması WEKA altında Sequential Minimal Optimization Regressor (SMOReg) adıyla functions grubu altında bulunmaktadır. Verileri ikililer şeklinde alarak, optimizasyon yaparak sonuca ulaşan bu algoritma çeşitli çekirdek fonksiyonlarına sahiptir (Baitharu ve ark 2015). Şekil 4.19'da SVR için seçim ekranı görülmektedir.



Şekil 4.19. Destek Vektör Regresyonu Seçim Ekranı

Explorer ara yüzü altında, Preprocess sekmesinden veri seti yüklenip, Classify sekmesinden, functions grubundan SVR algoritması ("weka.classifiers.misc. Input Mapped Classifier -I -trim -W weka.classifiers.functions.SMOreg --- C 0.4 -N 0 -I weka.classifiers.functions.Support Vector. Reg SMOImproved -T 0.025 -V -P 1.0E-12 -L 1.0E-4 -W 1" -K

"weka.classifiers.functions.supportVector.Puk -O 0.6 -S 1.5 -C 250007") şeklinde seçilip, algoritma özelliklerinden optimize edici olarak RegSMOImproved ile çekirdek türü olarak Puk tipi seçilip, çalıştırılmıştır.



Şekil 4.20. Destek Vektör Regresyonu Ayar Ekranı

Eğitim sonrası uygulanan test verileri SVR ile koşturulduğunda, D Grubu verilerle en yüksek tahmin değeri $R=0,89.26$ 'lık bir skor, C Grubu verilerle en düşük tahmin değeri $R=0,75.30$ 'luk bir skor tespit edilmiştir.

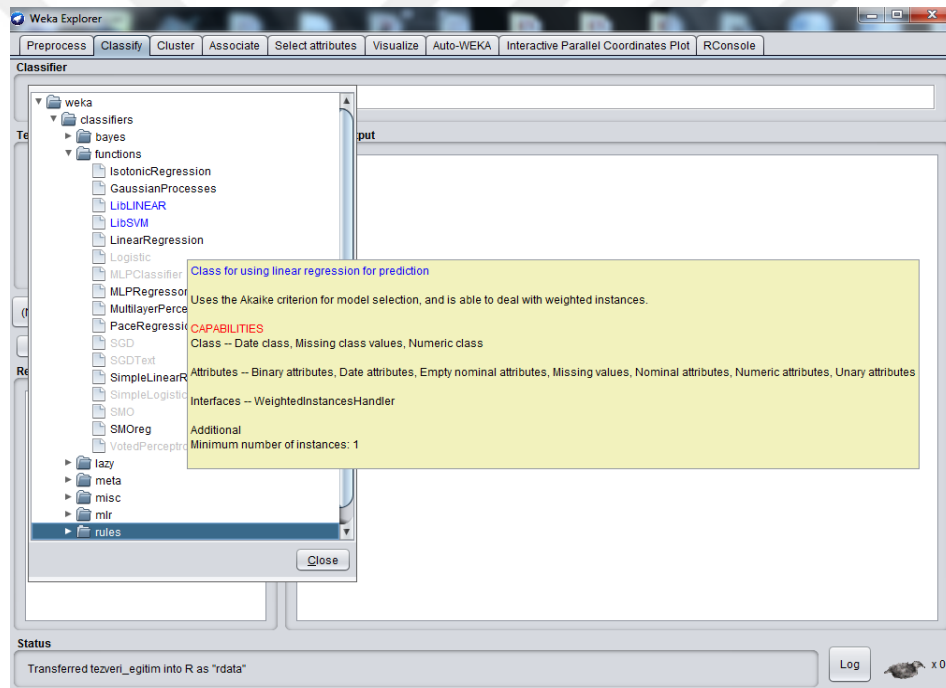
Çizelge 4.5. Destek Vektör Regresyonu Sonuçları

	R	RMSE
A Grubu	0.8892	113.1675
B Grubu	0.7538	166.6783
C Grubu	0.7530	166.9815
D Grubu	0.8926	111.9712
E Grubu	0.8882	113.7084
F Grubu	0.8902	113.0620

G Grubu	0.8905	112.5266
H Grubu	0.8877	114.5607

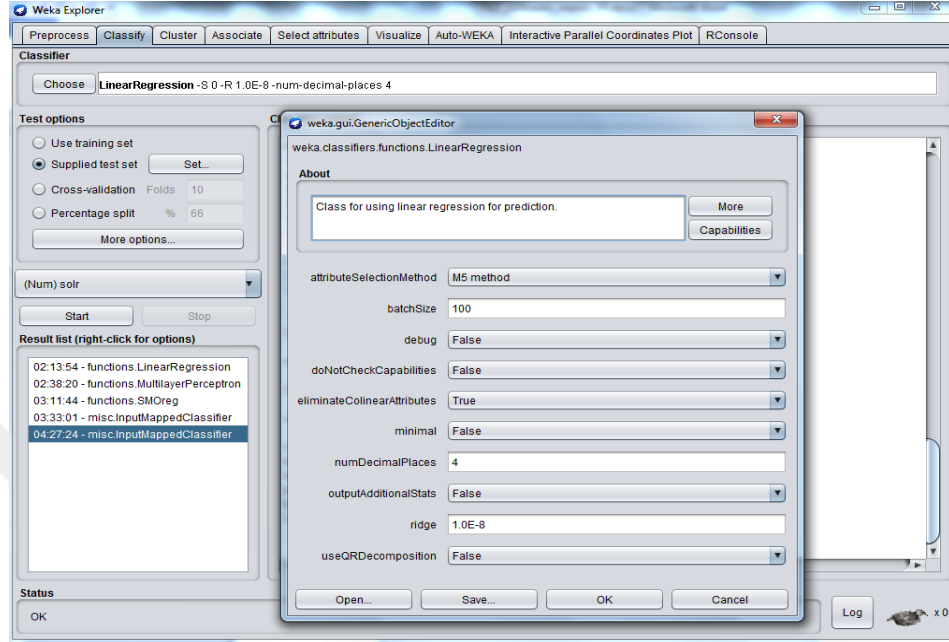
4.5. Regresyon Analizi Uygulaması

Explorer ara yüzü altında, Preprocess sekmesinden veri seti yüklenip, Şekil 4.21’de görüldüğü gibi Classify sekmesinden, functions grubundan lineer regresyon algoritması seçilmiştir.



Şekil 4.21. Regresyon Analizi Seçim Ekranı

Şekil 4.22’de görüldüğü gibi. regresyon analizi ayarlarından; (weka.classifiers.functions.LinearRegression -S 2 -R 1.0E-8 -num-decimal-places4) seçilip, özellik seçiminde GreedyMethod’u kullanıma eklenip, çalıştırılmıştır.



Şekil 4.22. Regresyon Analizi Ayar Ekranı

Eğitim sonrası uygulanan test verileri lineer regresyon algoritması ile koşurulduğunda, A Grubu verilerle en yüksek tahmin değerini $R=0,79.92$ 'lık bir skor ile, C Grubu verilerle en düşük tahmin değerini $R=0,49.66$ 'lık bir skor ile doğru tespit ettiği gözlenmiştir.

Çizelge 4.6. Lineer Regresyon Sonuçları

	R	RMSE
A Grubu	0.7992	148.9686
B Grubu	0.5910	200.9215
C Grubu	0.4966	219.7464
D Grubu	0.7686	157.6898
E Grubu	0.7811	154.768
F Grubu	0.7829	153.3795
G Grubu	0.7985	148.5875
H Grubu	0.7978	148.7078

4.6. Uygulama Sonuçları

Çalışmada kullanılan tüm veri setlerinin makine öğrenmesi algoritmalarında çalıştırılması ile elde edilen başarı (R) ve hata değerleri (RMSE) çizelge 4.7’de yer almaktadır.

Çizelge 4.7. Veri Setlerine göre Algoritmaların Başarı ve Hata değerleri

VERİ		MLP		SVR		LR	
Alanları	Grubu	r	RMSE	r	RMSE	r	RMSE
JDATE BAND1 BAND2 BAND3 BAND4 BAND5 T4-T5 LST PWAT TEMP SOLR	A Grubu	0.860	153.94	0.889	113.16	0.799	148.96
JDATE BAND1 BAND2 BAND3 BAND4 BAND5 T4-T5 LST SOLR	B Grubu	0.628	226.10	0.753	166.67	0.591	200.92
JDATE BAND1 BAND2 BAND3 BAND4 BAND5 SOLR	C Grubu	0.732	273.89	0.753	166.98	0.496	219.74
JDATE T4-T5 LST PWAT TEMP SOLR	D Grubu	0.817	187.54	0.892	111.97	0.768	157.68
JDATE BAND1 BAND2 BAND3 BAND4 BAND5 PWAT TEMP SOLR	E Grubu	0.884	116.27	0.888	113.70	0.781	154.76
JDATE BAND1 BAND2 T4-T5 LST PWAT TEMP SOLR	F Grubu	0.836	136.05	0.890	113.06	0.782	153.37
JDATE BAND3 BAND4 BAND5 T4-T5 LST PWAT TEMP SOLR	G Grubu	0.857	138.13	0.890	112.52	0.798	148.58
JDATE PWAT TEMP SOLR	H Grubu	0.869	140.11	0.887	114.56	0.797	148.70

Çizelge 4.8’de çalışmada kullanılan 3 algoritmanın en başarılı olduğu veri setlerinden tahmin edilen değerler ile gerçek ölçüm değerlerinin test veri setindeki ilk 25 verisinin çıktısı görülmektedir.

Çizelge 4.8. Algoritmalarla Göre En Başarılı Güneş Enerjisi Tahmin Sonuçları

MLP TAHMİNİ (W/m ²)				SVR TAHMİNİ (W/m ²)				LR TAHMİNİ (W/m ²)			
inst#	actual	predicted	error	inst#	actual	predicted	error	inst#	actual	predicted	error
1	507.5	474.074	-33.426	1	507.5	521.851	14.351	1	507.5	582.987	75.487
2	511.5	521.033	9.533	2	511.5	550.992	39.492	2	511.5	606.302	94.802
3	218.4	388.364	169.964	3	218.4	449.957	231.557	3	218.4	472.553	254.153
4	508.8	446.441	-62.359	4	508.8	518.68	9.88	4	508.8	534.943	26.143
5	281.6	393.923	112.323	5	281.6	444.336	162.736	5	281.6	513.383	231.783
6	564	506.199	-57.801	6	564	519.261	-44.739	6	564	630.019	66.019
7	556.1	402.583	-153.517	7	556.1	478.001	-78.099	7	556.1	550.663	-5.437
8	643.1	693.234	50.134	8	643.1	667.578	24.478	8	643.1	648.71	5.61
9	464.2	640.007	175.807	9	464.2	636.914	172.714	9	464.2	530.133	65.933
10	576.2	521.106	-55.094	10	576.2	589.969	13.769	10	576.2	595.298	19.098
11	654	652.659	-1.341	11	654	673.875	19.875	11	654	590.273	-63.727
12	648.2	710.857	62.657	12	648.2	699.504	51.304	12	648.2	656.39	8.19
13	676.4	722.686	46.286	13	676.4	722.939	46.539	13	676.4	710.372	33.972
14	662.4	734.156	71.756	14	662.4	722.62	60.22	14	662.4	767.821	105.421
15	574.2	428.132	-146.068	15	574.2	456.817	-117.383	15	574.2	514.051	-60.149
16	279.8	525.076	245.276	16	279.8	599.105	319.305	16	279.8	528.979	249.179
17	164.3	243.453	79.153	17	164.3	379.651	215.351	17	164.3	914.87	750.57
18	793.8	774.05	-19.75	18	793.8	770.896	-22.904	18	793.8	704.599	-89.201
19	795.9	684.967	-110.933	19	795.9	695.973	-99.927	19	795.9	664.72	-131.18
20	737.2	593.481	-143.719	20	737.2	633.779	-103.421	20	737.2	587.276	-149.924
21	562.6	696.555	133.955	21	562.6	709.202	146.602	21	562.6	628.778	66.178
22	879.4	882.12	2.72	22	879.4	740.83	-138.57	22	879.4	510.316	-369.084
23	641.9	692.132	50.232	23	641.9	685.717	43.817	23	641.9	585.106	-56.794
24	894.9	798.511	-96.389	24	894.9	816.615	-78.285	24	894.9	722.264	-172.636
25	616.3	790.448	174.148	25	616.3	825.446	209.146	25	616.3	762.527	146.227

Çalışmada kullanılan tüm veri setleri 8 ana başlık altında gruplanmış ve makine öğrenmesi metodlarının girdi verisi olarak belirlenmiştir. Bu gruplamanın amacı en iyi sonuçların hangi veri seti ile elde edilebileceğinin ve hangi verilerin

yer yüzeyine gelen güneş ışınımına etkisi olduğunun belirlenebilmesini sağlamaktır. Çizelgedeki sonuçları hem veriseti grubu hemde çalıştırılan makine öğrenmesi algoritmaları açısından irdelemek gerekmektedir.

Tüm girdi verilerinin tek bir grup olduğu A grubu veri setinde en iyi sonuçlar $R=0,889$ ve $RMSE= 113,16$ değerleri SVM algoritmasında elde edilirken MLP algoritmasının sonuçlarıda başarılı değerler vermiştir. LR algoritması ise bariz şekilde diğer iki algoritma bazında geri kalmıştır.

Sadece ham uydu verileri ve uydu verilerinden hesaplanmış girdilerin bulunduğu B grubunda veri setinde en iyi sonuçlar $R=0,753$ ve $RMSE= 166,67$ değerleri SVM algoritmasında elde edilmesine rağmen kullanılan üç algoritmanın sonuçlarının çok başarılı olduğundan bahsedilemez.

Sadece uydu verilerinin girdi olduğu C grubu veri setinde en iyi sonuçlar $R=0,753$ ve $RMSE= 166,98$ değerleri SVM algoritmasında elde edilirken LR $R=0,499$ ve $RMSE= 219,74$ değerleri ile gerçek anlamda uyumsuzluk sorunlarını ortaya koymuştur. Bu iki veri grubunu diğer veri setlerinden ayıran başlıca özellik iki setinde hiçbir meteorolojik ölçüm değerini girdi parametresi olarak kullanmamış olmasıdır. Çalışmada en başarısız tahmin değerleri B ve C grubu verilerinden oluşan setlerde bulunmuştur. Bu sonuçlar bizi bu tip bir çalışmada meteorolojik parametrelerin gerekli olduğunu göstermektedir

Uydu ham bant verilerinden hesaplanmış meteorolojik ürünler olan PWAT, T4-T5 ve LST ile meteoroloji verilerinin kullanıldığı D grubu veri setinde SVM algoritmasının $R=0,892$ ve $RMSE=111,97$ değerlerini verirken MLP algoritmasının sonuçlarıda $R=0,817$ ve $RMSE=187,54$ değerleri ile nispeten SVM'ye yakın ve başarılı olarak belirlenmiştir. Bu veri grubu hiçbir ham veri barındırmamaktadır. Ham uydu verisi yerine ham verilerden hesaplanmış ürünlerin kullanılmasının daha kullanışlı olduğunu ortaya çıkarmıştır.

Uydu ham bant verileri ve meteorolojik ölçüm değerlerinin kullanıldığı E grubu veri setinde SVM algoritması $R=0,888$ ve $RMSE= 113,70$ ve MLP algoritması $R=0,884$ ve $RMSE= 116,27$ değerleri ile başarılı sonuçlar ortaya

koymuştur. Bu veri setinde LR algoritmasında $R=0,781$ ve $RMSE= 154,76$ ile kullanılabilir sonuçlar ortaya koymuştur. Ham uydu verilerinin meteorolojik verilerle kullanılmasında meteorolojik ölçüm verilerinin başarımlarını arttırmada önemli rol oynadığını göstermektedir zira sadece ham verilerin kullanıldığı C grubu veri setine göre doğruluklarda ciddi farklar görülmektedir.

F ve G grubu veri setlerini birbirlerinden ayıran özellik F grubunda AVHRR dedektörünün görünür bölgede algılama yapan ilk 2 band verileri girdide yer alırken G grubunda ise AVHRR dedektörünün kızılötesi bölgede algılama yapan 3 band verileri meteorolojik ölçüm değerleri ve uydu verilerinden hesaplanan ürün değerleri ile algoritmalarda çalıştırılmıştır. Bu veri setlerinin oluşturulmasındaki amaç AVHRR dedektörünün görünür bölge ve kızılötesi bölge verilerini sonuçlar üzerindeki etkisini belirlemeye çalışmaktır. Sonuçlar irdelendiğinde SVM algoritmasında iki veri setinin aynı değeri verdiği görülürken MLP ve LR algoritmalarında kızılötesi uydu verilerini girdi olarak kullanmanın başarımlarını küçüktür olsa bir artış sağlandığı görülmektedir.

Sadece meteoroloji istasyonundan alınan ölçüm verilerinden oluşturulan H grubu veri setinden elde edilen başarımlar dikkate alınacak değerdedir. Bu veri setinde kullanılan üç algoritmadan SVM ve MLP sırası ile $R=0,887$ ve $RMSE= 114,54$ ve $R=0,869$ ve $RMSE= 140,11$ değerleri ile birbirlerine yakın sonuçlar verirken LR algoritması $R=0,797$ ve $RMSE= 148,70$ değerleri ile bu iki algoritmaya göre nispeten başarısız olmuştur. Bu başarımlar sadece uydu verilerini girdi parametresi olduğu C grubu veri setine göre bariz anlamda başarılı olmuştur. Bu sonuçlar bu veri setleri ve makine öğrenmesi algoritmalarında meteorolojik ölçüm değerlerinin sonuçlar üzerine ciddi etkileri olduğunu göstermektedir.

Yapılan tüm çalışmalarda JDATE verisi tüm veri setlerinden çıkarıldığında tahmin sonuçları başarımları düşmektedir. Bunun nedenini yer yüzeyine gelen güneş ışımasını hesaplama formülünde jülien tarih değerinin ana değişkenlerden biri

olmasıdır (Iqbal,1983). Bu nedenle JDATE verisi tüm veri setlerinde girdi parametresi olarak yer almaktadır.

Meteorolojik ölçüm verilerinin girdi parametresi olmadığı veri setlerinde LR algoritmasının iyi sonuçlar vermediği görülmektedir. Zira LR algoritması en iyi sonucunu $R=0,799$ ve $RMSE= 148,96$ ile tüm verilerin girdi parametresi olduğu A grubu veri setinde verirken en kötü sonuçları $R=0,496$ ve $RMSE= 219,74$ ve $R=0,591$ ve $RMSE= 200,92$ ile meteorolojik ölçüm değerlerinin kullanılmadığı C ve B grubu veri setinde vermiştir.

Çalışmada en büyük başarımlık değeri SVR algoritmasında elde edilmiştir. Bu algortimada genel olarak R değeri 0,888'in üzerinde tahmin doğruluk değeri elde edilmiş iken B ve C grubu veri setlerinde en düşük başarımlık değeri olan $R=0,753$ elde edilmiştir. Bu sonuç SVR algoritmasının sadece uydu verileri ve uydudan elde edilen ürünler ile çalıştırılmamasının doğruluğu arttıracak gerçeğini ortaya çıkarmıştır.

MLP algortiması genel olarak SVR algoritmasını takip etmekte ve benzer ama biraz daha düşük başarımlık oranları ortaya koymaktadır. Bu fark MLP ile SVR algortimalarının temellerinde yer alan istatistiksel kurulum ve işleyiş farklılığından kaynaklanmaktadır.



5. SONUÇ ve ÖNERİLER

Dünya üzerindeki herhangi bir konum için istenilen herhangi bir anda güneş enerjisi değeri ölçümü yapmanın zorlukları ve maliyeti düşünüldüğünde makine öğrenmesi ile uydu verileri kullanılarak bu tahminlerin gerçekleştirilmesi üzerine birçok çalışmalar yapılmakta olup tahminlerin doğruluk oranlarını arttırılmaya yönelik çalışmalar literatürde artan sayıda yer almaktadır.

Bu tez çalışmasında; 2010-2016 yıllarını kapsayan NOAA-AVHRR veriler kullanılarak, WEKA makine öğrenmesi yazılımında yer alan ve literatürde sıklık ile kullanılan MLP, SVR ve LR makine öğrenmesi algoritmaları kullanılarak çalışma bölgesi için yeryüzeyine gelen güneş ışıınımı değerleri tahmin edilmeye çalışılmıştır. Adana bölgesine ait 7 yıllık uydu verileri NOAA ‘nın sitesinden indirilerek, işlenmiş ve aralarında meteoroloji istasyonundan alınan veriler ile eşleşen, çalışmaya uygun olan toplam 932 adet farklı günlere ait uydu ve meteoroloji verisi, çalışmanın girdi ve çıktı değerleri olarak belirlenmiştir.

Makine öğrenmesi metotlarında, AVHRR dedektöründen elde edilen uydu görüntülerinin 5 bant değeri, band4 ile band5’in fark değeri, LST değeri, PWAT ve sıcaklık değerleri girdi parametresi olarak kullanılırken yer yüzeyine gelen güneş enerjisi değerleri çıktı değeri olarak kullanılmıştır. Girdilerin güneş enerjisi değerlerine etkisinin anlaşılabilmesi için 8 adet değişik veri seti kombinasyonları 3 algoritma ile çalıştırılarak Çizelge 4.7’deki sonuç tablosu elde edilmiştir. Çalışma genelinde SVR algoritması D grubu veri setinde $R=0,8926$ ve $RMSE=111,97$ değerleri ile çalışmadaki en başarılı sonucu vermiştir. Çalışmada kullanılan makine öğrenmesi metodları çıktı sonuçları açısından girdi parametrelerine göre kıyaslandığında ise MLP algoritması E grubu veri setinde $R=0,8842$ ve $RMSE=116,27$ değerini, LR algoritması A grubu veri setinde $R=0,7992$ ve $RMSE=148,96$ değerini vermiştir.

Yapılan çalışmada her üç makine öğrenmesi algoritmasında girdi parametresi olarak meteorolojik ölçüm verilerini kullanmanın başarımlarına

büyük oranda arttırdığını söylemek mümkündür.

Sadece ham uydu verilerini kullanıldığı veri seti gruplar ile yapılan makine öğrenmesi çalışmasının sonuçlarının tatminkar derecede başarılı güneş ışınımı değerleri tahmini vermediği görülmektedir. Bu tip meteoroloji tabanlı çalışmalarda meteorolojik verilerin uydu verileri ile kullanılmasının daha iyi sonuçlar verebileceği yapılacak diğer çalışmalar için önerilmektedir.

Başarı oranının artması model parametrelerine veri setinin büyüklüğü ve farklılığına bağlıdır. Bu tür problemlerde veri seti büyüklüğü uzun dönemli gözlemlerden elde edilen ve binlerce veriyi içeren veri setleri ile yapılarak hassasiyet ve doğruluk değerlerinde artış sağlanabilmektedir. Bu büyük veri seti tanımı ise bilimin yeni gelişen bir dalı olan “Big Data” tabiri olarak literatürde yer almaktadır.

Bu tip çalışmalarda yeni geliştirilecek makine öğrenmesi algoritmalarının daha kapsamlı veri setleri ile kullanımının ilerki dönemlere yapılacak bilimsel çalışmaların doğruluk değerlerini arttıracığı aşîkârdır.

Ayrıca bu çalışmada kullanılan AVHRR dedektörünün algılama yaptığı dalga boylarından farklı algılama aralığını kullanan dedektör verileri ile yapılacak çalışmalarda farklı sonuçlar elde dîlebilir. Bu nedenle ilerki çalışmalar farklı veri seti ve farklı algoritmalar ile desteklenirken bu çalışmanın sonuçları iyi bir altlık sunacağı düşünölmektedir.

KAYNAKLAR

- Ackerman, S.A., Knox, J. A., 2015, Meteoroloji Atmosferimizi Anlamak, Nobel Yayınevi, 3. Baskıdan çeviri / 602 Syf.
- Alpaydın, E., 2012, Yapay Öğrenme, Boğaziçi Üniversitesi Yayınevi, İstanbul, 496s.
- Aydemir, E., 2018, Weka ile Yapay Zekâ, Seçkin Yayınevi, Ankara, s216.
- Baitharu, T. R., Pani, S.K., Dhal, S.K., 2015, Comparison of Kernel Selection for Support Vector Machines Using Diabetes Dataset, Journal of Computer Sciences and Applications, Vol. 3, No. 6, 181-184.
- Becker, F., 1987, The impact of spectral emissivity on the measurement of land surface temperature., Quart. J. Roy. Meteorology soc., 109: 365-378.
- Becker, F. and Li, Z. L., 1990a "Temperature-independent spectral indices in thermal infrared bands," Remote Sensing Environ., vol. 32, pp. 17-33.
- Besharat, F., Dehghan, A.A., Faghieh, A.R., 2013, Empirical Models For Estimating Global Solar Radiation: A Review And Case, Study Renew. Sustain. Energy Rev., 21, pp. 798-821.
- Brownlee, J., 2016, Machine Learning Mastery With Weka, Jason Brownlee Ebook URL: <https://machinelearningmastery.com> Son Erişim Tarihi: 22.11.2021
- Cihlar, J., Li, H., Li, Z., Chen, J., Pokrant, H., and Huang, F. 1997. Multitemporal, multichannel AVHRR data sets for land biosphere studies - artifacts and corrections, Remote Sensing of Environment, 60, 35-57.
- Chen, J., Zhu, W., Yu, Q., 2021, Estimating half-hourly solar radiation over the Continental United States using GOES-16 data with iterative random forest Renewable Energy, Volume 178, November, Pages 916-929.
- Cornejo-Bueno, L., Casanova-Mateo, C., Sanz-Justo, J., Salcedo-Sanz, S., 2019 Machine learning regressors for solar radiation estimation from satellite data Solar Energy Volume 183, Pages 768-775.
- Cortes, C., Vapnik, V., 1995, Support-vector networks, Mach. Learn., 20 (3)

- (1995), pp. 273-297.
- Curran, P.J., 1985, Principles of Remote Sensing, Longman Company, London, p. 260.
- Czajkowski, K.P., Goward, S.N. and Ouaidrarı, H.,1998. Impact of AVHRR filter functions on surface temperature estimation from the split window approach, International Journal of Remote Sensing, 19 (10), 2007-2012
- Dinç, U., Yeğingil, İ., Peştemalcı, V., Dinç, O., Kandırmaz, H.M., 2001. “Uzaktan Algılamanın Temel Esasları ve Bazı Uygulamalar, Lisansüstü Yaz Okulu Ders Notları”, TÜBİTAK Bilim Adamı Yetiştirme Grubu-Çukurova Üniversitesi Bilimsel İşbirliği, Ç.Ü. Fen-Edebiyat Fakültesi Fizik Bölümü, Adana, 138 s.
- Dincer, I. ve Acar C. 2015, A review on clean energy solutions for better sustainability, International Journal of Energy Research, 39(5):585:606.
- Fallahia, S., Amanollahi, J., Tzanis, C. G., Ramlic, M.F., 2018, Estimating Solar Radiation Using NOAA/AVHRR And Ground Measurement Data, Atmospheric Research Volume 199, 1, Pages 93-102.
- Fente, D. N., Singh, D. K., 2018. Weather Forecasting Using Artificial Neural Network. In 2018 Second International Conference On Inventive Communication And Computational Technologies (ICICCT), IEEE, pp. 1757-1761.
- Hiregoudar, S., 2020, Ways to Evaluate Regression Models, URL: <https://towardsdatascience.com/ways-to-evaluate-regression-models-77a3ff45ba70> Son Erişim Tarihi: 22.11.2021
- Hochreiter, S., 1998. The vanishing gradient problem during learning recurrent neural nets and problem solutions. International Journal of Uncertainty.
- İnternet: Remote Sensing, URL: https://en.wikipedia.org/wiki/Remote_sensing Son Erişim Tarihi: 22.11.2021
- İnternet: Linear Regression, URL: https://en.wikipedia.org/wiki/Linear_regression Son Erişim Tarihi: 22.11.2021

- İnternet: Alfred Binet, URL: https://en.wikipedia.org/wiki/Alfred_Binet Son Erişim Tarihi: 22.11.2021
- Iqbal, M. 1983, An Introduction to Solar Radiation, Academic Press, Vancouver, p.220.
- Jiang, T., Gradus, J. L., & Rosellini, A. J., 2020. Supervised machine learning: a brief primer. Behavior Therapy, 51(5), 675-687
- Joselin, J., Dinesh, T., & Ashiq, M., 2018. A review on neural networks. Int. J. Trend Sci. Res. Dev.(IJTSRD), 2, 565-569.
- Kaba, K., Sarıgül, M., Avcı, M., Kandırmaz, H.M., 2018, Estimation Of Daily Global Solar Radiation Using Deep Learning Model. Energy Volume 162, 1, Pages 126-135.
- Kaelbling, L. P., Littman, M. L., & Moore, A. W., 1996. Reinforcement learning: A survey. Journal of artificial intelligence research, 4, 237-285.
- Karaman, Ö.A., Ağır, T.T., Arsel, İ., 2021, Estimation Of Solar Radiation Using Modern Methods, Alexandria Engineering Journal Volume 60, Issue 2, Pages 2447-2455.
- Kerr, Y. H., and Lagouarde, J .P.,1992. On the derivation of landsurface temperature from AVHRR data, in Kerr, Y.H., J.P. Lagouarde, and J. ImbeNon, Accurate Land Surface Temperature Retrieval from AVHRR Data with use of an Improved Split Window Algorithm, Remote Sensing of Environment, 41 (2/3), 197-209.
- Mueller, J.P., Massaron, L.,2018, Artificial Intelligence For Dummies, John Wiley & Sons, Inc., p.311.
- Orhan,U., 2021, Introduction to Machine Learning Lisans Ders Notları URL: <https://bmb.cu.edu.tr/uorhan/MachineLearning.htm> Son Erişim Tarihi: 22.11.2021
- Prata, A.J., and Platt, C.M.R.,1991 Land surface temperature measurements from the AVHRR, in 5th AVHRR Data users conference,, pp. 433-438, Tromso.
- Raha,S., Pal S., 2010, A study on the present environmental scenario due to

- pollution by conventional energy sources—and remedies: solar cell with nanotechnology, *Strat. Plann. Energy Environ.*, 30 (2), pp. 8-19.
- Ravinesh, C. D., Şahin, M., Adamowski, J. F., Mi, J. 2019, Universally Deployable Extreme Learning Machines Integrated With Remotely Sensed MODIS Satellite Predictors Over Australia To Forecast Global Solar Radiation: A New Approach., *Renewable and Sustainable Energy Rev.*, 104:235-261.
- Roodschild, M., Sardiñas, J. G., Will, A., 2020, A new approach for the vanishing gradient problem on sigmoid activation. *Progress in Artificial Intelligence*, 9(4), 351-360.
- Sobrino, J.A., Liang Lİ, Z.-, Stoll, M. P., 1991. Improvements in the Split-Window Technique for Land Surface Temperature Determination, *IEEE Transactions on Geoscience and Remote Sensing*, 32 (2), 243-253.
- Şahin, M., 2013, Comparison Of Modelling ANN And ELM To Estimate Solar Radiation Over Turkey Using NOAA Satellite Data, *International Journal of Remote Sensing*, 34:21, 7508-7533.
- Şahin, M., Kaya, Y., Uyar, M., 2012, Comparison Of ANN And MLR Models For Estimating Solar Radiation İn Turkey Using NOAA/AVHRR Data. *Advances In Space Research.*, 51: 891-904.
- Şahin, M., Kaya, Y., Uyar, M. and Yıldırım, S., 2014, Application Of Extreme Learning Machine For Estimating Solar Radiation From Satellite Data, *Int. J. Energy Res.*, 38: 205-212.
- Şeker, Ş.E., 2013, *Weka İle İş Zekası ve Veri Madenciliği*, Cinius Yayınları, İstanbul, s206 Syf.
- Senkal, O., 2018, Comparison Of Incoming Solar Radiation At Different Air Density Regimes Using Neural Network Model, *Russ. Meteorol. Hydrol.* 43, 49–55.
- Şenkal, O., 2010, Modeling Of Solar Radiation Using Remote Sensing And Artificial Neural Network İn Turkey, *Energy*, Volume 35, Issue 12, 2010, Pages 4795-4801, ISSN 0360-5442

- Şenkal, O., 2015, Solar Radiation And Precipitable Water Modeling For Turkey Using Artificial Neural Networks, *Meteorology and Atmospheric Physics*, 127, 481–488.
- Şenkal, O., Kuleli, T., 2009, Estimation Of Solar Radiation Over Turkey Using Artificial Neural Network And Satellite Data, *Applied Energy*, Volume 86, Issues 7–8, Pages 1222-1228, ISSN 0306-2619.
- Şenkal, O., Şahin, M., Peştimalcı, V., 2010, The Estimation of Solar Radiation for Different Time Periods, *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 32:13, 1176-1184.
- Van De Griend, A.A., and Owe, M., 1993 : On the relationship between thermal emissivity and the normalized difference vegetation index for natural surfaces *Int. J. Rem. Sens.*, 14, 1119-1137.
- Vapnik, V., Golowich, S.E., Smola, A., 1996, *Support Vector Method for Function Approximation, Regression Estimation, And Signal Processing*, *Adv. Neural Inf. Process. Syst.*, pp. 281-287
- Wan, Z., 1999., *MODIS Land-Surface Temperature Algorithm Theoretical Basis Document (LST ATBD)*, pp. 75, Institute for Computational Earth System Science, Santa Barbara.
- Wan, Z., and Dozier, J., 1996. A Generalized Split-Window Algorithm for Retrieving Land-Surface Temperature from Space, *IEEE Transactions on Geoscience and Remote Sensing*, 34(4), 892-905.
- Yıldız, B.Y., 2009, Türkiye'nin Bölgesel Atmosferik Değişkenlere Bağlı LST Algoritmalarının Oluşturulması, *Ç.Ü. Fen Bilimleri Enstitüsü Fizik Ana Bilim Dalı Doktora Tezi*, Adana.
- Zhou, Y., Liu, Y., Wang, D., Liu, X., Wang, Y., 2021, A Review On Global Solar Radiation Prediction With Machine Learning Models In A Comprehensive Perspective, *Energy Conversion And Management*, Volume 235, 113960.



ÖZGEÇMİŞ

1999 yılında Gaziantep Üniversitesi MYO Kontrol Sistemleri ve Teknolojisi ön lisans, 2009 yılında Mersin Üniversitesi MYO Bilgisayar Teknolojisi ve Programlama ön lisans programını bitirdi. Eskişehir Anadolu Üniversitesi İktisat bölümünden 2010 yılında, İşletme bölümünden 2013 yılında mezun oldu. Sakarya Üniversitesi Bilgisayar ve Bilişim Mühendisliği anabilim dalında 2012 yılında tezsiz yüksek lisans yaptı. Halen bilişim alanında çalışmaktadır.