

**BAŐKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĐİ ANABİLİM DALI
ELEKTRİK ELEKTRONİK MÜHENDİSLİĐİ TEZLİ YÜKSEK
LİSANS PROGRAMI**

**RADYOSONDE RASATLARI İLE MAKİNA ÖĐRENMESİ TABANLI
HAVA DURUMU KESTİRİMİ**

HAZIRLAYAN

ERALP GÖĐEN

YÜKSEK LİSANS TEZİ

TEZ DANIŐMANI

DR. ÖĐR. ÜYESİ SELDA GÜNEY

ANKARA - 2022

**BAŐKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĐİ ANABİLİM DALI
ELEKTRİK ELEKTRONİK MÜHENDİSLİĐİ TEZLİ YÜKSEK
LİSANS PROGRAMI**

**RADYOSONDE RASATLARI İLE MAKİNA ÖĐRENMESİ TABANLI
HAVA DURUMU KESTİRİMİ**

HAZIRLAYAN

ERALP GÖĐEN

YÜKSEK LİSANS TEZİ

TEZ DANIŐMANI

DR. ÖĐR. ÜYESİ SELDA GÜNEY

ANKARA - 2022

BAŐKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Elektrik/Elektronik Mühendisliđi Anabilim Dalı Elektrik/Elektronik Mühendisliđi Tezli Yüksek Lisans Programı çerçevesinde Eralp GÖĞEN tarafından hazırlanan bu çalıřma, aőađıdaki jüri tarafından Yüksek Lisans Tezi olarak kabul edilmiřtir.

Tez Savunma Tarihi: 19/08/2022

Tez Adı: Radyosonde Rasatları ile Makine Öğrenmesi Tabanlı Hava Durumu Kestirimi

Tez Jüri Üyeleri (Unvanı, Adı - Soyadı, Kurumu)	İmza
Prof. Dr. Hasan Őakir BİLGE (Gazi Üniversitesi)
Prof. Dr. Hamit ERDEM (Baőkent Üniversitesi)
Dr. Öğr. Üyesi Selda GÜNEY (Baőkent Üniversitesi)

ONAY

Prof. Dr. Ömer Faruk ELALDI
Fen Bilimleri Enstitüsü Müdürü

Tarih : ... / ... / 2022

BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLER ENSTİTÜSÜ
YÜKSEK LİSANS / DOKTORA TEZ ÇALIŞMASI ORIJİNALLIK RAPORU

Tarih: ... / ... / 20...

Öğrencinin Adı, Soyadı : Eralp GÖĞEN

Öğrencinin Numarası : 21810358

Anabilim Dalı : Elektrik-Elektronik Mühendisliği Anabilim Dalı

Programı : Elektrik-Elektronik Mühendisliği Tezli Yüksek Lisans Programı

Danışmanın Unvanı/Adı, Soyadı : Dr. Öğr. Üyesi Selda GÜNEY

Tez Başlığı : Radyosonde Rasatları İle Makine Öğrenmesi Tabanlı Hava Durumu Kestrimi

Yukarıda başlığı belirtilen Yüksek Lisans/Doktora tez çalışmamın; Giriş, Ana Bölümler ve Sonuç Bölümünden oluşan, toplam 88 sayfalık kısmına ilişkin, 29 / 08 / 2022 tarihinde şahsım/tez danışmanım tarafından turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 7'dir. Uygulanan filtrelemeler:

1. Kaynakça hariç
2. Alıntılar hariç
3. Beş (5) kelimedenden daha az örtüşme içeren metin kısımları hariç

"Başkent Üniversitesi Enstitüleri Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Usul ve Esaslarını" inceledim ve bu uygulama esaslarında belirtilen azami benzerlik oranlarına tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Öğrenci İmzası:.....

ONAY

Tarih: ... / ... / 20...

Dr. Öğr. Üyesi Selda GÜNEY

.....

Bu tezi savunma sanayisinde yıllarını büyük bir fedakârlıkla geçiren ve ülkemizin savunma sektöründe ilerlemesine katkı sağlayan tüm tasarım mühendislerine ithaf ediyorum.

Eralp GÖĞEN

Ankara – 2022

TEŐEKKÜR

Lisans ve yüksek lisans eđitim hayatım boyunca, bana danıőmanlık ederek, yolumu aydınlatan ve yardımlarını esirgemeyen danıőman hocam Dr. Selda GÜNEY'e sonsuz teőekkürlerimi sunarım.

Bu çalıőmanın testlerinin gerçekleştirilmesini sađlayan ve bu konuda yardımlarını esirgemeyen T.C. Tarım ve Orman Bakanlığı, Meteoroloji Genel Müdürlüğü ve MAGNETRON Firması ve deđerli çalıőanlarına sonsuz teőekkürlerimi sunarım.

Tez süresi boyunca ihtiyaç duyduğum her noktada yardımlarını asla esirgemeyen Türkiye Amatör Uydu Teknolojileri Derneđi'nin (TAMSAT) saygıdeđer üyelerine teőekkür ederim.

Bu çalıőma boyunca engin bilgileriyle bana yol gösteren, tecrübe ve deneyimleri ile beni aydınlatan ve çalıőmalarım sırasında yanımda olan Yüksek Elektrik/Elektronik Mühendisi Sayın İlter KARADEDE ve Sayın Baran AKBIYIK'a teőekkür ederim.

Öđrenim hayatım boyunca maddi ve manevi desteklerini hiçbir zaman esirgemeyen, hep yanımda olan aileme yürekten sonsuz teőekkürlerimi sunarım.

ÖZET

ERALP GÖĞEN

**RADYOSONDE RASATLARI İLE MAKİNA ÖĞRENMESİ TABANLI HAVA DURUMU
TAHMİNİ**

Başkent Üniversitesi Fen Bilimleri Enstitüsü

Elektrik/Elektronik Mühendisliği Anabilim Dalı

2022

Geçmişten günümüze hava tahmini insanlık için önem arz etmektedir. Hava tahmininin hassas gerçekleştirilebilmesi sel, tsunami vb. doğal afetlere karşı önlemler alınarak oluşacak olumsuz etkileri en düşük seviyeye indirmeyi sağlayabilmektedir. Bu çalışma kapsamında radyosonde verilerini kullanarak hava durumu kestirimi yapılmaktadır. Bu kestirimde en yüksek ve en düşük sıcaklık tahmini yapılmaktadır. Makine Öğrenmesi Algoritmaları kullanarak kestirim gerçekleştirilmiştir. Daha önce literatürde bulunan sıcaklık tahmini çalışmalardan farklı olarak 3 yıllık Radyosonde rasat verileri kullanılmıştır. Bu sayede yerden 40 km yüksekliğe kadar 1mbar aralıklarla ölçülmüş veriler ile atmosfer, literatürdeki diğer çalışmalara göre çok daha hassas olarak modellenmiştir. Bu modelde ertesine ait en yüksek ve en düşük sıcaklık değerleri kestirilmiştir. Bu aşamada normalizasyon ve öznelilik çıkarma veya seçmenin sonuçlara etkileri analiz edilerek tahmin için en uygun model belirlenmiştir. MATLAB ortamında gerçekleştirilen yazılım ile farklı regresyon yöntemleri karşılaştırılmıştır. Bu analizler sonucunda Gauss Süreci Regresyonu yöntemini kullanarak 1,2 Ortalama Karekök Sapması ile ertesine ait en yüksek sıcaklık tahmini en yüksek doğrulukla elde edilmiştir. Aynı yöntem kullanarak 2,4 Ortalama Karekök Sapması oranı ile en düşük sıcaklık tahmini yapılmıştır. Sonuçlar literatürdeki çalışmalardan daha başarılı sıcaklık tahmini yapıldığını göstermektedir.

ANAHTAR KELİMELELER: Radyosonde, Hava Durumu Tahmini, Regresyon.

ABSTRACT

ERALP GÖĞEN

**ARTIFICIAL INTELLIGENCE BASED WEATHER FORECAST WITH RADIOSONDE
OBSERVATIONS**

Başkent University Institute of Science

Department of Electrical and Electronics Engineering

2022

Weather forecasting from past to present is important for humanity. The precise realization of the weather forecast can ensure that the negative effects that will occur are minimized by taking precautions against natural disasters such as floods, tsunamis, etc. Within the scope of this study, weather forecasting is made using radiosonde data. In this estimation, the highest and lowest temperatures are estimated. Estimation was made using Machine Learning Algorithms. Unlike the temperature estimation studies previously in the literature, 3-year radiosonde observation data were used. In this way, the atmosphere was modeled much more precisely than other studies in the literature with the data measured at 1mbar intervals up to 40 km above the ground. In this model, the highest and lowest temperature values for the next day are estimated. At this stage, the most appropriate model for the prediction is determined by analyzing the effects of normalization and attribute extraction or voter on the results. Different regression methods were compared with the software performed in MATLAB environment. As a result of these analyzes, the highest temperature estimate for the next day was obtained with the highest accuracy with 1.2 Mean Square Root Deviation using the Gaussian Process Regression method. Using the same method, the lowest temperature estimate was made with an average Square Root Deviation rate of 2.4. The results show that more successful temperature estimation is made than the studies in the literature.

KEYWORDS: Radiosonde, Weather Forecast, Regression.

İÇİNDEKİLER

TEŞEKKÜR.....	i
ÖZET	ii
ABSTRACT	iii
İÇİNDEKİLER.....	iv
EKLER	vi
ŞEKİLLER LİSTESİ	viii
SİMGELER VE KISALTMALAR LİSTESİ	x
1. GİRİŞ.....	1
1.1. Literatür Araştırması	3
2. HAVA DURUMU KESTİRİMİ.....	5
3. YÜKSEK İRTİFA METEOROLOJİ ÖLÇÜM CİHAZI RADYOSONDE	10
4. 3 BOYUTLU ATMOSFER VERİ SETİ	16
4.1 İstasyon Numarası:	18
5. KULLANILAN YÖNTEMLER	20
5.1 Regresyon:	20
5.1.1 Doğrusal Regresyon (Linear Regression):.....	21
5.2 Karar Ağacı Algoritması:	22
5.2.1 Regresyon Karar Ağacı Analizi:	22
5.3 Destek Vektör Makinesi:.....	23
5.4 Gauss Süreci Regresyonu:.....	25
6. DENEYSEL SONUÇLAR VE TARTIŞMA.....	27
6.1 Eğitim-1:	29
6.2 Eğitim-2:	33
6.3 Eğitim-3:	36
6.4 Eğitim-4:	39
6.5 Eğitim-5:	42
6.6 Eğitim-6:	45
6.7 Eğitim-7:	48

6.8	Eđitim-8:	51
6.9	Eđitim-9:	54
6.10	Eđitim-10:	57
6.11	Eđitim-11:	60
6.12	Literatür Karşılařtırması	65
7.	SONUÇ	67
8.	KAYNAKLAR	69



EKLER

EK 1 Örnek Tam bir TEMP KODU.....	72
-----------------------------------	----



TABLULAR LİSTESİ

Tablo 1 Radyosonde Özellikleri [12]	12
Tablo 2 Gerçekleştirilen Eğitim Sonuçları Özet Tablosu.....	28
Tablo 3 Eğitim-1 Sonucu Elde Edilen Başarım Sonuçları Tablosu	30
Tablo 4 Eğitim-2 Sonucu Elde Edilen Başarım Sonuçları Tablosu	33
Tablo 5 Eğitim-3 Sonucu Elde Edilen Başarım Sonuçları Tablosu	36
Tablo 6 Eğitim-4 Sonucu Elde Edilen Başarım Sonuçları Tablosu	39
Tablo 7 Eğitim-5 Sonucu Elde Edilen Başarım Sonuçları Tablosu	42
Tablo 8 Eğitim-6 Sonucu Elde Edilen Başarım Sonuçları Tablosu	45
Tablo 9 Eğitim-7 Sonucu Elde Edilen Başarım Sonuçları Tablosu	48
Tablo 10 Eğitim-8 Sonucu Elde Edilen Başarım Sonuçları Tablosu	51
Tablo 11 Eğitim-9 Sonucu Elde Edilen Başarım Sonuçları Tablosu	54
Tablo 12 Eğitim-10 Sonucu Elde Edilen Başarım Sonuçları Tablosu	57
Tablo 13 Eğitim-11 Sonucu Elde Edilen Başarım Sonuçları Tablosu	60
Tablo 14 Gerçekleştirilen Eğitim Sonuçları Tablosu	63
Tablo 15 Hava Durumu Kestirim Literatür Karşılaştırma Tablosu.....	65

ŞEKİLLER LİSTESİ

Şekil 2.1 Hava Durumu Verisi Toplama Yolları [6]	5
Şekil 2.2 Prognostik Çizelge Örneği [2].....	7
Şekil 2.3 850mb, 700mb, 500mb ve 300mb için GFS (Global Forecast System) Modelini Kullanan Bilgisayarla Çizilmiş 24 Saatlik Proglar [2]	8
Şekil 3.1 Meteoroloji Genel Müdürlüğü Tarafından Kullanılmakta Olan Yabancı Radyosonde Meteorolojik Ölçüm Cihazları.....	10
Şekil 3.2 Radyosonde Blok Şeması.....	11
Şekil 3.3 GNSS Tabanlı Yer İstasyonu Birimi Görseli	13
Şekil 3.4 GNSS Tabanlı Yer Alıcı Sistemine Ait Görsel	14
Şekil 3.5 Yer Alıcı Antenleri.....	14
Şekil 3.6 MGM 9. Bölge Müdürlüğü ile Birlikte Yapılan Rasat ve İncelemeler	15
Şekil 4.1 Radyosonde Veri Toplama Formatı	16
Şekil 4.2 Dünyanın Yuvarlaklığının Yer İstasyonu Antenine Olan Etkisi [13]	18
Şekil 4.3 MGM Bünyesinde Kurulu Bulunan Radyosonde İstasyonları ve İstasyon Numaraları [6]	19
Şekil 5.1 Radyosonde Verilerine Uygulanan Yöntemlerin Blok Diyagramı	20
Şekil 5.2 Örnek Bir Regresyon Yöntemi.....	20
Şekil 5.3 Örnek Bir Doğrusal Regresyon Grafiği	22
Şekil 5.4 SVM Sınıflandırma Örneği	24
Şekil 6.1 Eğitim-1 Sonucu Elde Edilen Tepki Grafiği	31
Şekil 6.2 Eğitim-1 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği	31
Şekil 6.3 Eğitim-1 Sonucu Elde Edilen Hata Grafiği.....	32
Şekil 6.4 Eğitim-2 Sonucu Elde Edilen Tepki Grafiği	34
Şekil 6.5 Eğitim-2 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği	34
Şekil 6.6 Eğitim-2 Sonucu Elde Edilen Hata Grafiği	35
Şekil 6.7 Eğitim-3 Sonucu Elde Edilen Tepki Grafiği	37
Şekil 6.8 Eğitim-3 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği	37
Şekil 6.9 Eğitim-3 Sonucu Elde Edilen Hata Grafiği	38

Şekil 6.10 Eğitim-4 Sonucu Elde Edilen Tepki Grafiği	40
Şekil 6.11 Eğitim-4 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği.....	40
Şekil 6.12 Eğitim-4 Sonucu Elde Edilen Hata Grafiği.....	41
Şekil 6.13 Eğitim-5 Sonucu Elde Edilen Tepki Grafiği	43
Şekil 6.14 Eğitim-5 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği.....	43
Şekil 6.15 Eğitim-5 Sonucu Elde Edilen Hata Grafiği.....	44
Şekil 6.16 Eğitim-6 Sonucu Elde Edilen Tepki Grafiği	46
Şekil 6.17 Eğitim-6 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerler Grafiği ..	46
Şekil 6.18 Eğitim-6 Sonucu Elde Edilen Hata Grafiği.....	47
Şekil 6.19 Eğitim-7 Sonucu Elde Edilen Tepki Grafiği	49
Şekil 6.20 Eğitim-7 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği.....	49
Şekil 6.21 Eğitim-7 Sonucu Elde Edilen Hata Grafiği.....	50
Şekil 6.22 Eğitim-8 Sonucu Elde Edilen Tepki Grafiği	52
Şekil 6.23 Eğitim-8 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği.....	52
Şekil 6.24 Eğitim-8 Sonucu Elde Edilen Hata Grafiği	53
Şekil 6.25 Eğitim-9 Sonucu Elde Edilen Tepki Grafiği	55
Şekil 6.26 Eğitim-9 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği.....	55
Şekil 6.27 Eğitim-9 Sonucu Elde Edilen Hata Grafiği.....	56
Şekil 6.28 Eğitim-10 Sonucu Elde Edilen Tepki Grafiği	58
Şekil 6.29 Eğitim-10 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği	58
Şekil 6.30 Eğitim-10 Sonucu Elde Edilen Hata Grafiği	59
Şekil 6.31 Eğitim-11 Sonucu Elde Edilen Tepki Grafiği	61
Şekil 6.32 Eğitim-11 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği.....	61
Şekil 6.33 Eğitim-11 Sonucu Elde Edilen Hata Grafiği.....	62

SİMGELER VE KISALTMALAR LİSTESİ

BTK	Bilgi Teknolojileri ve İletişim Kurumu
DMO	Dünya Meteoroloji Organizasyonu
DOST	Gelişmiş Bilim ve Teknoloji Enstitüsü
GNSS	Küresel Konum Belirleme Sistemi
GPS	Küresel Konumlama Sistemi
GSR	Gaussian Süreç Regresyon
KAÖ	Kuzey Amerika Orta Ölçekli
Km	Kilometre
KSVM	Kübik SVM
KTS	Küresel Tahmin Sistemi
LSTM	Uzun kısa süreli bellek
mbar	Milibar
MGM	Meteoroloji Genel Müdürlüğü
RMSE	Ortalama Karekök Sapması
SVM	Destek Vektör Makineleri
SVR	Support Vector Machine Regresyon
TEMP	Radyosonde rasat verisi
UÇTM	Ulusal Çevresel Tahmin Merkezleri
UHDM	Ulusal Hava Durumu Merkezi
UHDS	Ulusal Hava Durumu Servisi
UPS	Kesintisiz Güç Kaynakları
UTC	Eşgüdümlü Evrensel Zaman
UTM	Ulusal Tahmin Merkezlerine
YÇHY	Yüksek Çözünürlüklü Hızlı Yenileme

1. GİRİŞ

Belirli bir ülke, bölge veya merkezde, bir zaman dilimi içinde görülebilecek meteorolojik olayların gözlem ve analizlere dayanılarak sübjektif veya objektif yöntemler kullanılarak önceden öngörülme çalışmaları hava durumu tahmini olarak adlandırılır. Hava tahmini, gözlemler analiz ve tahmin olmak üzere üç aşamada belirlenir. Hava durumu atmosferdeki belirli şartların özel yerler ve zamanlarda bir araya gelmesinden oluşur. Hava durumunu sıcaklık, nem, rüzgâr hızı, bulutluluk, yağmur, atmosfer basıncı ve bunun gibi birçok faktör etkilemektedir. Mümkün olduğunca doğru hava durumunu tahmin etmek her zaman meteorologların bir numaralı önceliği olmuştur. Hava durumunun hatasız tahmin edilmesi dünya genelinde çok önemlidir. Hatalı gerçekleştirilen tahminler ülkelere ekonomik olarak milyarlarca liralık zararlar oluşturabilir. Bu nedenle hava durumunu tahmin etmek için matematiksel ve istatistiksel tabanlı denklemler kullanılır. Ancak, küresel ısınmadan dolayı artık hava durumu daha da karışık ve kaotik olmaya başlamıştır. Bu sebepten dolayı matematiksel yöntemler yeterli olmamaktadır. Hâlâ istatistiksel yöntemler kullanılsa da teknolojinin ve yapay zekânın gelişmesiyle bu yöntemler yerini ya hibrit yöntemlere ya da yapay zekâyâ dayanan yöntemlere bırakmıştır. Bu nedenle Makine Öğrenme Algoritmaları ile hava durumu kestirimi önemli bir çalışma konusudur [1].

Dünya genelinde farklı ülkelere ait meteoroloji birimleri dünya ile eş zamanlı olarak meteoroloji balonları ile rasatlar gerçekleştirirler. Bu rasatlar sabah rasatları ve akşam rasatları olarak adlandırılırlar. Gerçekleştirilen bu rasatlarda yerden 40 km irtifaya kadar 1mbar aralıklarla her bir yükseklik noktasına ait sıcaklık, işba sıcaklığı, nem, rüzgâr hızı ve rüzgâr yönü değerleri ölçülür. Ardından gerçekleştirilen bütün bu rasatlar Dünya Meteoroloji Organizasyonu'na gönderilerek 3-boyutlu atmosfer modeli elde edilir. Bu model kullanılarak istatistiksel hava durumu tahmini gerçekleştirilir.

Radyosonde rasatlarının yanı sıra uydu, radar ve otomatik yer gözlem istasyonu ölçümleri de gerçekleştirilmektedir. Ancak, hava durumu kestirimi için farklı irtifalara ait ölçümleri barındırdığından dolayı en baskın rasat türü Radyosonde rasatlarıdır.

Hava durumu kaotik bir durumdur ve bunu basit yöntemlerle kestirmek yanlış sonuçlara neden olur bu sebepten ötürü meteorologlar bu kaotik olayı anlamlandırıp tahmin yürütmek için istatistiksel tabanlı ve yapay zekâ tabanlı denklemler ile çözmeye çalışmışlardır.

Günümüzde yapay zekâ algoritmalarının gelişmesiyle yapay zekâ tabanlı kestirimler istatistiksel tabanlı kestirimlere göre daha keskin cevap vermeye başlamıştır [2].

Yapay Sinir Ağları (ANN) hava tahmini benzeri kaotik sorunları çözmek için uygun yöntemlerden biridir. Bunun sebebi ANN insan beyni çalışmasından esinlenerek gerçekleştirilmiş ve doğrusal olmayan problemlerde de etkin sonuçlar vermektedir [1].

Bu çalışmada Yer Gözlem İstasyonundan alınan sabit irtifalı veriler yerine yüksek irtifa meteoroloji ölçüm cihazı olan Radyosonde ile gerçekleştirilen 3-boyutlu rasat sonuçları kullanılarak Makine Öğrenmesi Algoritmaları eğitilmiştir. Gerçekleştirilen eğitimler sonucunda elde edilen modeller ile ertesi güne ait en yüksek ve en düşük hava sıcaklığı değerleri tahmin edilmiştir.



1.1. Literatür Araştırması

Hava durumu tahmini konusunda daha önce yapılan çalışmalar incelenmiştir. Bu çalışmalar içerisinde bazı çalışmalar hava durumunu kestirmek için matematiksel denklemler kullanırken bazı araştırmacılar ise bu kestirimi makine öğrenmesi ile yapmıştır. Literatür araştırmasında; öncelikle istatistiksel tabanlı çalışmalar, ardından da makine öğrenmesi tabanlı çalışmalar incelenmiştir.

Fente çalışmasında hava durumu verilerini Ulusal Hava Durumu Merkezi'nden (UHDM) toplamıştır [3]. Uzun Kısa Vadeli Bellek (Long Short Term Memory, LSTM) algoritması kullanılarak farklı kombinasyonlarla eğitilip test edilmiştir. Kullandığı veriler UHDM'den alınmış 10 yıllık verilerdir. Bu veriler; sırası ile sıcaklık, yağış, rüzgâr hızı, basınç, sis yoğunluğu ve nem verisinden oluşmuş ve otomatik istasyonlar tarafından toplanmıştır. Fente'nin çalışmasında kestirmeye çalıştığı günlük ortalama sıcaklık değeri ve başarısı %95'tir.

Sobrevilla çalışmasında Tiwi adasındaki hava durumunu tahmin etmek için Esnek Yayılma algoritması ile eğitim yapmıştır [1]. Bu çalışma günlük hava tahminini bulmak içindir. Bu çalışmada kullanılan veriler Gelişmiş Bilim ve Teknoloji Enstitüsü'nden (DOST – ASTI) toplanan 2012-2015 yılları arasındaki ölçümlerdir. Kullandığı verilerde bozulmalar ve boş alanlar olduğundan ilk olarak bu alanları Fourier Uyumlandırma ve Çoklu Atama yöntemleri ile doldurarak eğitim gerçekleştirmiştir. İkinci eğitimde ise bu bozuk satırları tamamen eğitim verisinden çıkararak eğitim gerçekleştirmiştir. Yaptığı bu iki çalışma sayesinde düzeltmek yerine tamamen çıkarmanın daha etkili olduğu sonucuna ulaşmıştır. Bu çalışma 17000 yinelemede 10 gizli nöronla optimal modele ulaşmıştır. Bu çalışmada yüzde 98,96743 günlük hava tahmini için ortalama tahmin doğruluğu elde etmiştir.

Nandar çalışmasında Mynmardaki aylık yağmur oranlarını tahmin etmek için uğraşmıştır. Kullandığı öğrenme algoritması Bayesian Ağ (Likelihood weight) algoritmasıdır [4]. Algoritmanın başarısını da kök ortalama kare hata (Root Mean Square Error, RMSE) başarı kriterini kullanarak hesaplamıştır. Kullandığı veri seti Myanmar'ın 1990 ile 2006 yıllarındaki yağmur yağış verileridir. Veri setini parçalama şekli 1990 ile 2005 arasını eğitim için kullanıp, 2006 yılındaki veriler ile oluşturulan modeli test etmiştir. Yağmur oranlarını kestirmede RMSE oranı %3 ile tahmin etmiştir.

Rahul çalışmasında Ulusal İklim Veri Merkezi'nden yer gözlem istasyonları tarafından ölçülen Kasım 2007'den Ekim 2017'ye kadarki hava durumu verileri ile eğitim

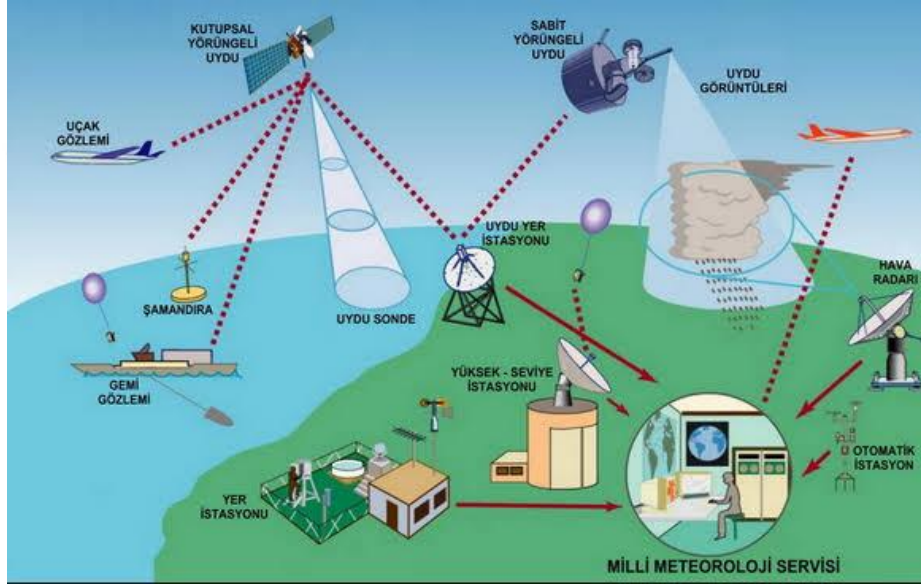
yapmıştır [5]. Veri seti; sıcaklık, nem, çiy noktası, basınç, görünürlük ve yağış gibi birçok hava durumu özelliğinden oluşmaktadır. LSTM tekniği kullanılarak uygun bir doğruluk ile hava durumu tespit edilmek istenmiştir. Hava durumu %84 başarı ile doğru hesaplamıştır.



2. HAVA DURUMU KESTİRİMİ

Her gün, Ulusal Tahmin Merkezlerine (UTM) iletilen binlerce gözlem, bilgisayarlar tarafından haritalandırılır ve yüzey ve hava haritalarında gösterilir [2]. Bu gözlemlerin toplanma şekli çeşitlilik gösterir. Hava tahminlerinde gözlemler ve kullanılan ürünler aşağıdaki gibidir.

- Yer Gözlemleri,
- Gemi Gözlemleri,
- Yüksek Atmosfer Gözlemleri,
- Radar Ürünleri,
- Uydu Görüntüleri,
- Otomatik Meteoroloji İstasyonları Verileri.



Şekil 2.1 Hava Durumu Verisi Toplama Yolları [6]

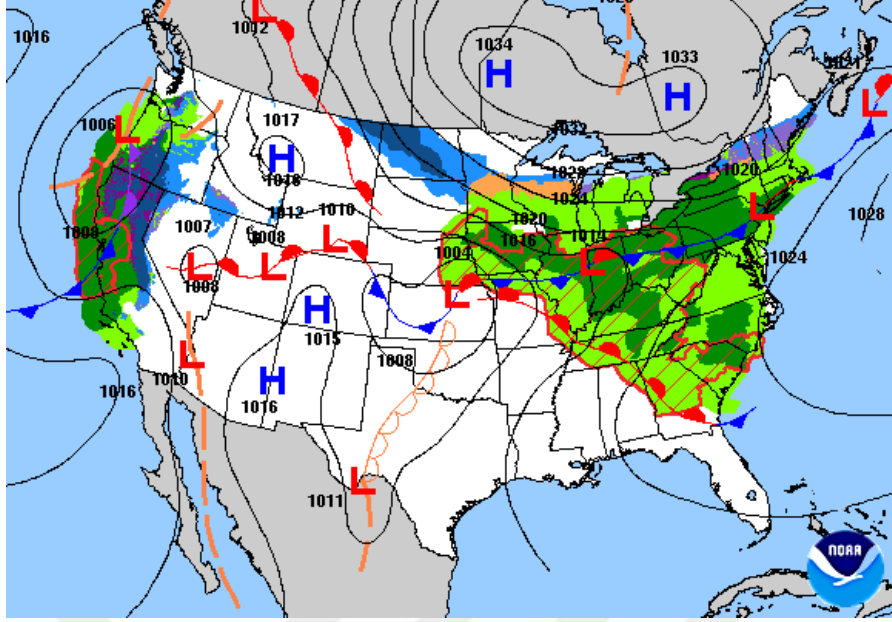
Meteorologlar, ölçüm sonuçlarından elde edilen hava modellerini yorumlar ve ardından da mevcut olabilecek hataları tespit ederler. Ancak bilgisayarlar yalnızca verileri çizip analiz etmekle kalmaz, aynı zamanda hava durumunu da tahmin ederler; bu çok daha zorlu bir görevdir. Günümüzün gelişmiş bilgisayarları, saniyede trilyonlarca hesaplama yaparak büyük miktarda veriyi son derece hızlı bir şekilde analiz edebilir. Atmosfer çok karmaşık olduğu için, dünyadaki en güçlü bilgisayarlardan bazıları hava ve iklim tahminine

adanmıştır. Bilgisayarlar tarafından matematiksel denklemler kullanılarak rutin günlük hava tahmini, sayısal hava tahmini olarak bilinir hâle gelmiştir [7].

Birçok hava değişkeni sürekli değiştiği için, meteorologlar tarafından atmosferin mevcut durumunu tanımlayan atmosferik modeller tasarlanmıştır. Bunlar gelişmekte olan bir fırtınanın resmini çizen fiziksel modeller değildir. Daha çok atmosferik sıcaklığın, basıncın, rüzgârların ve nemin zamanla nasıl değişeceğini tanımlayan birçok denklemden oluşan matematiksel modellerdir. Bu modeller gerçek atmosferi tam olarak temsil etmemektedir. Çünkü dünya çapında her an meydana gelen süreçler, tam olarak temsil edilemeyecek kadar karmaşıktır. Bunun yerine modeller, atmosferin davranışının en önemli yönlerini korumak için formüle edilmiş çok faydalı yaklaşımlardır.

Bu modellerde öncelikle denklemler karmaşık yazılımlara dönüştürülür. Sıcaklık, basınç, nem, rüzgâr ve hava yoğunluğunun yüzey ve üst hava gözlemleri düzenli aralıklarla bu denklemlere girdi olarak verilir. Bu verileri sayısal modellere entegre etme işlemine veri asimilasyonu denir. Modellere giderek daha fazla veri türü asimile edildiğinden, model tahminlerinin kalitesi genellikle iyileşir.

Bu kilit meteorolojik değişkenlerin her birinin nasıl değişeceğini belirlemek için, her bir denklem, her biri belirli bir mesafede bulunan grid noktaları olarak adlandırılan çok sayıdaki konum için gelecekteki 5 dakika için çözülür. Ayrıca, her denklem atmosferde 50 farklı seviye için çözülür. Bu hesaplamaların sonuçları daha sonra orijinal denklemlere verilerek, yeni “verilerle” denklemler yeniden çözülerek, sonraki beş dakika boyunca hava durumu tahmini yapılmaktadır. Bu prosedür, gelecekte istenen bir zamana genellikle 6, 12, 24, 36 ve 84 saate ulaşana kadar tekrar tekrar yapılır. Bilgisayar daha sonra verileri analiz eder ve izobarları veya kontur çizgileri ile basınç sistemlerinin öngörülen konumlarını çizer. Belirli bir gelecek zamanda atmosferi temsil eden nihai tahmin çizelgesine prognostik çizelge veya kısaca prog denir. Bu prognostik çizelgeye örnek Şekil 2.2 görülmektedir. Bilgisayarla çizilen proglar, “makine yapımı” tahminler olarak bilinmektedir [2].



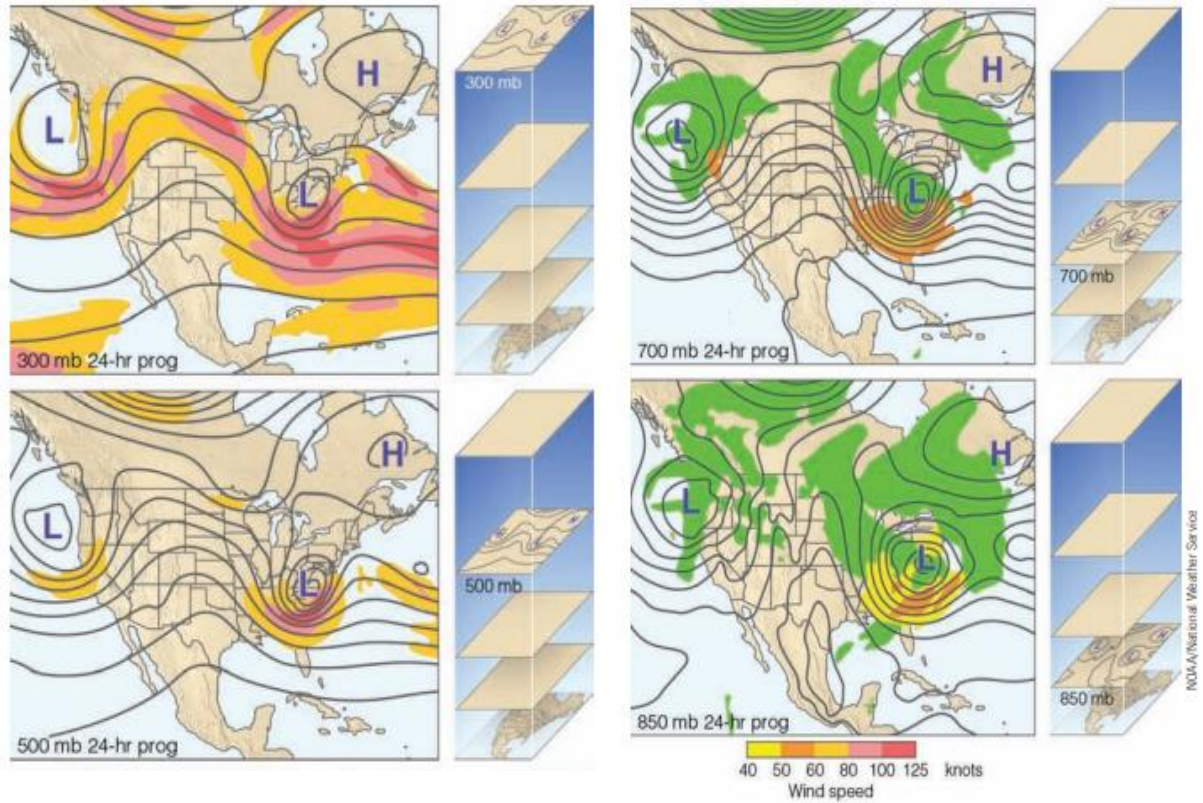
Şekil 2.2 Prognostik Çizelge Örneği [2]

Dünyanın dört bir yanındaki büyük meteoroloji merkezlerinin her biri kendi bilgisayar modellerine sahiptir. Bu merkezler tüm Kuzey veya Güney Yarımkürenin tamamı için genellikle bir hafta olacak şekilde veya daha uzun sürelerle yönelik çeşitli büyük ölçekli model tahminlerini paylaşmaktadır. Ulusal Çevresel Tahmin Merkezleri (UÇTM), birkaç modeli operasyonel modda çalıştırır; bu durumda hava tahminlerinin kapsamlı bir şekilde test edildiklerinin ve günlük kullanım için yeterince güvenilir olduklarını ifade etmektedirler. Diğer modeller deneysel modda çalıştırılır; bu modeller umut vaat etmekte ve tahminler tarafından danışılabilir. Ancak henüz operasyonel kullanım için yeterince uygun olmayabilirler. Günümüzde kullanılan UÇTM operasyonel modelleri; altı saatte bir çalıştırılan ve koşulları üç saatlik aralıklarla 84 saate kadar tahmin eden Kuzey Amerika Orta Ölçekli (KAOÖ) modelini ve her 12 saatte bir çalıştırılan Küresel Tahmin Sistemi (KTS) modelini içerir. 384 saate veya 16 güne kadar uzanan tahminler yayınlamaktadırlar. Yakın zamanda UÇTM'de operasyonel kullanıma dönüştürülen deneysel Yüksek Çözünürlüklü Hızlı Yenileme (YÇHY) modeli, her saat başı çalıştırılarak, 15 saate kadar saatlik zaman dilimlerinde hava durumunu keskin ayrıntılarla tahmin etmek için tasarlanmıştır [2].

Günümüzde kullanılan gelişmiş bilgisayarlar, atmosferik hareket denklemlerini elle yapılabilecekten çok daha hızlı ve verimli bir şekilde çözebilirler. Örneğin, sadece Kuzey Yarımküre için 24 saatlik bir tahmin çizelgesi oluşturmak için yüz milyonlarca matematiksel

hesaplama gerekir. Bu nedenle, tek bir çizelge oluşturmak için bir grup meteoroloğun tam zamanlı olarak çalışarak el ile hesaplamaları çok uzun süreceğinden, hava tahmini hazır olduğunda o günün hava durumu çoktan geçmiş olacaktır. Dolayısıyla bugün bilgisayar modellerinin öneminin daha da arttığı gözlenmektedir. Aslında, bazı durumlarda bir bilgisayar modeli sakın havalarda yüksek ve düşük sıcaklıkları tahmin etmede bir insan kadar başarılı olabilir. Ancak, tahminçiler herhangi bir modelin tahminini müjde gerçeği olarak almamaya dikkat etmelidir. Tahminçiler, tahmin sürecine kendi bilgi ve deneyimlerini katmadan modellere çok fazla güvenirlerse, “Meteorolojik Kanser” olarak adlandırılan hastalığın kurbanı olabilirler. En olağandışı ve tehditkâr hava durumlarında bilgisayarın hesapladığı en iyi tahminler deneyimli bilgili bir meteorolog tarafından gerektiği gibi dikkatli bir şekilde hesaplandığında ortaya çıkar.

Yukarıda da belirtildiği gibi, tahminçilerin artık seçebilecekleri, her biri aynı öngörülen zaman ve atmosferik seviye için havanın biraz farklı bir yorumunu üreten, giderek artan modelleri (ve dolayısıyla program) bulunmaktadır. Şekil 2.3, atmosferdeki farklı seviyeler için 24 saat geleceğe yönelik dört ilerlemeyi göstermektedir.



Şekil 2.3 850mb, 700mb, 500mb ve 300mb için GFS (Global Forecast System) Modelini Kullanan Bilgisayarla Çizilmiş 24 Saatlik Proglar [2]

Belirli bir gelecek zamanda atmosferi temsil eden nihai tahmin çizelgesine prognostik çizelge veya kısaca prog denir. Proglar arasındaki farklar, modellerin denklemleri kullanma biçiminden veya çözünürlük olarak adlandırılan ızgara noktaları arasındaki mesafeden kaynaklanabilir. Bazı modeller, belirli özellikleri diğerlerinden daha iyi tahmin eder. Bir model, üst düzey grafiklerdeki çukurların konumunu tahmin etmede en iyi sonucu verebilirken, bir diğeri yüzey diplerinin konumunu oldukça iyi tahmin edebilir. Daha yüksek çözünürlüğe sahip bir model daha fazla ayrıntı sağlayabilse de bu tür gelişmiş ayrıntılar, modelin daha doğru olduğu anlamına gelmez.

İyi bir tahmin için, her modelin kendine özgü özelliklerine göre tüm progları dikkate almak gerekir. Tahminci; bilgisayardan gelen rehberliğe, hava durumu hakkındaki yorumuna ve belirli tahmin alanı içindeki hava durumunu etkileyen herhangi bir yerel coğrafi özelliğe dayanarak bir tahminde bulunur. Günümüzdeki tahmin modelleri, daha uzun aralıklarla, azalan becerilerle 4 ila 6 gün sonraki hava durumunu oldukça iyi tahmin etmektedir. Modeller, sıcaklık ve jet akışı modellerini tahmin etmede yağıştan daha iyi bir iş yapma eğilimindedir. Bununla birlikte, her zamankinden daha güçlü bilgisayarlar tarafından sağlanan hava durumu tahminindeki tüm modern gelişmelere rağmen, Ulusal Hava Durumu Servisi (UHDS) ve özel şirketler tarafından yapılan tahminler bazen yanlış sonuçlar verebilmektedir. En iyi istatistik tabanlı denklemden bile yüksek bir başarı elde edilmektedir. Bunun sebebi hava durumunun kaotik bir olay olması, doğrusal olmadığı için de mükemmel bir şekilde matematiğe dökülmesinin çok zor olmasındandır. Her sonucun bir uzman tarafından yorumlanıp iyileştirilmesi gerekmektedir. Makine öğrenmesi tabanlı tahminlerle de tahmin başarıları daha yükselmeye başlamıştır. Bunun sebebi olarak makine öğrenmesi yöntemlerinin doğrusal olmayan sistemleri başarılı yorumlamasından kaynaklanmaktadır [2].

3. YÜKSEK İRTİFA METEOROLOJİ ÖLÇÜM CİHAZI RADYOSONDE

Hava tahmininin gerçekleştirilebilmesi için atmosferin 3 boyutlu olarak modellenmesi gerekmektedir. Atmosferin modellenebilmesi için havanın yerden 40 km yüksekliğe kadar her bir yükseklik değeri karşılığında; sıcaklık, bağıl nem, rüzgâr hızı ve rüzgâr yönü bilgilerinin ölçülmesi gerekmektedir. Ancak bu bilgiler kullanılarak atmosferin modellenebilmesi ve bu modelin anlamlı bir şekilde kullanılabilmesi için en az 300 km aralıkla bütün dünyanın hava durumunun aynı anda ölçülmesi gerekmektedir. Ayrıca, bu ölçümlerin günde en az iki kere tekrarlanması gereklidir.

Meteoroloji Genel Müdürlüğü (MGM), bütün dünya ile aynı anda Türkiye’de 9 (Ankara, İstanbul, Diyarbakır, Samsun, Adana, İzmir, Isparta, Erzurum ve Kayseri) farklı noktadan bu ölçümleri gerçekleştirmektedir. Ölçümler öğlen 14:30 ve gece 02:30’da Dünya ülkeleri ile eşzamanlı olarak yapılmaktadır. Şekil 3.1’de MGM tarafından kullanılan bu cihazlar görülmektedir.



a) Vaisala (Finlandiya) [8]



b) Meteo Modem (Fransa) [9]



c) Meisei (Japonya) [10]



d) Graw (Almanya) [11]

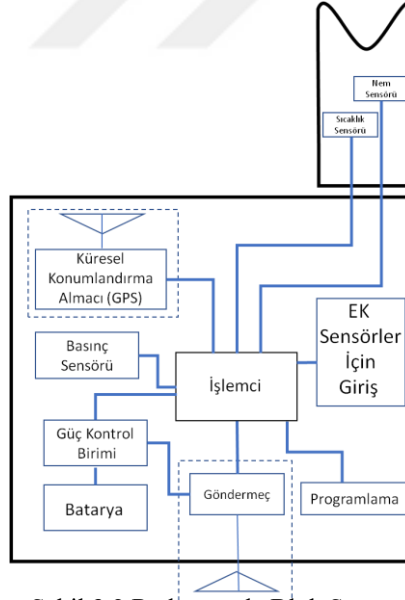
Şekil 3.1 Meteoroloji Genel Müdürlüğü Tarafından Kullanılmakta Olan Yabancı Radyosonde Meteorolojik Ölçüm Cihazları

Radyosonde içerisinde bulundurduğu Küresel Konum Belirleme Sistemi (Global Navigation Satellite Systems, GNSS) sistemi sayesinde 1559 Mhz – 1610 Mhz aralığında yayın yaparak bütün küresel konumlandırma uydularına aynı anda bağlanabilmektedir.

Yükseklik ölçümü GNSS algılayıcısı tarafından ayrı ayrı veya beraber yapılabilmektedir. Bu sayede sinyal takibinin zor olduğu ortamlarda yükseklik bilgisi kopmadan rasata devam edilebilmektedir.

Dahili ısıtıcılı nem sensörü sayesinde bulut içi nemden kaynaklı buzlanmaların önüne geçilebilmekte ve ölçüm doğruluğu artırılmaktadır.

Bilgi Teknolojileri ve İletişim Kurumu (BTK) tarafından meteoroloji uygulamaları için tahsil edilmiş 400 MHz ile 406 MHz frekans aralığında yayın yapılabilmekte ve çıkış gücü 27 dBm'e kadar ayarlanabilmektedir. Şekil 3.2'de Radyosonde blok şeması görülmektedir. Radyosonde elektronik olarak; Elektronik Baskı Devre Kartı, GNSS Alıcısı, Nem Sensörü, Sıcaklık Sensörü, İşlemci, Güç Kontrol Birimi, Batarya, Göndermeç ve Opsiyonel Basınç Sensöründen mekanik olarak ise köpük kutusu, ip sargısı 30 metre ip ve antenden oluşmaktadır.



Şekil 3.2 Radyosonde Blok Şeması

Tablo 1'de Radyosonde üzerinde bulunan sensörler ve bu sensörlere ait Dünya Meteoroloji Organizasyonu tarafından belirlenmiş sahip olması gereken ölçüm aralıkları görülmektedir.

Tablo 1 Radyosonde Özellikleri [12]

Sıcaklık	Ölçüm Aralığı	-95° C ile +60°°C
	Doğruluk	< 0,4° C
	Çözünürlük	0,1° C
	Bildirim Süresi	<0,4s
Nem	Ölçüm Aralığı	1% RH ile 100% RH
	Doğruluk	< 5% RH
	Çözünürlük	0,1 % RH
	Bildirim Süresi	<0,2s
Küresel Konumlandırma Almacı (GPS)	Frekans	1559 MHz ile 1610 MHz
	Bağlanılabilir Uydu Sayısı	72 Kanal
Yükseklik	Ölçüm Aralığı	-500 m ile 50.000 m
	Çözünürlük	0,1 m
Yükseklik Ölçümü	GPS veya Basınç Sensörü (opsiyonel)	
Rüzgâr Yönü	Ölçüm Aralığı	0° ile 360°°C
	Çözünürlük	0,01°
Rüzgâr Hızı	Ölçüm Aralığı	0 m/s ile 200 m/s
	Çözünürlük	0,01 m/s
Göndermeç	Frekans Aralığı	400 MHz ile 406 MHz
	Bant Genişliği	1 kHz
	Çıkış Gücü	27 dBm (Ayarlanabilir)
Modülasyon	Tipleri	GFSK/OOK/MOK/FSK
	Baud Rate	9.600 / 4.800
	Mesafe	250 km
	Örnekleme	0,5 s
Batarya	Pil Tipi	2 x CR123
Ağırlık	110 gr	
Süre	Min. 6 Saat	
Boyut	115(E)x80(B)x55(Y) mm	

GNSS Tabanlı Yer İstasyonu Birimi, rasat sırasında Radyosonde tarafından yer istasyonuna yollanan sinyalleri anlamlandırıp bilgisayarda gözlemlenebilmesi için çeviren ara birimdir; dâhili GPS almacı, 4 adet anten girişi, 2 adet RS-232 portu, 1 adet RS-485 portu, 1 adet 100 Mbit Ethernet portu, Radyosonde programlama çıkışı, güç girişi ve çıkışları, nem sensörü ve sıcaklık sensöründen oluşur. Birim 5 kg'ın altında ağırlığa sahiptir. Gövde malzemesi paslanmaz çeliktir ve üzeri korozyona karşı mat siyah boya ile boyatılmıştır. Ek olarak birim üzerinde ledler bulunur. Bu sayede rasata ait bilgiler birim üzerindeki ledler tarafından da takip edilebilir. Şekil 3.3'de GPS Tabanlı Yer İstasyonu Birimi'ne ait önden ve arkadan görünüş görülmektedir.



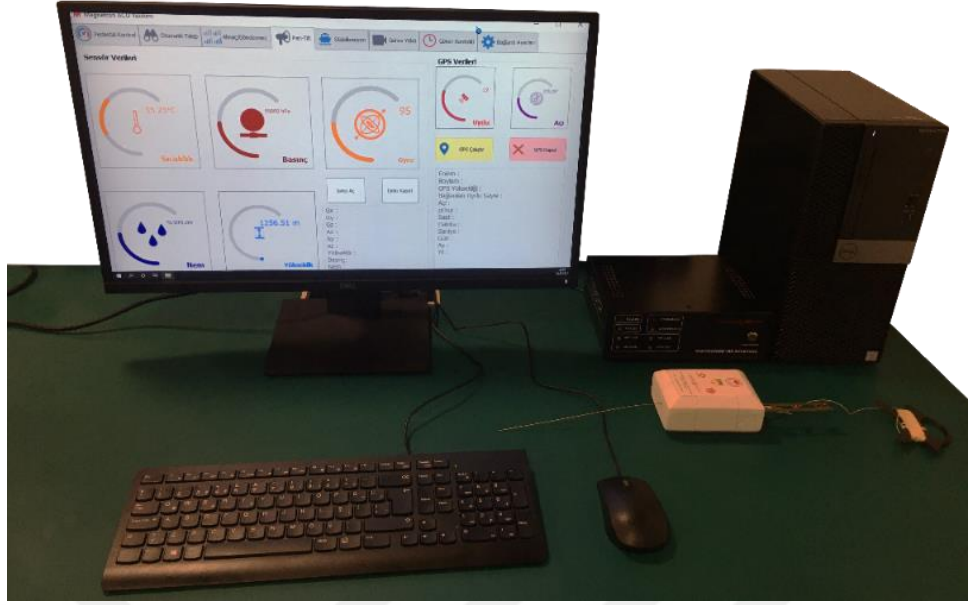
a) Önden Görünüş



b) Arkadan Görünüş

Şekil 3.3 GNSS Tabanlı Yer İstasyonu Birimi Görseli

GNSS Tabanlı Yer İstasyonu Birimi rasat yazılımı yüklü bilgisayar ile ethernetten haberleşir. Rasata başlamadan önce rasat programından gelen veriler doğrultusunda Radyosonde programlanır. Atıştan önce Radyosonde sensörleri ile yer istasyonu sensörleri karşılaştırılır. Eğer büyük bir farklılık varsa operatör uyarılır. Şekil 3.4'te Rasata hazır sistem görülmektedir. Ayrıca sistemle birlikte UPS bulunur. Böylelikle sistem en az 12 saat boyunca aralıksız çalışabilir.



Şekil 3.4 GNSS Tabanlı Yer Alıcı Sistemine Ait Görsel

Sistemde toplam 4 farklı anten bulunur. 3 adet anten Radyosonde verilerini almak için 1 adet anten ise GNSS bilgileri için kullanılır. Antenlerin her biri 5 kg'ın altında ağırlığa sahip ve dış ortam çevre koşullarına dayanabilecek yapıdadırlar. Şekil 3.5'te Radyosonde takip antenleri görülmektedir. Bu antenler yerli ve millî olarak Magnetron tarafından geliştirilmiştir.



a) Yönsüz Dikey Polarize Anten



b) Yönsüz Dairesel Polarize Anten



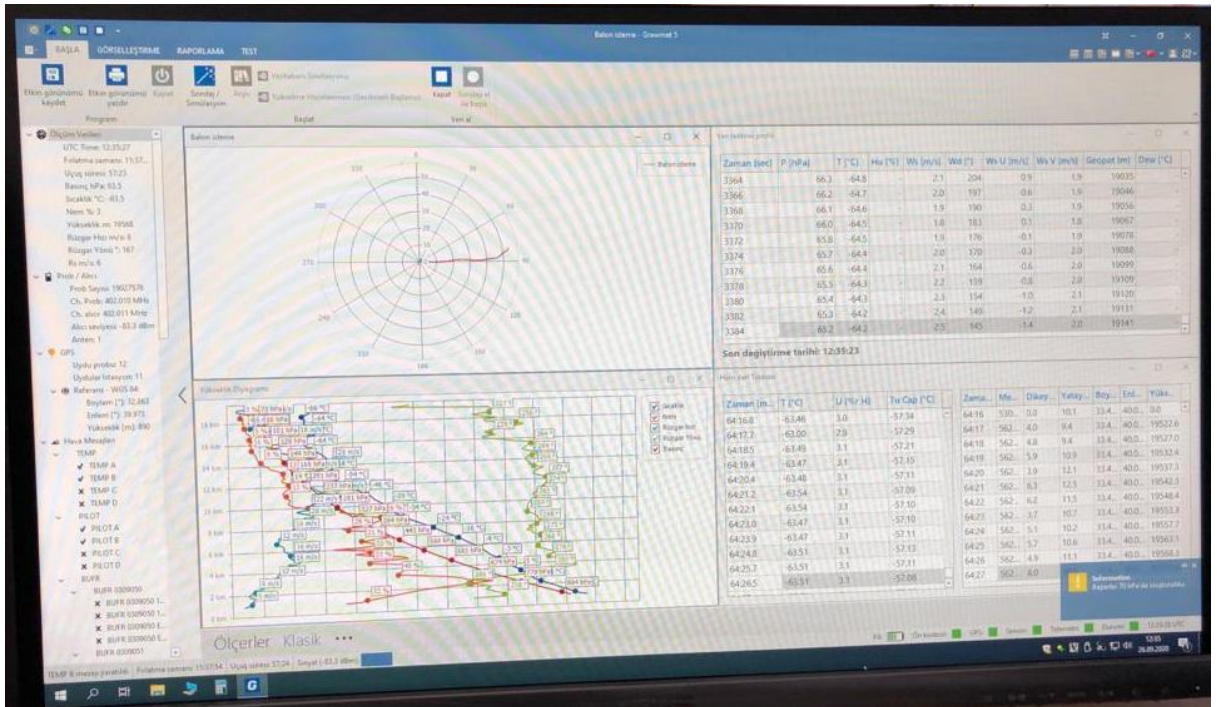
c) Yönlü Dikey Polarize Anten

Şekil 3.5 Yer Alıcı Antenleri

Meteoroloji Mühendisleri ile rasatlar gerçekleştirilmiştir. Şekil 3.6'da MGM 9. Bölge Müdürlüğü ile yapılan incelemeler ve çalışmalar görülmektedir.



a) MGM 9. Bölge Müdürlüğü Radyosonde Rasatı



b) MGM 9. Bölge Müdürlüğü Tarafından Kullanılan Graw Firmasına Ait Rasat Yazılımı

Şekil 3.6 MGM 9. Bölge Müdürlüğü ile Birlikte Yapılan Rasat ve İncelemeler

4. 3 BOYUTLU ATMOSFER VERİ SETİ

Radyosonde rasat verisi TEMP adı verilen özel bir format ile toplanır. Rasat verisinde verileri ayırmak için ayraç olarak “|” işareti kullanılır. Herhangi bir nedenden dolayı anlık veri doğru alınmadı ise o bölge boş bırakılmaz “/” işareti konulur. Veri toplama işlemi balon yükselirken her 1mb aralıklı olacak şekilde toplanır ve Şekil 4.1’de görülmektedir. Her bir satır sırası:

- İstasyon Numarası
- Yıl
- Gün
- Rast Saati
- Referans Saati
- Ölçümün Enlemi
- Ölçümünü Boylamı
- Basın Seviyesi
- Yükseklik
- Sıcaklık
- Nem
- İşba Sıcaklığı
- Rüzgâr Yönü
- Rüzgâr Hızı

Olmak üzere aşağıdaki şekilde tutulur.

```
istno|yıl|ay|gün|rasat_saati|referans_saati|ölçümün_enlemi|ölçümün_boylamı|basınç_seviyesi  
|yükseklik|sıcaklık|nem|işba_sıcaklığı|rüzgar_yönü|rüzgar_hızı
```

Şekil 4.1 Radyosonde Veri Toplama Formatı

Rasat verisi bir balonun ilgili yerden serbest bırakılıp patlayana kadar süredeki her bir 1mb artışında Şekil 4.1’deki paylaşılan veri formatında topladığı hâlden oluşur. Bunu daha iyi açıklamak için örnek bir satır açıklaması aşağıda verilmektedir:

Örnek:

17220|2014|8|5|0|2021-08-04 23:32:01|38.7949|27.0819|1014.6|47|9.2|73|3.8|159|12.3

Şeklinde bir satırı açıklayacak olursak;

17220	: istasyon no
2014	: rasadın ait olduğu yıl
8	: rasadın ait olduğu ay
5	: rasadın ait olduğu gün
0	: rasadın saati (G-T - yalnızca 00 veya 12 olabilir)
2021-08-04 23:32:01	: referans zamanı (balondan kaydın alındığı zaman)
38.7949	: ölçüm alınan enlem derecesi (ondalık derece, WGS 84 standardına göre)
27.0819	: ölçüm alınan boylam derecesi (ondalık derece, WGS 84 standardına göre)
1014.6	: ölçüm alınan basınç seviyesi (mb) (Sadece Yer seviyesinde basınç bir ondalıklı olmalı, üst seviyelerde tamsayı olmalıdır.)
47	: yükseklik (jpm) jeo potansiyel metre
9.2	: sıcaklık (°C cinsinden sıcaklığın bir ondalıklı değeri.)
73	: nem (% ve tamsayı olarak. Gerçek değer %'73'tür.)
3.8	: işba sıcaklığı (°C cinsinden, bir ondalıklı değeri)
159	: rüzgâr yönü (1-360° üzerinden)
12.3	: rüzgâr hızı (knot olarak, bir ondalıklı değeri)

Örnek kayıt:

17220|2014|8|5|0|2014-08-04 23:32:01|38.7949|27.0819|1014.6|47|9.2|73|3.8|159|12.3

17220|2014|8|5|0|2014-08-04 23:33:02|38.7950|27.0818|1014|54|9.1|72|3.7|160|12.1

17220|2014|8|5|0|2014-08-04 23:33:41|38.7951|27.0817|1013|64|9.0|70|3.6|161|12.0

17220|2014|8|5|0|2014-08-04 23:34:03|38.7953|27.0816|1012|71|8.9|70|3.7|160|12.4

4.1 İstasyon Numarası:

Dünya Meteoroloji Organizasyonu (DMO)'nun dünyadaki rasat yapılan belli bölgelere verdiği bir numaradır. Türkiye'de 9 bölge bulunmaktadır. Bunlar;

- SAMSUN, 17030
- İSTANBUL, 17064
- ANKARA, 17130
- İZMİR, 17220
- ISPARTA, 17240
- DİYARBAKIR,17281
- ADANA,17351
- ERZURUM,17095
- KAYSERİ,17196

Bu bölgeler, DMO tarafından Türkiye'yi etkileyen hava kütleleri ve standartlara uygun olarak belirlenmiştir. Ayrıca bu bölgelerin seçilme sebebi dünyanın geoit şeklinden dolayı aralarındaki mesafenin belli bir seviyede olmasıdır. Şekil 4.2'de de görüldüğü gibi dünya yuvarlak olduğu için anten, anteni görerek haberleşme en yüksek 250 km'dir. Bu yüzden Radyosonde atış bölgeleri arası en yüksek mesafenin 500 km olarak belirlenmesi Dünya Meteoroloji Organizasyonu tarafından şart koşulmuştur.



Şekil 4.2 Dünyanın Yuvarlaklığının Yer İstasyonu Antenine Olan Etkisi [13]

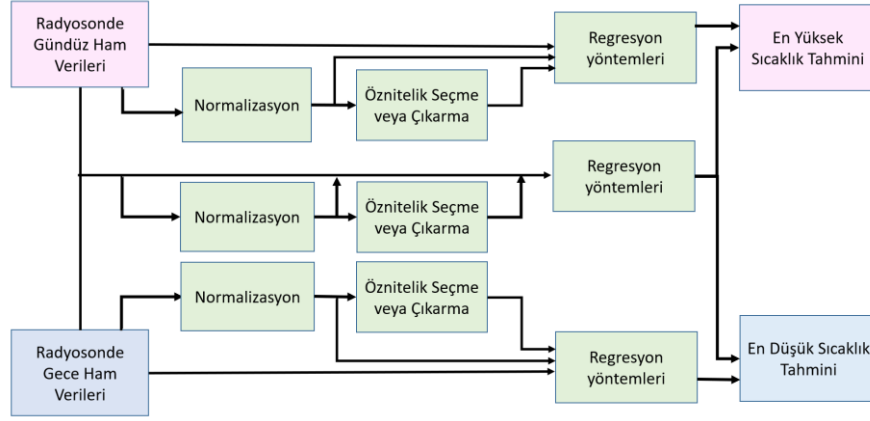
DMO bu istasyonların günde 4 defa rasat yapmasını tavsiye etmektedir. Fakat ülkemizdeki istasyonlar bütün dünya genelinde uygulandığı gibi, 00 UTC ve 12 UTC olmak üzere günde iki defa rasat yapmaktadır. Rasat yapılan bölgeler Türkiye’de 9 adettir ve bu bölgeler Şekil 4.3’de görülebilmektedir.



Şekil 4.3 MGM Bünyesinde Kurulu Bulunan Radyosonde İstasyonları ve İstasyon Numaraları [6]

5. KULLANILAN YÖNTEMLER

Radyosonde verilerini eğitmek için literatürde en çok kullanılan yöntemleri makine öğrenmesi yöntemlerini deneyerek bu eğitim yöntemlerinin kaotik olan Radyosonde rasatları üzerindeki başarıları incelenmiştir. Bunlar Regresyon, Karar Ağacı Algoritması, Destek Vektör Makinesi Algoritması ve Gauss Süreci Regresyonu yöntemleridir. Şekil 5.1’de bu tez çalışmasında uygulanan yöntemlerin blok diyagramı görülmektedir.



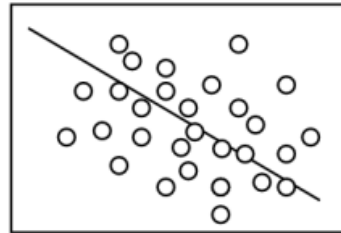
Şekil 5.1 Radyosonde Verilerine Uygulanan Yöntemlerin Blok Diyagramı

5.1 Regresyon:

Regresyon, değişkenler arasındaki ilişkiyi tanımlamak için kullanılan istatistiksel bir analiz yöntemidir. İlişki bağımlı ve bağımsız değişkenler arasında tanımlanabilir.

$$Y = f(X, \beta) \quad (1)$$

Denklem 1’de Y: bağımlı bir değişken, X: bağımsız bir değişkendir. B: bilinmeyen bir parametredir. Şekil 5.2’de regresyon örneği görülmektedir.



Şekil 5.2 Örnek Bir Regresyon Yöntemi

Değişken bağımlılık, tek değişkenli veya çok değişkenli regresyon şeklinde olabilir. Tek değişkenli regresyon denklemi aşağıdaki gibidir.

$$y = a + bx + \varepsilon \quad (2)$$

Denklem 2’de y: bağımlı bir değişken iken x ise b katsayısının katı olan bağımsız değişkendir, a ise bir sabit sayıdır. Çok değişkenli regresyon ise aynı anda birkaç değişken arasındaki bağımlılığı belirlemek için de kullanılır. Denklem 2 ifadesi çok değişkenli olunca denklem 3’e dönüşür.

$$y = a + b_1x + b_2x + b_3x + \dots + b_nx + \varepsilon \quad (3)$$

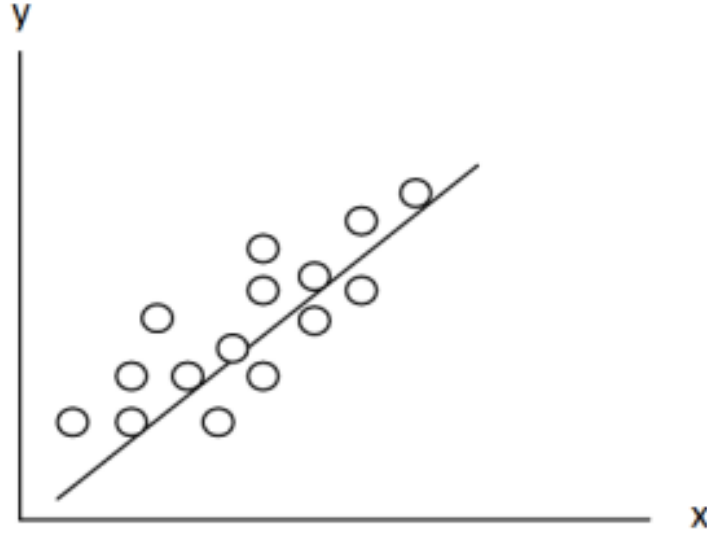
Regresyon modelleri, bağımlı değişkenlerin sonucunu bağımsız değişkenlerden tahmin eder. Karmaşık problemlerin üstesinden gelmek için regresyon analizi kullanılan bir yöntemdir [14].

5.1.1. Doğrusal Regresyon (Linear Regression):

Doğrusal Regresyon, değişkenler arasındaki ilişkiyi tanımlamak için en çok kullanılan tahmin modelidir. Lineer regresyon, basit lineer veya çoklu lineer regresyon olmak üzere ikiye ayrılır.

$$y = x\beta + \varepsilon \quad (4)$$

Denklem 4’te y: devamlı veya kategorik değer alabilen bağımsız bir değişkendir, x: her zaman sürekli bir değer alan bağımlı bir değişkendir. Olasılık dağılımı ile analiz edilir ve esas olarak çok değişkenli analiz ile koşullu olasılık dağılımına odaklanır. Şekil 5.3’te temsil edilen basit doğrusal regresyon, Denklem 2’de açıklandığı gibi tek değişkenli regresyon analizi olan tek bağımsız değişken kullanan tahmin sürecidir. Basit doğrusal regresyon, iki değişken arasındaki ilişkiyi korelasyona benzer şekilde genişletmek için bağımlı değişkenleri ve bağımsız değişkenleri ayırır ancak korelasyon bağımlı ve bağımsız değişkenleri ayırmaz [14].



Şekil 5.3 Örnek Bir Doğrusal Regresyon Grafiği

5.2 Karar Ağacı Algoritması:

Karar ağaçları algoritmaları genelde veri madenciliğinde kullanılır. Amaç girdi değişkeni ile uygun bir hedef değerini elde eden bir model çıkararak kestirim yapmaktır. Algoritma bir ağaç gibidir ve bölgelerden oluşur bu bölgeler; kök, dallar ve yapraklardır.

Ağacın her yaprağı bir sınıfla veya sınıflar üzerinde bir olasılık dağılımıyla hesaplanır. Bunun için entropi kullanılır. En belirgin özelliği belirlemek için entropi hesaplanır ve hesaplar arasında en yüksek olan bir alt dal olarak seçilir.

$$E = \sum_{i=1}^n p(x_i) \log_2\left(\frac{1}{p(x_i)}\right) - \sum_{i=1}^n p(x_i) \log_2(p(x_i)) \quad (5)$$

Denklemden E: entropiyi, $p(x_i)$: olasılık ağırlık fonksiyonunu belirtir. Karar Ağacı Algoritmaları Gürültüye karşı dayanıklı algoritmalarıdır.

5.2.1 Regresyon Karar Ağacı Analizi:

Regresyon karar ağacı analizi ise; tahmin edilen sonuç bir grup yerine gerçek bir sayı olarak kabul edildiği zamandır. (Örneğin, yarın hava kaç derece olacak, bir evin fiyatı veya bir hastanın hastanede kalış süresi).

5.3 Destek Vektör Makinesi:

Destek Vektör Makineleri (Support Vector Machine, SVM), Bell Laboratuvarlarında geliştirilmiş bir makine öğrenme algoritmasıdır. İlk denemeler, test örneklerinin ikili sınıflandırmasına ve örüntü tanımaya odaklanmıştır. Genel olarak, SVM'ler, giriş vektörlerini sonsuz boyutlu bir öznitelik uzayına eşleyerek bir karar alanı oluşturmak için bir algoritma kullanır. Sonrasında, sonsuz veya yüksek boyutlu özellik uzayında doğrusal bir regresyon yürütülür. Çoğu zaman, çok boyutlu girdi vektörü x ile çıktı y arasındaki ilişki bilinmediği ve çoğunlukla doğrusal olmadığı için bu eşleme işlemi gereklidir. Support Vector Machine Regresyon (SVR), çok boyutlu girdi vektörlerini çıktı değerlerine uyduran doğrusal bir hiperdüzlem bulmayı amaçlar. Sonuç daha sonra bir test setinde yer alan gelecekteki çıktı değerlerini tahmin etmek için kullanılır. $X(i)$ veri noktasının giriş vektörü (i), $a(i)$ gerçek değer ve n veri noktası sayısı ile $P = (x_i, a_i)$, $i = 1, \dots, n$ bir dizi veri noktası tanımlar. Doğrusal fonksiyonlar f için, SVR tarafından oluşturulan hiper düzlem aşağıdaki gibi belirlenir [15]. Denklem (6) tahmin edilen değeri hesaplamak için kullanılır.

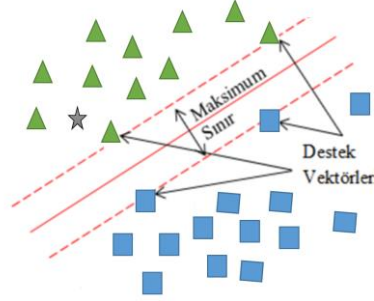
$$f(x) = wx + b \quad (6)$$

Denklem (6)'da doğrusal bir regresyon modeline benzerlikler gösterir. Tahmin edilen değer, $f(x)$, bir w eğimine ve bir b kesişmesine bağlıdır. Genel olarak, iyi bir genelleme davranışını sürdürürken girdiler ve çıktılar arasındaki ilişkiyi öğrenmek arasında bir denge kurmak istenir. Eğitim hatalarını en aza indirmeye çalışmak aşırı odaklanma ve ezberlemeye neden olabilir. Güven aralığı, öğrenme modelinin karmaşıklığı olarak düşünülebilecek olan SVM'nin Vapnik-Chervonenkis boyutuyla ilgilidir. Bu nedenle, ek eğitim hataları pahasına güven aralığı geliştirilerek iyileştirilmiş genelleme elde edilebilir[16]. Bu takası kontrol etmek için birincil araç önemini açıklayan ceza parametresi C 'dir. İyi eğitim ve genelleme davranışı arasındaki denge Denklem' 7'de yansıtılır.

$$R = \frac{C}{n} \sum_{i=1}^n L_{\epsilon}(a_i, f(x_i)) + \frac{1}{2} w^2 \quad (7)$$

Burada R , eğitim hatalarından ve model karmaşıklığından kaynaklanan bileşik riski ifade eder. Doğal olarak, R riskinin mümkün olduğu kadar düşük tutulması gerekir. Şekil

5.4' te görüldüğü üzere SVM'ye ait iki sınıfın birbirinden ayrılmasına örnek gösterilmiştir. SVM her iki sınıfın da bulunduğu en uç noktaları destek vektörü olarak atar. Bu destek vektörleri ile iki sınıf arasındaki uzaklık en büyük olacak şekilde bir hiper düzlem belirlemeye çalışır. Eklenen her yeni veri bu hiper düzleme göre sınıflandırılır.



Şekil 5.4 SVM Sınıflandırma Örneği

Verileri doğrusal olmayan bir şekilde ayırmak için ikili bir optimizasyon formu ve Çekirdek Hilesi (kernel trick) kullanılmaktadır. Geometrik marjının, karar fonksiyonunun gradyanının normunun tersine eşit olduğundan, fonksiyonel marj, ikinci dereceden fonksiyonun denklemidir. Kernel SVM'nin ikinci dereceden bir optimizasyon ayarına yerleştirildiği kanıtlanmıştır. Bu ayar, ikili formun kullanılmasını veya Çekirdek Hilesi'nin kullanılmasını gerektirmez [17].

Kübik SVM sınıflandırma tekniği, düşük bellek alanı ikilemiyle karşılaşıldığında yararlıdır. SVM, çok boyutlu uzayda sınıfları mümkün olan en iyi şekilde bölen bir hiper düzlem bulur.

$$K(x_i, x_j) = (x_i^T x_j)^3 \quad (8)$$

Denklem 8'de sınıflandırıcının çekirdek fonksiyonunun $K(x_i, x_j)$ olarak verildiği kübik SVM tipi sınıflandırıcı kullanılır [18].

Fine Gaussian yönteminin örüntü tanım için güvenilir olduğu kanıtlanmış ve olasılık dağılımı hesaplamaktadır.

$$G(p, q) = \frac{\exp(-(p-q)'(p-q))}{2\sigma^2} \quad (9)$$

Denklem 9, σ : varyansını temsil eder. Örüntü tanıma için güvenilir bir eğitim olarak ilan edilmiştir ve olasılık dağılımını hesaplar. Yalnızca eğitim setindeki verilerin değil, test setinin verilerinin de yanlış sınıflandırılma riskini azaltır. Tercihen, performansı artırma ihtiyacı dışında mükemmel genel performansı nedeniyle tavsiye edilir. İkili sınıfların uygulanması SVM aracılığıyla gerçekleştirilir. Genellikle birçok sınıflı yöntem için görüntü tanıma ve nesne algılama durumunda zayıf kalır ve farklı sınıflarla işlem yapılması ihtiyacı doğar. Bugüne kadar nerdeyse tüm yöntemleri kapsayan bir ayırım bulunamamıştır [19]. SVM’yi fonksiyon regresyon tahmininde uygulamak, konveks ikinci dereceden bir programlamayı çözdüğü düşünülen ve teorik olarak global optimal çözümü elde edebilen SVM regresyon sorusunu oluşturur [20]. Algoritma, doğrusal olmayan dönüştürme yoluyla asıl sorunu yüksek boyutlu uzaya dönüştürür ve orijinal uzayda doğrusal olmayan gerilemeyi gerçekleştirmek için yüksek boyutlu uzayda doğrusal gerilemeyi sürdürür. Algoritma doğası, SVM regresyon modelinin iyi bir genelleme yeteneğine sahip olmasını garanti etmiştir. SVM regresyon algoritması, eğitim örneğinin kısmına göre regresyon fonksiyonunu yapılandırır.

5.4 Gauss Süreci Regresyonu:

Denetimli öğrenme problemleri sınıflandırma ve regresyon olarak iki ana kategoriye ayrılabilir. Sınıflandırma çıktıları ayrık etiketler iken, regresyon sürekli niceliklerin tahmini ile ilgilidir. Gauss süreci, herhangi bir sonlu alt kümesinin ortak bir Gauss dağılımına sahip olduğu rastgele değişkenlerin bir koleksiyonudur[21]. Gauss süreci, tamamen ortalama ve kovaryans fonksiyonlarıyla belirtilir.

$$m(x) = E[f(x)] \quad (10)$$

$$k(x, x') = E[(f(x) - m(x))(f(x') - m(x')))] \quad (11)$$

Denklem 10 ve Denklem 11’de $f(x)$: fonksiyonu $m(x)$: sürecinin ortalama fonksiyonunu, $k(x, x')$: kovaryans fonksiyonunu belirtir. Gauss süreci aşağıdaki şekilde gösterilir:

$$f(X) \sim GP(m(X), k(X, X')) \quad (12)$$

Denklem 12’de GP : gauss proses fonksiyonu belirtmiştir.

Tahmini dağılım, Gauss varsayımına ve Gauss sürecinin sıra dışı özelliğine dayanarak türetilir; bu, ortak dağılımın Gauss olduğunu ima eder.

Gauss süreçleri regresyonu, bir kovaryans fonksiyonu seçimini gerektirir. Kovaryans fonksiyonu ve bazı parametreleri, öykü hakkındaki ön bilgileri analize yansıtmanın bir yoludur [8].

$$k(X, X') = \sigma_f^2 \exp \left[\frac{-(x-x')^2}{2l^2} \right] \quad (13)$$

Denklem 13 ‘de $k(x, x')$: kovaryans fonksiyonunu, σ : ise varyansı belirtir. Varyansın kare üstel kovaryans fonksiyonudur ve yapısı yakın noktalara büyük önem verir. Böylece düzgün fonksiyon benzer yerel davranışa sahiptir, burada l : karakteristik uzunluk ölçeğini tanımlar.

6. DENEYSEL SONUÇLAR VE TARTIŞMA

Bu çalışmada kullanılan veriler 17130 numaralı Ankara bölgesinde bulunan istasyon tarafından ölçülen 3 yıllık (2016-2017-2018) sabah ve akşam atılan rasat verilerinden oluşmaktadır. Bu veriler 40km irtifaya kadar toplanan basınç seviyesi, yükseklik, sıcaklık, nem, işba sıcaklığı, rüzgâr yönü ve rüzgâr hızından oluşmaktadır. Ancak gerçekleştirilen rasatlarda rasat balonu farklı irtifalarda atmosferik etkilerden dolayı patladığı için bütün rasatlar farklı irtifalarda sonlanmaktadır. Bu sebepten dolayı 95mbar'a kadar gerçekleştirilen ölçümler hava tahmininde kullanılmıştır. 95mbar'ın altında kalan veriler veri setine dâhil edilmemiştir. Ayrıca sinyal karıştırıcı ve farklı sebeplerden dolayı kullanıma uygun olmayan rasatlarda veri setine dâhil edilmemiştir.

Veri seti, 3 yıllık ölçüm sonucunda elde edilen 944 ölçümden oluşmaktadır. Bu veri setini eğitim ve test olarak ayırmak için çapraz doğrulama uygulanmıştır. K çaprazlama doğrulamadaki k değeri 5, 10 ve 15 olarak üç farklı değer seçilip bu üç değerde eğitimler yapılmıştır. Gece ve gündüz sıcaklık tahminleri için iki ayrı model oluşturulmuştur. En düşük sıcaklık tahmini için 00.00 UTC de toplanan veriler kullanılırken, En yüksek sıcaklık tahmini için 12.00 UTC'de toplanan veriler kullanılmıştır. Yöntemlerin başarısını analiz etmek için farklı eğitim setleri kullanarak başarıya etkisi analiz edilmiştir.

Veri setinde yükseklik ölçümleri ve her yükseklikte ölçülen sıcaklık ölçümleri, nem ölçümleri, işba sıcaklık ölçümleri, rüzgâr yönü ölçümleri, rüzgâr hızı ölçümleri bulunmaktadır.

Bu tez kapsamında, başlangıç olarak eğitimlerde üç yıllık sabah ve akşam toplanmış rasat verileri kullanılarak gün içerisindeki en yüksek ve en düşük sıcaklık kestirimi gerçekleştirilmiştir. Ardından rasat verilerine öz nitelik çıkarımı yöntemi uygulanmıştır. Daha sonra veri seti yerden farklı irtifalara kadar ayrıştırılarak eğitimler gerçekleştirilmiştir. Son olarak veri seti sabah ve akşam rasatları olarak ayrıştırılmış ve sabah veri seti kullanılarak en yüksek hava sıcaklığı, akşam veri seti kullanılarak ise en düşük hava sıcaklığı tahmin edilmiştir.

Her eğitimde 16 farklı makine öğrenme algoritması K=5, K=10 ve K=15 değerleri için ayrı ayrı eğitilmiştir. Eğitimlerin özet tablosu Tablo 2'de verilmektedir.

Tablo 2 Gerçekleştirilen Eğitim Sonuçları Özet Tablosu

	Veriler	Parametreler
Eğitim 1	3 yıllık gündüz ve akşam rasat verileri. Veriler normalize edilmiştir.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.
Eğitim 2	3 yıllık gündüz ve akşam rasat verileri. Veriler normalize edilmiştir.	PCA değerinin başarıya etkileri incelenmiştir. PCA 97 de en iyisidir.
Eğitim 3	3 yıllık gündüz ve akşam rasat verileri. Veriler normalize edilmemiştir.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.
Eğitim 4	3 yıllık gündüz ve akşam rasat verileri. Veriler normalize edilmemiştir.	PCA değerinin başarıya etkileri incelenmiştir. PCA 97 de en iyisidir.
Eğitim 5	3 yıllık sabah ve akşam rasat verileri kullanılmıştır. Veriler normalize edilmiştir. Bu rasat verileri yer seviyesinden ilk 400mbar kadar olan bölge alınmıştır.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.
Eğitim 6	3 yıllık sabah ve akşam rasat verileri kullanılmıştır. Veriler normalize edilmiştir. Bu rasat verileri yer seviyesinden ilk 500mbar kadar olan bölge alınmıştır.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.
Eğitim 7	3 yıllık sabah ve akşam rasat verileri kullanılmıştır. Veriler normalize edilmiştir. Bu rasat verileri yer seviyesinden ilk 600mbar kadar olan bölge alınmıştır.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.
Eğitim 8	3 yıllık sabah datası ile en yüksek sıcaklık kestirimi yapılmıştır. Veriler normalize edilmiştir.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.
Eğitim 9	3 yıllık sabah datası ile en yüksek sıcaklık kestirimi yapılmıştır. Veriler normalize edilmemiştir.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.
Eğitim 10	3 yıllık akşam datası ile en yüksek sıcaklık kestirimi yapılmıştır. Veriler normalize edilmiştir.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.
Eğitim 11	3 yıllık akşam datası ile en yüksek sıcaklık kestirimi yapılmıştır. Veriler normalize edilmemiştir.	K Çaprazlama Değeri 10 Seçilmiştir. PCA yok.

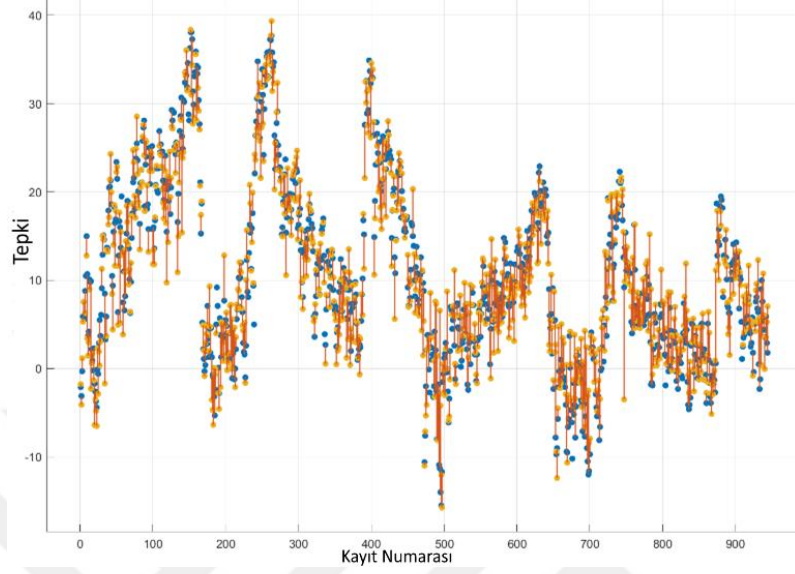
6.1 Eđitim-1:

Bu eđitimde, üç yıllık sabah ve akřam toplanmıř rasat verileri kullanılarak gn iindeki en yksek ve en dřk sıcaklık kestirimi yapılmıřtır. Kullanılan veriler normalize edilmiř verilerdir. Tablo 3'te derece cinsinden elde edilen RMSE deęerleri verilmektedir. Farklı regresyon algoritmalarından en bařarılı sonu Gaussian Sre Regresyon (GSR) yntemi ile elde edilmiřtir. GSR parametrik bir algoritma olmadıęı iin, ok az veri mevcut olduęunda bile ok eřitli denetimli ęrenme problemlerinde etkili sonular retildięi gzlenmiřtir [22]. Tm veriler kullanılarak aynı anda en yksek ve en dřk sıcaklık tahmini yapılmıřtır. Buradaki sıcaklık tahmin problemi doęrusal olmadıęı iin en bařarısız sonu Doęrusal Regresyon modeli ile elde edilmiřtir.

Tablo 3 Eğitim-1 Sonucu Elde Edilen Başarım Sonuçları Tablosu

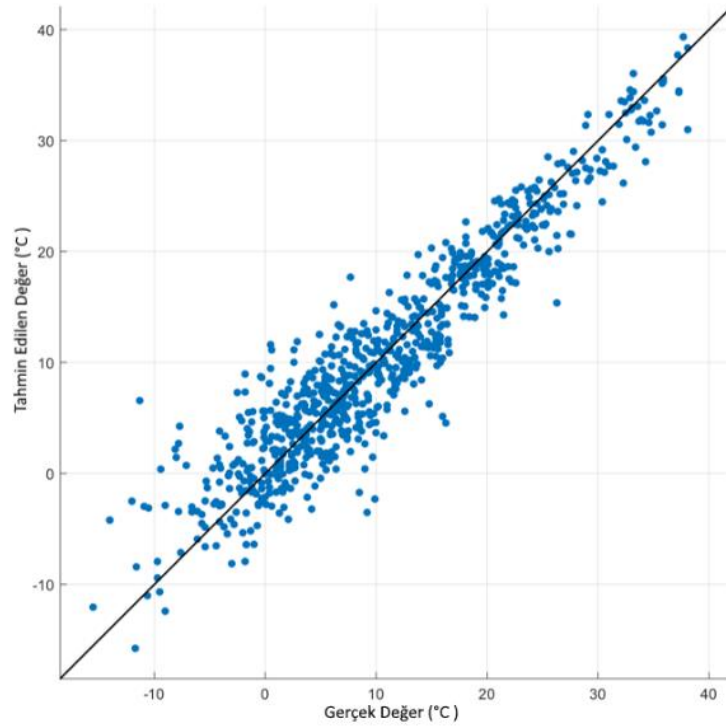
	K=5	K=10	K=15
Algoritma Adı	Başarı (RMSE)	Başarı (RMSE)	Başarı (RMSE)
Doğrusal Regresyon (Doğrusal) / Linear Regression (Linear)	47,965	61,668	64,166
Ağaç (İnce Ağaç) / Tree (Fine Tree)	6,5676	6,4842	6,4026
Ağaç (Orta Ağaç) / Tree (Medium Tree)	6,2353	6,1863	6,2282
Ağaç (Kaba Ağaç) / Tree (Coarse Tree)	6,3356	6,1135	6,1089
SVM (Doğrusal SVM) / SVM (Linear SVM)	4,6946	4,5686	4,5325
SVM (Kuadratik SVM) / SVM (Quadratic SVM)	3,8679	3,9542	4,0064
SVM (Kübik SVM) / SVM (Cubic SVM)	6,4365	5,2164	5,1027
SVM (İnce Gauss SVM) / SVM (Fine Gaussian SVM)	10,285	10,275	10,277
SVM (Orta Gauss SVM) / SVM (Medium Gaussian SVM)	4,3089	4,2873	4,2494
SVM (Kaba Gauss SVM) / SVM (Coarse Gaussian SVM)	4,9040	4,8585	4,7623
Topluluk (Güçlendirilmiş Ağaçlar) / Ensemble (Boosted Trees)	4,6998	4,6574	4,6220
Topluluk (Torbalı Ağaçlar) / Ensemble (Bagged Trees)	4,7730	4,7457	4,6929
Gauss Süreç Regresyonu (Kareli Üstel GPR) / Gaussian Process Regression (Squared Exponential GPR)	10,290	10,280	10,283
Gauss Süreç Regresyonu (Anne 5/2 GPR) / Gaussian Process Regression (Matern 5/2 GPR)	10,262	10,249	10,249
Gauss Süreç Regresyonu (Üstel GPR) / Gaussian Process Regression (Exponential GPR)	8,6169	8,2138	8,8522
Gauss Süreç Regresyonu (Rasyonel Kuadratik GPR) / Gaussian Process Regression (Rational Quadratic GPR)	3,5269	3,5194	3,5101

Şekil 6.1’de Eğitim-1 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait tepki grafiği görülmektedir.



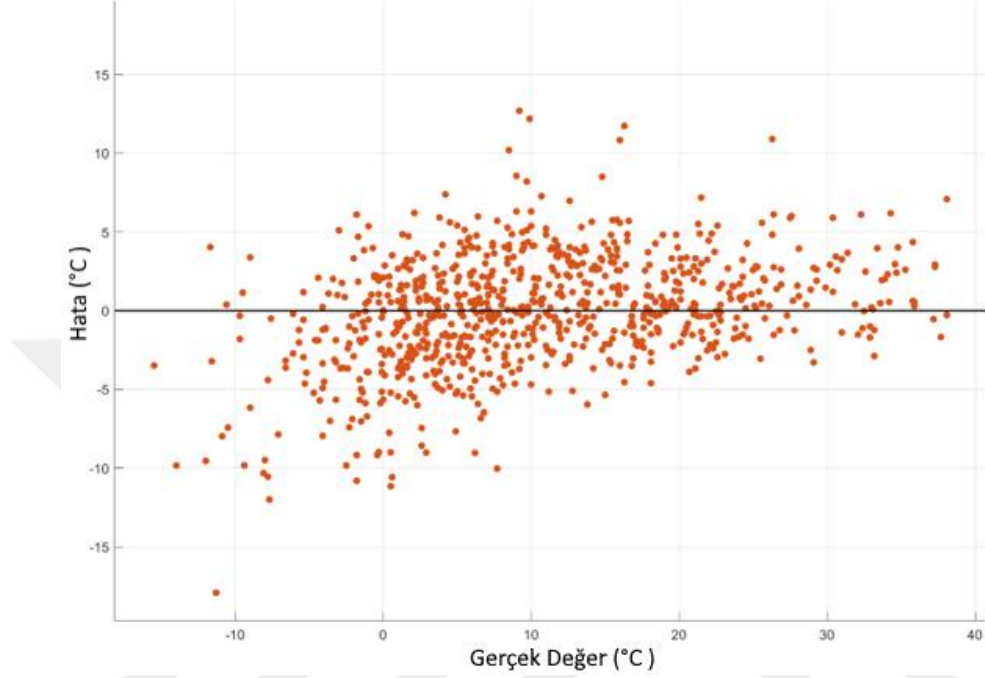
Şekil 6.1 Eğitim-1 Sonucu Elde Edilen Tepki Grafiği

Şekil 6.2’te Eğitim-1 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait sapma grafiği görülmektedir.



Şekil 6.2 Eğitim-1 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği

Şekil 6.3'te Eğitim-1 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait hata grafiği görülmektedir.



Şekil 6.3 Eğitim-1 Sonucu Elde Edilen Hata Grafiği

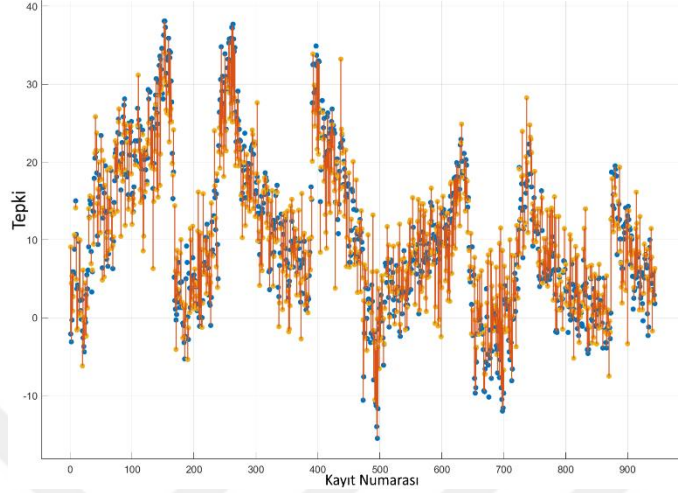
6.2 Eğitim-2:

Bu eğitimde, üç yıllık sabah ve akşam toplanmış rasat verileri kullanılarak gün içindeki en yüksek ve en düşük sıcaklık kestirimi yapılmıştır. Eğitim 2’de, Eğitim 1’den farklı olarak kullanılan ham veriden PCA yöntemi ile öznelik çıkartılmış ve temel bileşen toplam varyans değerleri değiştirilerek başarıya etkisi incelenmiştir. K Çaprazlama değeri 15 olarak seçilmiştir. Kullanılan veriler normalize edilmiş verilerdir. Tablo 4’de derece cinsinden elde edilen RMSE değerleri verilmektedir. Temel bileşen analizi, verilerin toplam varyansının farklı değerleri için verilere uygulanmış ve farklı regresyon analiz yöntemleri karşılaştırılmıştır. Tablo 4 incelendiğinde en başarılı yöntemin Gauss Süreç Regresyonu (Rational Quadratic GSR) olduğu görülmektedir.

Tablo 4 Eğitim-2 Sonucu Elde Edilen Başarım Sonuçları Tablosu

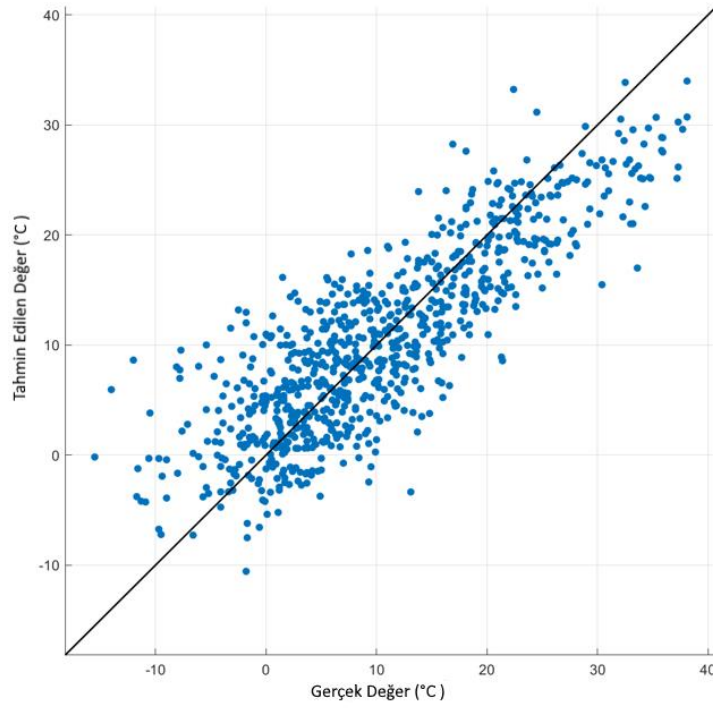
	PCA=90	PCA=95	PCA=97
Algoritma Adı	Başarı (RMSE)	Başarı (RMSE)	Başarı (RMSE)
Linear Regression (Linear)	6,4158	6,0539	5,8841
Tree (Fine Tree)	9,6774	9,9387	9,9749
Tree (Medium Tree)	8,9799	9,0726	9,2128
Tree (Coarse Tree)	8,6421	8,6351	8,7520
SVM (Linear SVM)	6,4722	6,0778	6,0349
SVM (Quadratic SVM)	6,0970	5,7496	5,6702
SVM (Cubic SVM)	6,3490	5,8338	5,7714
SVM (Fine Gaussian SVM)	6,4438	6,0551	5,9834
SVM (Medium Gaussian SVM)	8,9350	8,7782	8,7450
SVM (Coarse Gaussian SVM)	10,307	10,298	10,295
Ensemble (Boosted Trees)	7,5080	7,5486	7,6037
Ensemble (Bagged Trees)	7,6163	7,7289	7,8594
Gaussian Process Regression (Squared Exponential GPR)	5,9127	5,6358	5,5601
Gaussian Process Regression (Matern 5/2 GPR)	76257	8,5834	9,6936
Gaussian Process Regression (Exponential GPR)	8,0597	8,3906	8,9674
Gaussian Process Regression (Rational Quadratic GPR)	5,8851	5,6102	5,5603

Şekil 6.4'te Eğitim-2 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait tepki grafiği görülmektedir.



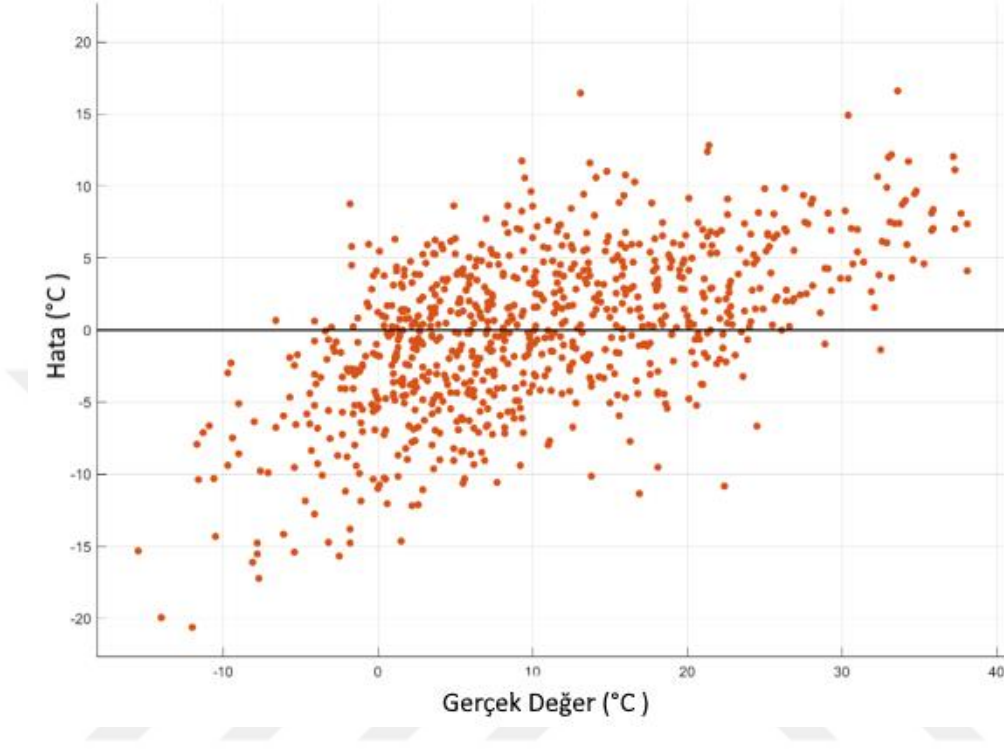
Şekil 6.4 Eğitim-2 Sonucu Elde Edilen Tepki Grafiği

Şekil 6.5'te Eğitim-2 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait sapma grafiği görülmektedir.



Şekil 6.5 Eğitim-2 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği

Şekil 6.6'da Eğitim-2 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait hata grafiği görülmektedir.



Şekil 6.6 Eğitim-2 Sonucu Elde Edilen Hata Grafiği

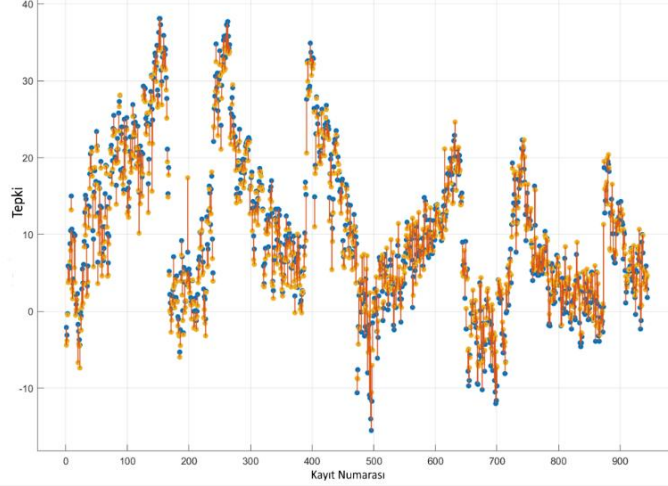
6.3 Eğitim-3:

Bu eğitimde, üç yıllık sabah ve akşam verileri toplanmış rasat verileri kullanılarak gün içindeki en yüksek ve en düşük sıcaklık kestirimi yapılmıştır. Bu eğitimde farklı olarak Eğitim 1’de ki ham veriler normalize edilmeden kullanılmıştır. Kullanılan veriler normalize edilmeden uygulandığında yöntemlerin başarı değerlerinin değişimi Tablo 5’te gözlemlenmektedir. Veriler normalize edilmediği için eğitimde en baskın veriler basınç değerleri olmuştur. Bu baskınlıktan dolayı başarı artmıştır. Basınç değeri sıcaklık tahmininde etkin değerlerden biridir.

Tablo 5 Eğitim-3 Sonucu Elde Edilen Başarım Sonuçları Tablosu

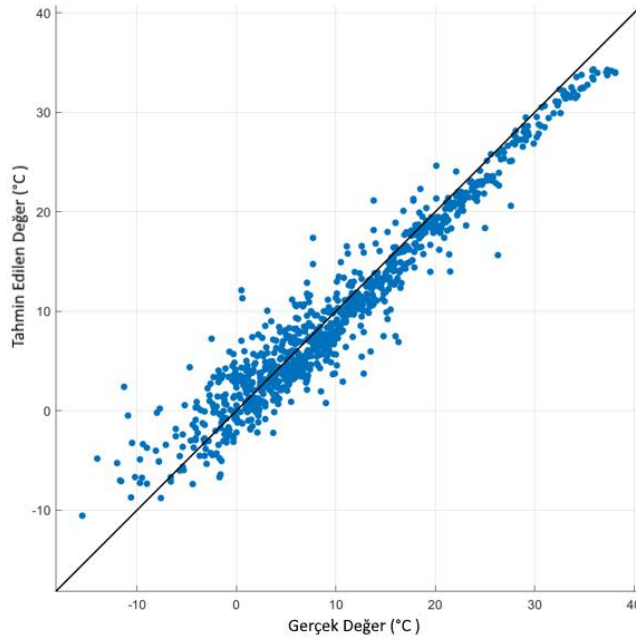
	K=5	K=10	K=15
Algoritma Adı	Başarı (RMSE)	Başarı (RMSE)	Başarı (RMSE)
Linear Regression (Linear)	26,364	33,323	40,293
Tree (Fine Tree)	3,5775	3,7211	3,4828
Tree (Medium Tree)	3,3155	3,3603	3,1735
Tree (Coarse Tree)	3,3278	3,2318	3,2197
SVM (Linear SVM)	3,3254	3,2709	3,2854
SVM (Quadratic SVM)	22276e07	1961e07	1,7259e07
SVM (Cubic SVM)	13168e15	87204e12	2,8817e07
SVM (Fine Gaussian SVM)	10,102	10,068	10,064
SVM (Medium Gaussian SVM)	4,1075	4,0234	3,9490
SVM (Coarse Gaussian SVM)	4,2585	4,1677	4,1374
Ensemble (Boosted Trees)	2,7464	2,7660	2,7496
Ensemble (Bagged Trees)	2,7909	2,7484	2,7723
Gaussian Process Regression (Squared Exponential GPR)	10,290	10,280	10,278
Gaussian Process Regression (Matern 5/2 GPR)	10,262	10,247	10,247
Gaussian Process Regression (Exponential GPR)	7,3045	4,8411	5,7861
Gaussian Process Regression (Rational Quadratic GPR)	3,2014	3,1438	3,1420

Şekil 6.7’de Eğitim-3 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait tepki grafiği görülmektedir.



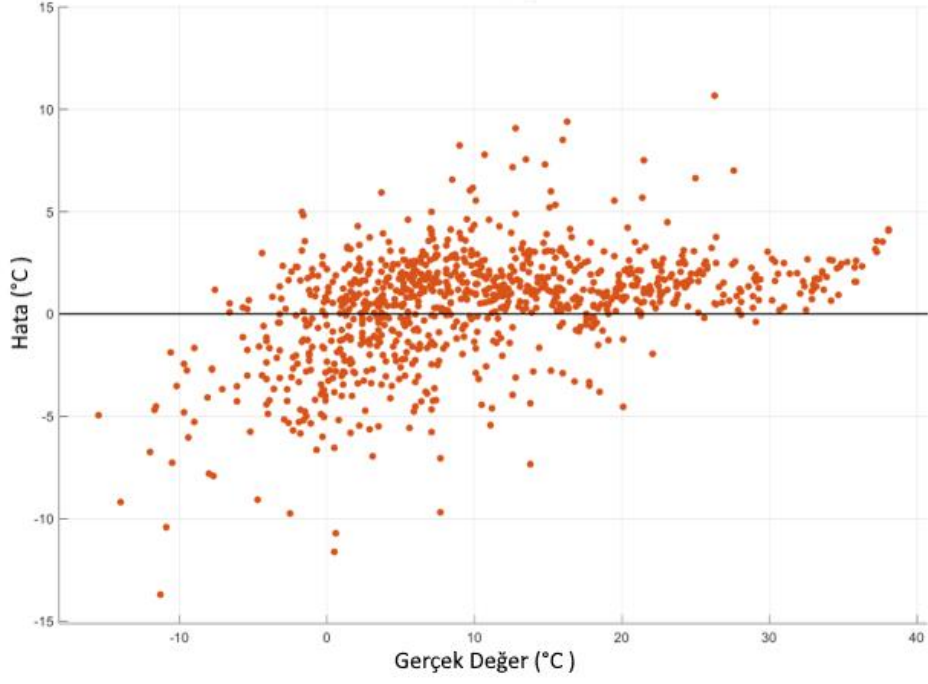
Şekil 6.7 Eğitim-3 Sonucu Elde Edilen Tepki Grafiği

Şekil 6.8’de Eğitim-3 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait sapma grafiği görülmektedir.



Şekil 6.8 Eğitim-3 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği

Şekil 6.9’da Eğitim-3 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait hata grafiği görülmektedir.



Şekil 6.9 Eğitim-3 Sonucu Elde Edilen Hata Grafiği

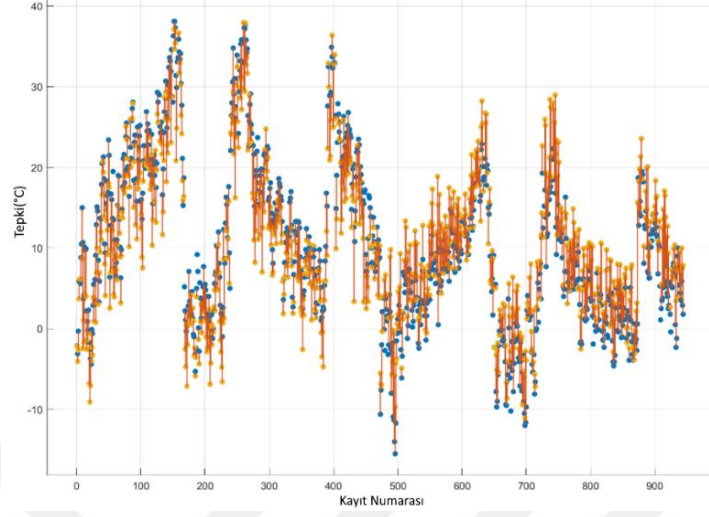
6.4 Eğitim-4:

Bu eğitimde, üç yıllık sabah ve akşam verileri toplanmış rasat verileri kullanılarak gün içindeki en yüksek ve en düşük sıcaklık kestirimi yapılmıştır. K Çaprazlama değeri 15 olarak seçilmiştir. Kullanılan veriler normalize edilmemiş verilerdir. Eğitim sonucu elde edilen başarılar Tablo 6’da görülmektedir. En yüksek başarı Gauss Süreç Regresyonu yöntemi ile sağlanmıştır. Eğitim 4’te Eğitim 3’ten farklı olarak Eğitim 3’te ki verilere PCA yöntemi uygulanarak öznelik çıkartılmıştır. Temel bileşenlerin toplam varyansları değişikçe başarı sonuçlarına etkisi analiz edilmiştir.

Tablo 6 Eğitim-4 Sonucu Elde Edilen Başarım Sonuçları Tablosu

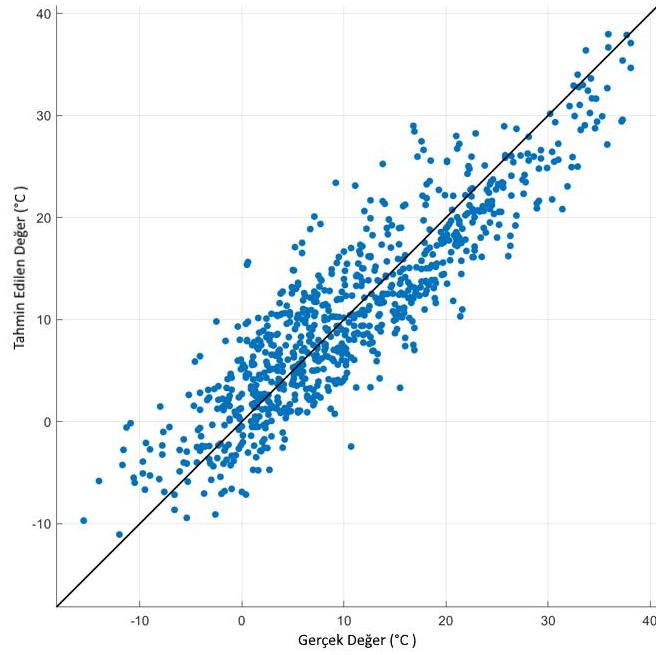
	PCA=90	PCA=95	PCA=97
Algoritma Adı	Başarı (RMSE)	Başarı (RMSE)	Başarı (RMSE)
Linear Regression (Linear)	6,2389	5,6559	4,6426
Tree (Fine Tree)	7,9452	7,5293	7,4144
Tree (Medium Tree)	7,2893	6,9623	7,0601
Tree (Coarse Tree)	7,0596	6,8304	6,8346
SVM (Linear SVM)	6,4732	5,9784	4,7183
SVM (Quadratic SVM)	6,4821	5,9823	5,1350
SVM (Cubic SVM)	8,1910	7,6934	5,6785
SVM (Fine Gaussian SVM)	6,5025	5,6359	4,7523
SVM (Medium Gaussian SVM)	8,9090	8,7401	8,4435
SVM (Coarse Gaussian SVM)	10,301	10,293	10,271
Ensemble (Boosted Trees)	6,7852	6,3750	6,1379
Ensemble (Bagged Trees)	6,8013	6,4295	6,1934
Gaussian Process Regression (Squared Exponential GPR)	5,9372	5,4130	4,6012
Gaussian Process Regression (Matern 5/2 GPR)	7,1898	6,5966	5,6730
Gaussian Process Regression (Exponential GPR)	6,7393	6,2415	5,7194
Gaussian Process Regression (Rational Quadratic GPR)	5,9362	5,4130	4,5963

Şekil 6.10'da Eğitim-4 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait tepki grafiği görülmektedir.



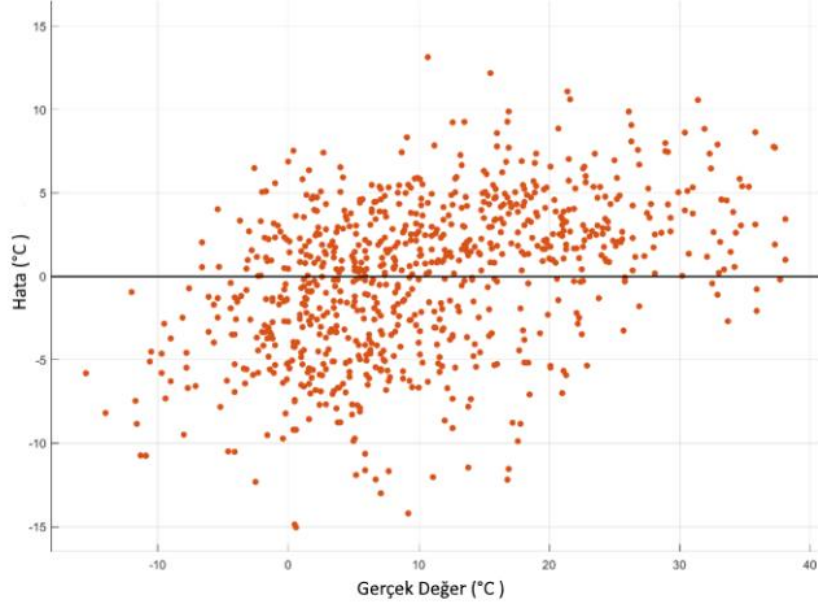
Şekil 6.10 Eğitim-4 Sonucu Elde Edilen Tepki Grafiği

Şekil 6.11'de Eğitim-4 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait sapma grafiği görülmektedir.



Şekil 6.11 Eğitim-4 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği

Şekil 6.12’de Eğitim-4 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait hata grafiği görülmektedir.



Şekil 6.12 Eğitim-4 Sonucu Elde Edilen Hata Grafiği

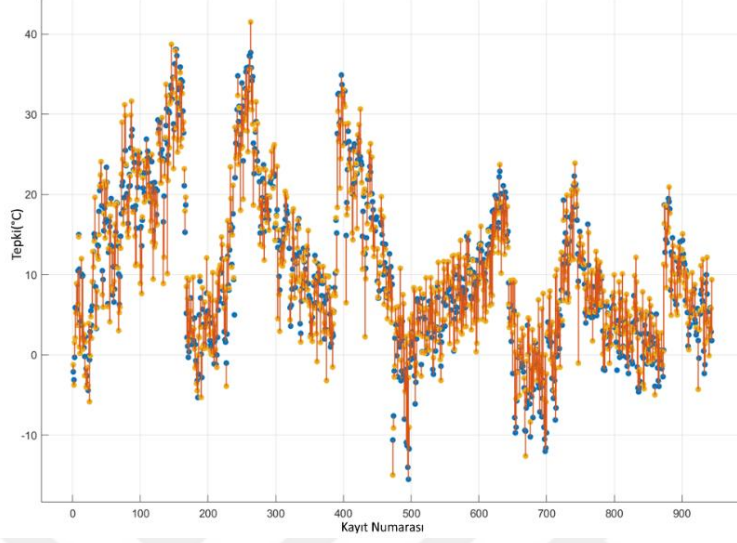
6.5 Eğitim-5:

Bu eğitimde, üç yıllık sabah ve akşam verileri toplanmış rasat verileri kullanılarak gün içindeki en yüksek ve en düşük sıcaklık kestirimi yapılmıştır. Bu veriler 400mbar'a kadar toplanmıştır. 850mbar ile 400mbar arasındaki hava durumu verileri hava tahmininde baskın durumdadır [23]. Bu doğrultuda 850mbar ile 400mbar arasındaki hava durumu verileri kullanılmıştır. Tablo 7'da yer seviyesinden 400mbar'a kadar olan seviyedeki veriler ile eğitilmiş başarılar görülmektedir. Tablo 7'de görüldüğü gibi en yüksek başarı GSR ile sağlanmıştır. Kullanılan veriler normalize edilmiş verilerdir. Bu eğitimde Eğitim 1'den farklı olarak meteorolojide önemli kabul edilen noktalardan biri olan 400mbar seviyesine kadar olan veriler ile eğitim yapıp başarıya etkisi incelenmiştir [23].

Tablo 7 Eğitim-5 Sonucu Elde Edilen Başarım Sonuçları Tablosu

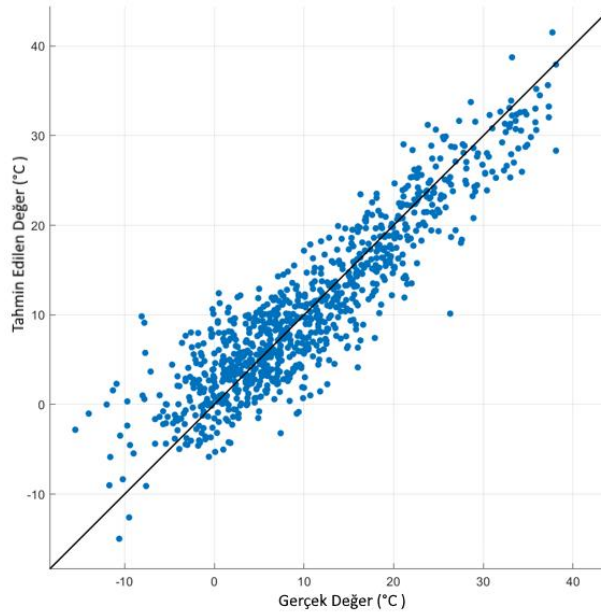
	K=5	K=10	K=15
Algoritma Adı	Başarı (RMSE)	Başarı (RMSE)	Başarı (RMSE)
Linear Regression (Linear)	80,902	105,09	116,92
Tree (Fine Tree)	7,2181	7,3977	7,0554
Tree (Medium Tree)	6,8189	6,8841	6,6986
Tree (Coarse Tree)	6,5190	6,4644	6,3454
SVM (Linear SVM)	5,6076	5,4648	5,4781
SVM (Quadratic SVM)	5,0417	5,0125	4,8630
SVM (Cubic SVM)	13,472	7,3060	6,0140
SVM (Fine Gaussian SVM)	10,192	10,197	10,188
SVM (Medium Gaussian SVM)	5,0175	5,0311	4,9930
SVM (Coarse Gaussian SVM)	5,7844	5,7067	5,7264
Ensemble (Boosted Trees)	5,1884		5,1344
Ensemble (Bagged Trees)	5,4050	5,3477	5,3126
Gaussian Process Regression (Squared Exponential GPR)	10,245	10,255	10,255
Gaussian Process Regression (Matern 5/2 GPR)	10,242	10,253	13,252
Gaussian Process Regression (Exponential GPR)	7,8371	7,5158	6,8566
Gaussian Process Regression (Rational Quadratic GPR)	4,4034	4,3300	4,2084

Şekil 6.13'te Eğitim-5 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait tepki grafiği görülmektedir.



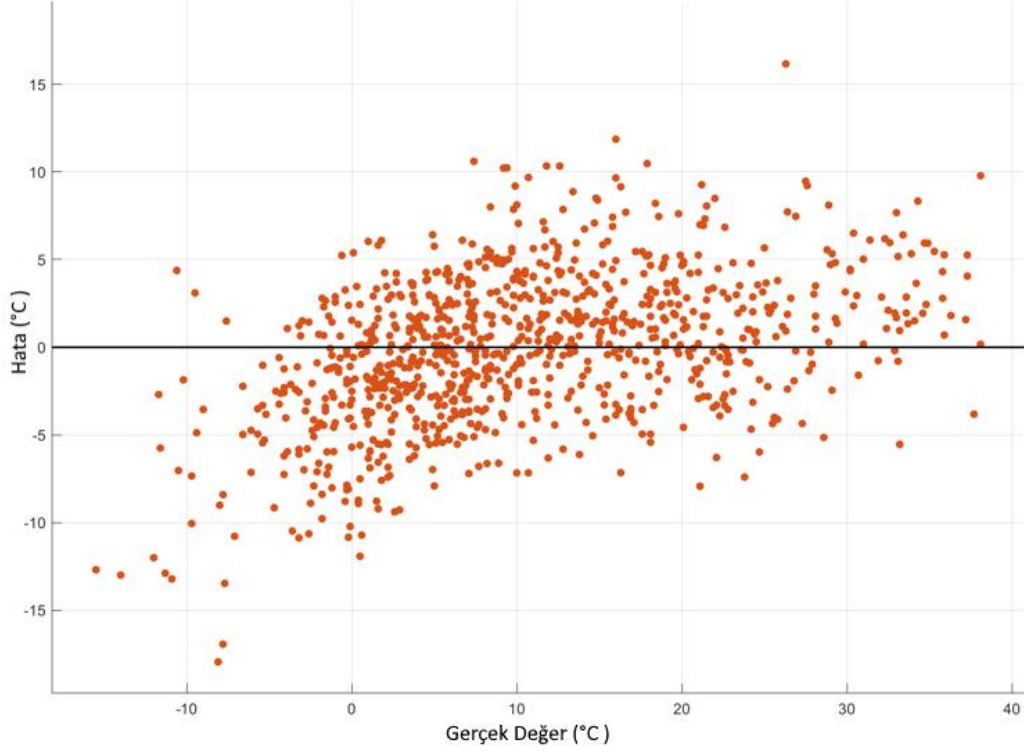
Şekil 6.13 Eğitim-5 Sonucu Elde Edilen Tepki Grafiği

Şekil 6.14'de Eğitim-5 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait sapma grafiği görülmektedir.



Şekil 6.14 Eğitim-5 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği

Şekil 6.15'te Eğitim-5 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait hata grafiği görülmektedir.



Şekil 6.15 Eğitim-5 Sonucu Elde Edilen Hata Grafiği

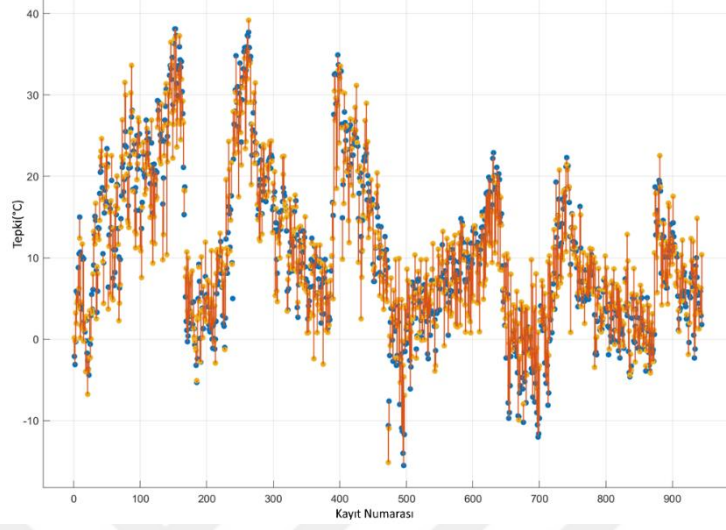
6.6 Eğitim-6:

Bu eğitimde üç yıllık sabah ve akşam verileri toplanmış rasat verileri normalize edilmeden kullanılarak gün içindeki en yüksek ve en düşük sıcaklık kestirimi yapılmıştır. Eğitim 5'ten farklı olarak yere daha yakın olan 500mbar'a kadar verilerle eğitim gerçekleştirilmiştir. Eğitimde kullanılan veriler normalize edilmiş verilerdir. K Çaprazlama değeri 15 olarak seçilmiştir. Tablo 8'te görüldüğü üzere eğitimdeki en yüksek başarı GSR ile sağlanmıştır. Bu eğitimde Eğitim 1'den farklı olarak meteorolojide önemli olan noktalardan bir diğeri olan 500mbar seviyesine kadar olan veriler ile eğitim yapıp başarıya etkisi incelenmiştir [23].

Tablo 8 Eğitim-6 Sonucu Elde Edilen Başarım Sonuçları Tablosu

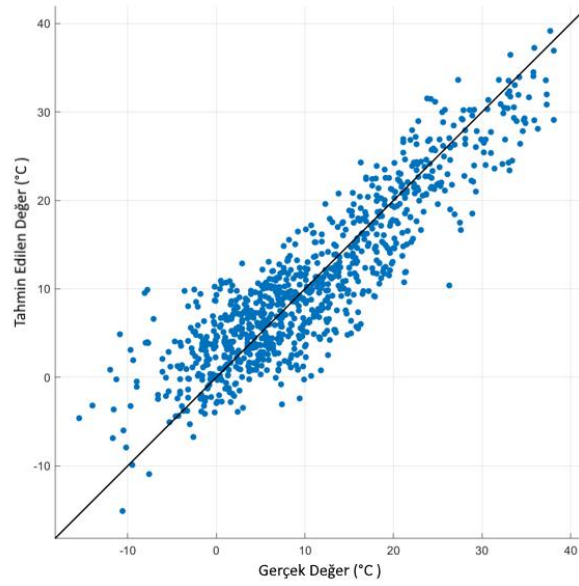
	PCA=90	PCA=95	PCA=97	K=15
Algoritma Adı	Başarı (RMSE)	Başarı (RMSE)	Başarı (RMSE)	Başarı (RMSE)
Linear Regression (Linear)	7,4248	7,2808	7,2938	155,38
Tree (Fine Tree)	10,075	10,155	10,570	7,1011
Tree (Medium Tree)	8,9656	9,1522	9,3445	6,7177
Tree (Coarse Tree)	8,5388	8,6113	8,6183	6,3345
SVM (Linear SVM)	7,4599	7,3418	7,4196	5,9897
SVM (Quadratic SVM)	7,3199	7,1516	7,2560	5,1703
SVM (Cubic SVM)	7,5075	7,6729	7,8456	7,4699
SVM (Fine Gaussian SVM)	7,3150	7,1316	7,1012	10,173
SVM (Medium Gaussian SVM)	8,3696	8,2845	8,2490	5,3529
SVM (Coarse Gaussian SVM)	10,210	10,200	10,198	6,0761
Ensemble (Boosted Trees)	7,7931	7,6544	7,7447	5,3406
Ensemble (Bagged Trees)	7,7114	7,7009	7,8670	5,4038
Gaussian Process Regression (Squared Exponential GPR)	7,1043	7,0027	6,8906	10,250
Gaussian Process Regression (Matern 5/2 GPR)	7,1508	7,2167	7,1262	10,248
Gaussian Process Regression (Exponential GPR)	7,1356	7,2232	7,2374	6,3373
Gaussian Process Regression (Rational Quadratic GPR)	7,0920	7,0069	6,8361	4,6212

Şekil 6.16’da Eğitim-6 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait tepki grafiği görülmektedir.



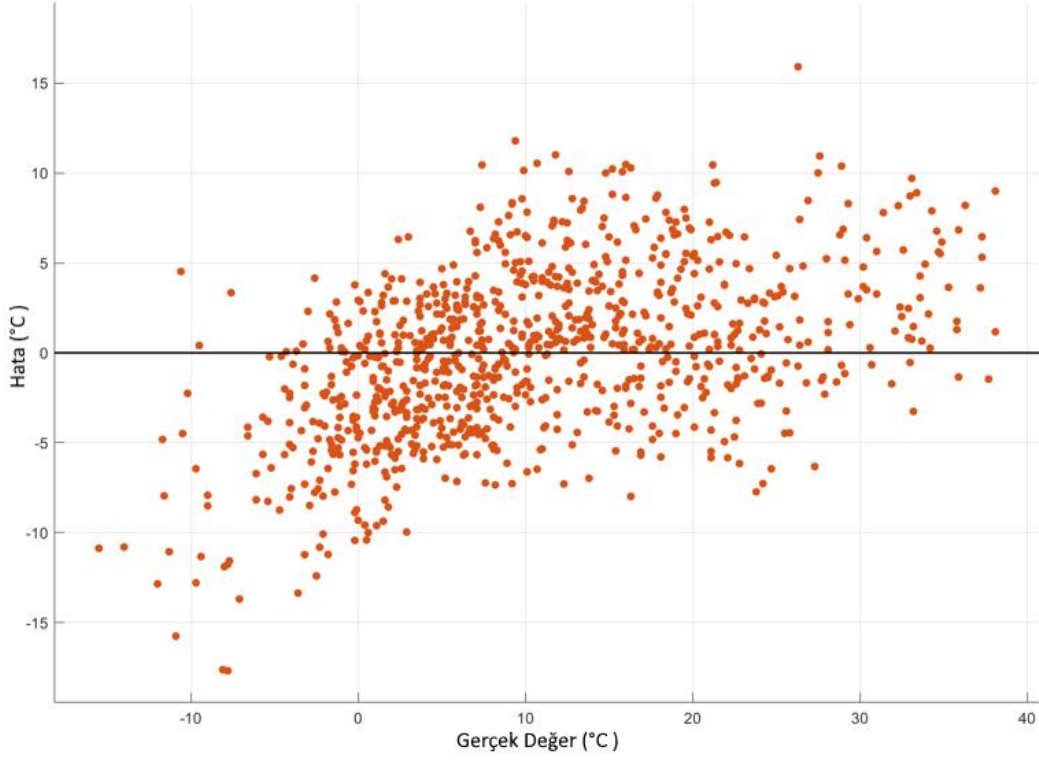
Şekil 6.16 Eğitim-6 Sonucu Elde Edilen Tepki Grafiği

Şekil 6.17’de Eğitim-6 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait sapma grafiği görülmektedir.



Şekil 6.17 Eğitim-6 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerler Grafiği

Şekil 6.18'te Eğitim-6 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait hata grafiği görülmektedir.



Şekil 6.18 Eğitim-6 Sonucu Elde Edilen Hata Grafiği

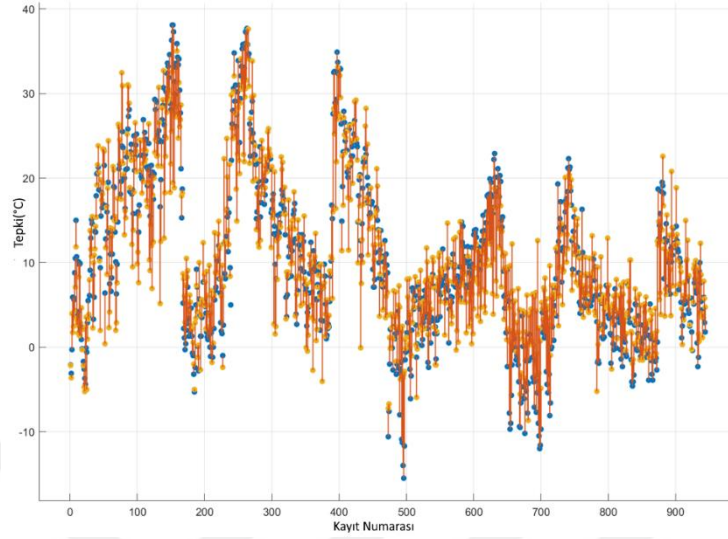
6.7 Eğitim-7:

Bu eğitimde, üç yıllık sabah ve akşam verileri toplanmış rasat verileri normalize edilmeden kullanılarak gün içindeki en yüksek ve en düşük sıcaklık kestirimi yapılmıştır. Eğitim 6'daki farklı olarak yere daha yakın olan 600mbar'a kadar verilerle eğitim gerçekleştirilmiştir. Eğitimde kullanılan veriler normalize edilmiş verilerdir. K Çaprazlama değeri 15 olarak seçilmiştir. Tablo 9'da görüldüğü gibi eğitimdeki en yüksek başarı GSR ile sağlanmıştır. Bu eğitimde Eğitim 1'den farklı olarak meteorolojide önemli olan noktalardan bir diğeri olan 600mbar seviyesine kadar olan normalize edilmiş veriler ile eğitim yapıp başarıya etkisi incelenmiştir [23].

Tablo 9 Eğitim-7 Sonucu Elde Edilen Başarım Sonuçları Tablosu

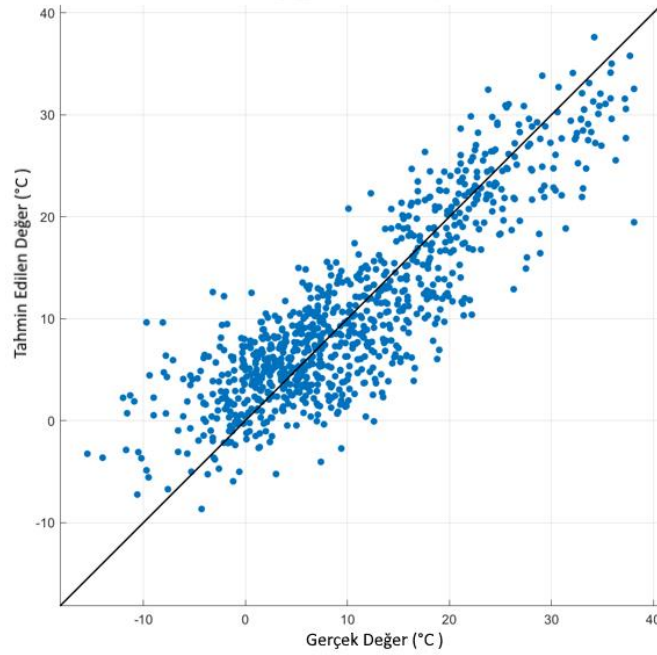
	PCA=90	PCA=95	PCA=97	K=15
Algoritma Adı	Başarı (RMSE)	Başarı (RMSE)	Başarı (RMSE)	Başarı (RMSE)
Linear Regression (Linear)	5,9202	5,6583	5,6204	189,87
Tree (Fine Tree)	8,8012	9,1022	9,0575	7,3130
Tree (Medium Tree)	8,3649	8,5055	8,6111	6,7547
Tree (Coarse Tree)	8,1974	8,1869	8,1801	6,4428
SVM (Linear SVM)	5,9833	5,6476	5,6663	6,4603
SVM (Quadratic SVM)	5,7473	5,3887	5,1992	5,7493
SVM (Cubic SVM)	6,1499	5,4838	5,2643	8,3135
SVM (Fine Gaussian SVM)	6,1928	5,8281	5,8745	9,9582
SVM (Medium Gaussian SVM)	9,1291	9,0766	9,0597	5,5971
SVM (Coarse Gaussian SVM)	10,060	10,057	10,055	6,1404
Ensemble (Boosted Trees)	6,9281	6,9606	6,9317	5,3855
Ensemble (Bagged Trees)	7,2660	7,2801	7,4683	5,3708
Gaussian Process Regression (Squared Exponential GPR)	5,7057	5,3917	5,2073	10,248
Gaussian Process Regression (Matern 5/2 GPR)	7,5214	8,4064	7,8529	10,245
Gaussian Process Regression (Exponential GPR)	7,1121	7,8044	7,9202	5,6053
Gaussian Process Regression (Rational Quadratic GPR)	5,7057	5,3917	5,2073	5,1316

Şekil 6.19’da Eğitim-7 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait tepki grafiği görülmektedir.



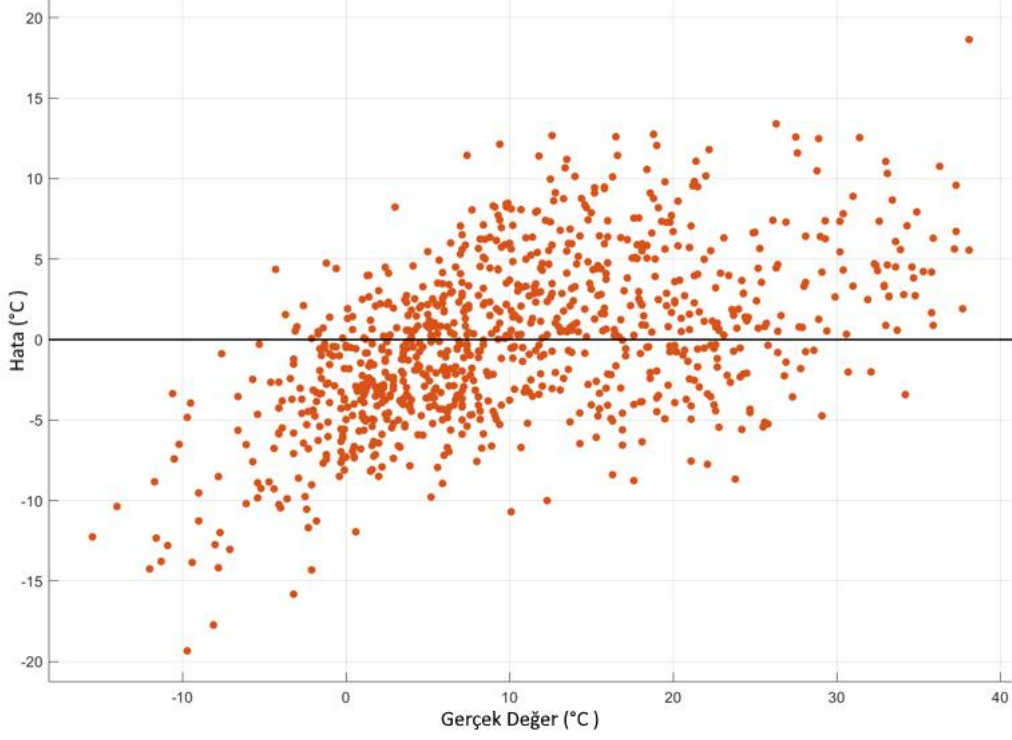
Şekil 6.19 Eğitim-7 Sonucu Elde Edilen Tepki Grafiği

Şekil 6.20’de Eğitim-7 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait sapma grafiği görülmektedir.



Şekil 6.20 Eğitim-7 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği

Şekil 6.21’de Eğitim-7 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait hata grafiği görülmektedir.



Şekil 6.21 Eğitim-7 Sonucu Elde Edilen Hata Grafiği

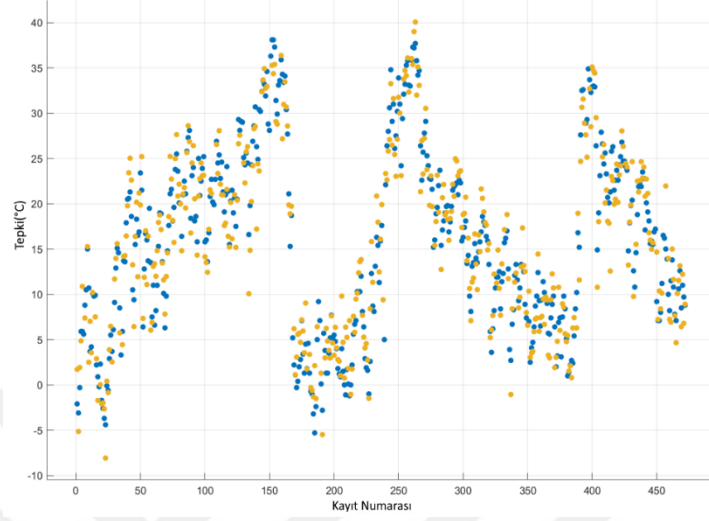
6.8 Eğitim-8:

Bu eğitimde, sadece üç yıllık sabah rastına ait veriler kullanılmıştır. Gün içindeki en yüksek sıcaklık kestirimi yapılmıştır. Bunun sebebi en yüksek ve en düşük sıcaklık tahmini için 2 farklı model önererek tahmin başarısını artırmayı hedeflemektir. Kullanılan veriler normalize edilmiş verilerdir. K Çaprazlama değeri 15 olarak seçilmiştir. Tablo 10'daki en yüksek başarı Gauss Süreç Regresyonu ile sağlanmıştır. Eğitim 8'de önceki eğitimlerden farklı olarak en yüksek ve en düşük sıcaklık değerlerini aynı anda tahmin etmek yerine, en yüksek sıcaklık tahmini için ayrı, en düşük sıcaklık tahmini için ayrı olmak üzere 2 model önerilmektedir. Bu eğitimde sadece en yüksek sıcaklık değeri tahmin edilmeye çalışılmıştır.

Tablo 10 Eğitim-8 Sonucu Elde Edilen Başarım Sonuçları Tablosu

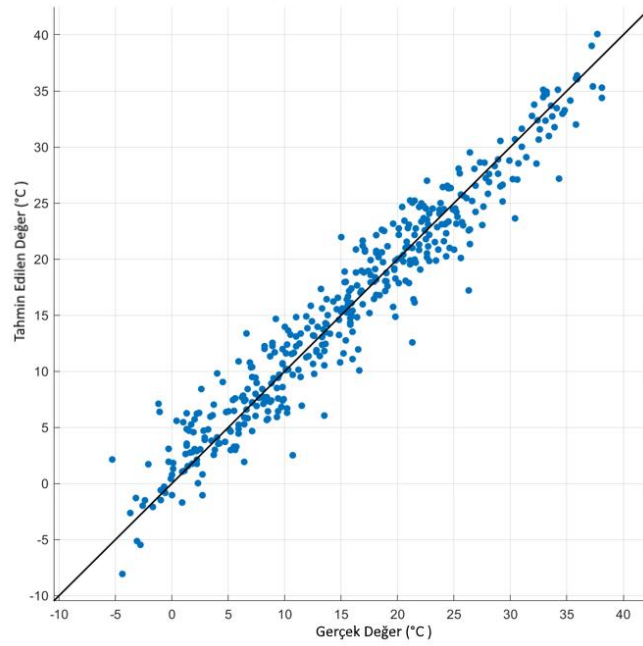
	PCA=90	PCA=95	PCA=97	Ham veri
Algoritma Adı	Başarı (RMSE)	Başarı (RMSE)	Başarı (RMSE)	Başarı (RMSE)
Linear Regression (Linear)	5,9202	5,6583	5,6204	37,992
Tree (Fine Tree)	8,8012	9,1022	9,0575	6,1062
Tree (Medium Tree)	8,3649	8,5055	8,6111	5,7071
Tree (Coarse Tree)	8,1974	8,1869	8,1801	5,8898
SVM (Linear SVM)	5,9833	5,6476	5,6663	3,6972
SVM (Quadratic SVM)	5,7473	5,3887	5,1992	3,3546
SVM (Cubic SVM)	6,1499	5,4838	5,2643	3,3850
SVM (Fine Gaussian SVM)	6,1928	5,8281	5,8745	10,097
SVM (Medium Gaussian SVM)	9,1291	9,0766	9,0597	3,8124
SVM (Coarse Gaussian SVM)	10,060	10,057	10,055	4,4335
Ensemble (Boosted Trees)	6,9281	6,9606	6,9317	4,2045
Ensemble (Bagged Trees)	7,2660	7,2801	7,4683	4,3394
Gaussian Process Regression (Squared Exponential GPR)	5,7057	5,3917	5,2073	10,107
Gaussian Process Regression (Matern 5/2 GPR)	7,5214	8,4064	7,8529	10,091
Gaussian Process Regression (Exponential GPR)	7,1121	7,8044	7,9202	7,8677
Gaussian Process Regression (Rational Quadratic GPR)	5,7057	5,3917	5,2073	2,6187

Şekil 6.22’de Eğitim-8 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait tepki grafiği görülmektedir.



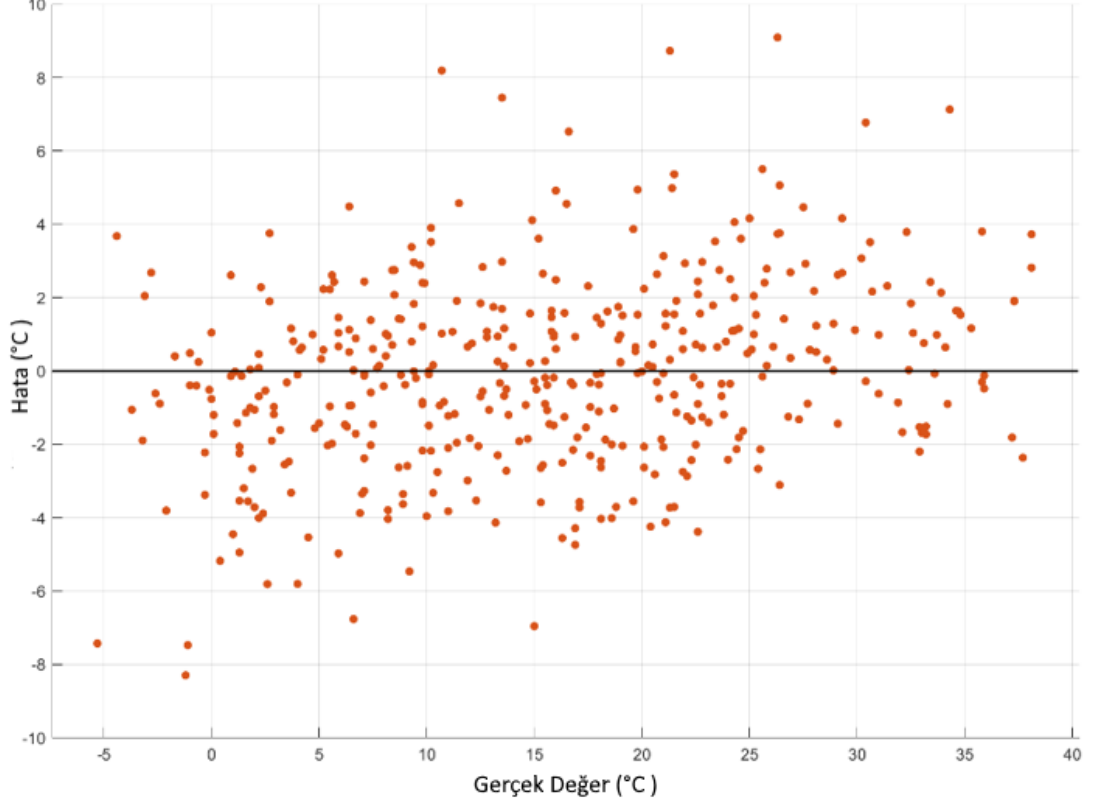
Şekil 6.22 Eğitim-8 Sonucu Elde Edilen Tepki Grafiği

Şekil 6.23’te Eğitim-8 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait sapma grafiği görülmektedir.



Şekil 6.23 Eğitim-8 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği

Şekil 6.24'te Eğitim-8 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait hata grafiği görülmektedir.



Şekil 6.24 Eğitim-8 Sonucu Elde Edilen Hata Grafiği

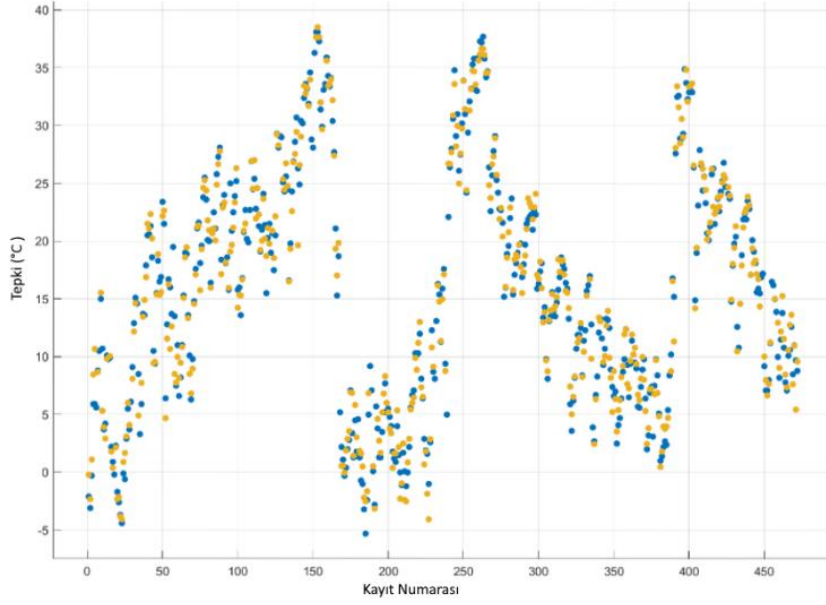
6.9 Eğitim-9:

Bu eğitimde, sadece üç yıllık sabah rastına ait veriler kullanılmıştır. Tablo 11’de kullanılan veriler normalizasyon yapılmadan ağırlıklandırma yöntemi ile başarı artırılmaya çalışılmıştır. Gün içindeki en yüksek sıcaklık kestirimi yapılmıştır. Bunun sebebi en yüksek ve en düşük sıcaklık tahminini ayırarak Eğitim 8’deki gibi başarıyı yükseltmektir. K Çaprazlama değeri 15 olarak seçilmiştir. Tablo 11’de görüldüğü gibi eğitimdeki en yüksek başarı Gauss Süreç Regresyonu ile sağlanmıştır. Bu eğitim ile çalışmadaki en yüksek başarı sağlanmıştır. Eğitim 9 bir önceki Eğitim 8’den farklı olarak veriler normalize edilmeden ağırlıklandırma yapılarak eğitime dâhil edilmiş, bu koşullar altındaki başarısı analiz edilmiştir.

Tablo 11 Eğitim-9 Sonucu Elde Edilen Başarım Sonuçları Tablosu

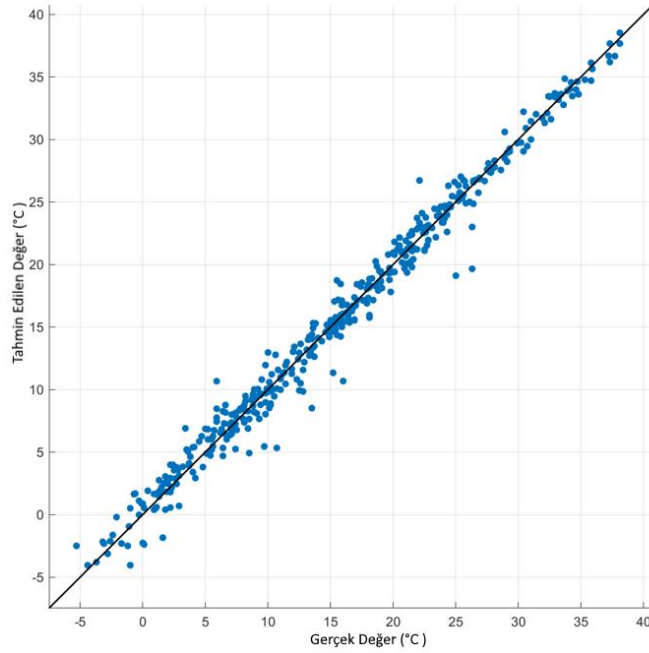
	PCA=90	PCA=95	PCA=97	K=15
Algoritma Adı	Başarı (RMSE)	Başarı (RMSE)	Başarı (RMSE)	Başarı (RMSE)
Linear Regression (Linear)	3,6727	3,0441	1,9403	9,3472
Tree (Fine Tree)	5,2122	5,2018	5,2408	1,7243
Tree (Medium Tree)	5,0896	5,1326	5,2406	1,6906
Tree (Coarse Tree)	5,2751	5,2751	5,2751	2,3511
SVM (Linear SVM)	3,8949	3,3715	1,9682	1,4074
SVM (Quadratic SVM)	3,2145	2,7061	2,0450	1,6604
SVM (Cubic SVM)	4,1980	3,4565	2,3148	1,8868
SVM (Fine Gaussian SVM)	3,5661	3,0649	2,1717	9,9421
SVM (Medium Gaussian SVM)	8,8767	8,8190	8,7439	2,1351
SVM (Coarse Gaussian SVM)	10,073	10,071	10,068	1,8352
Ensemble (Boosted Trees)	4,6483	4,6005	4,5826	1,5600
Ensemble (Bagged Trees)	4,7132	4,7156	4,7353	1,3667
Gaussian Process Regression (Squared Exponential GPR)	2,9585	2,5843	1,7962	10,121
Gaussian Process Regression (Matern 5/2 GPR)	4,5634	4,5477	22,484	10,099
Gaussian Process Regression (Exponential GPR)	4,2859	4,1791	3,8174	4,9283
Gaussian Process Regression (Rational Quadratic GPR)	2,9199	2,5843	1,7962	1,2893

Şekil 6.25'te Eğitim-9 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait tepki grafiği görülmektedir.



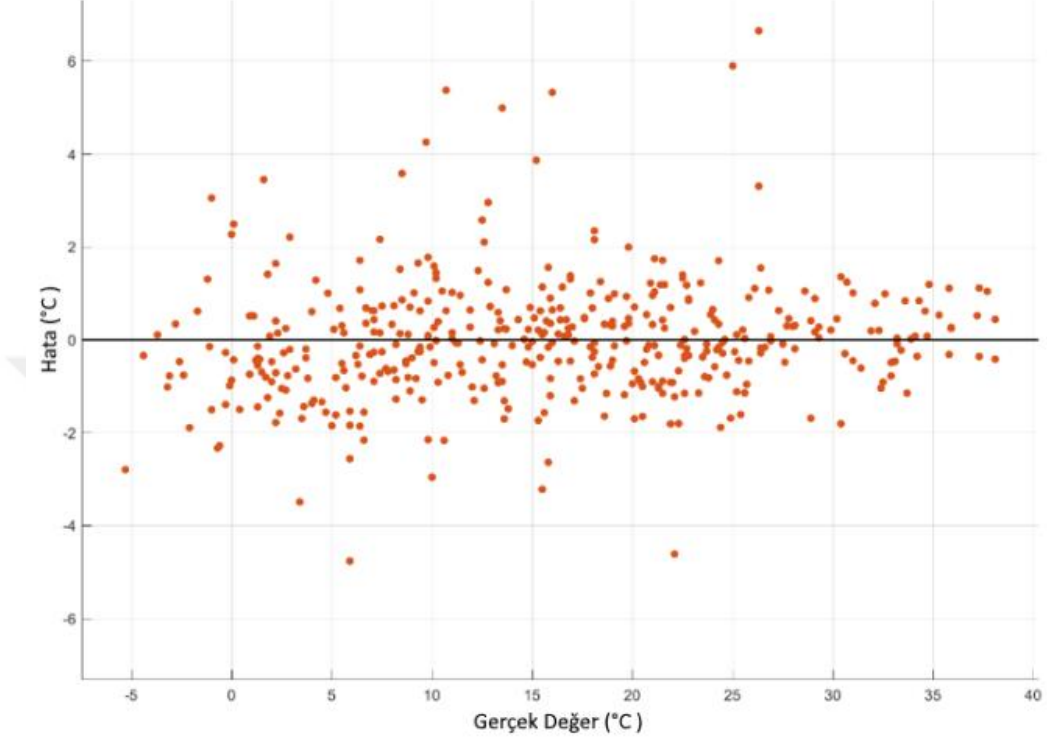
Şekil 6.25 Eğitim-9 Sonucu Elde Edilen Tepki Grafiği

Şekil 6.26'da Eğitim-9 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait sapma grafiği görülmektedir.



Şekil 6.26 Eğitim-9 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği

Şekil 6.27’de Eğitim-9 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait hata grafiği görülmektedir.



Şekil 6.27 Eğitim-9 Sonucu Elde Edilen Hata Grafiği

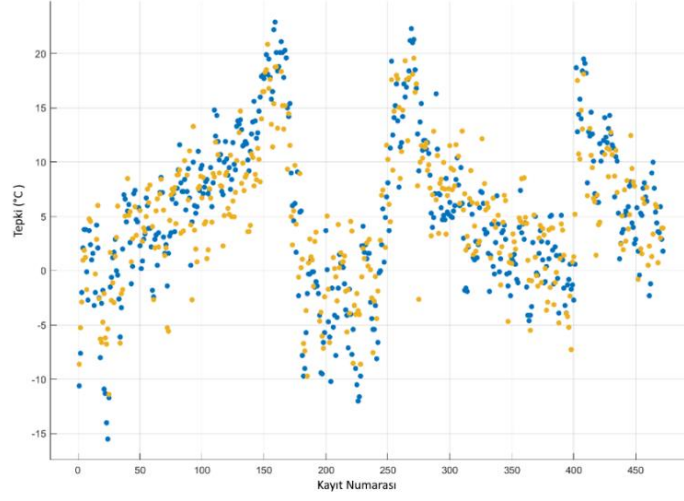
6.10 Eğitim-10:

Bu eğitimde, üç yıllık akşam toplanmış rasat verileri kullanılarak gün içindeki en düşük sıcaklık kestirimi yapılmıştır. Kullanılan veriler normalize edilmiş verilerdir. Tablo 12’de de görüldüğü üzere kullanılan veriler normalizasyon yapılarak başarı artırılmaya çalışılmıştır. Gün içindeki en düşük sıcaklık kestirimi yapılmıştır. K Çaprazlama değeri farklı değerlerde eğitim yapılmıştır. Tablo 12’deki en yüksek başarı K Çaprazlama değeri 15 olarak seçilerek Gauss Süreç Regresyonu ile başarı sağlanmıştır. Eğitim 10’da önceki eğitimlerden farklı olarak sadece en düşük sıcaklık tahmini için model oluşturulmuştur. Farklı algoritmaların başarı üzerindeki etkisi analiz edilmiştir.

Tablo 12 Eğitim-10 Sonucu Elde Edilen Başarım Sonuçları Tablosu

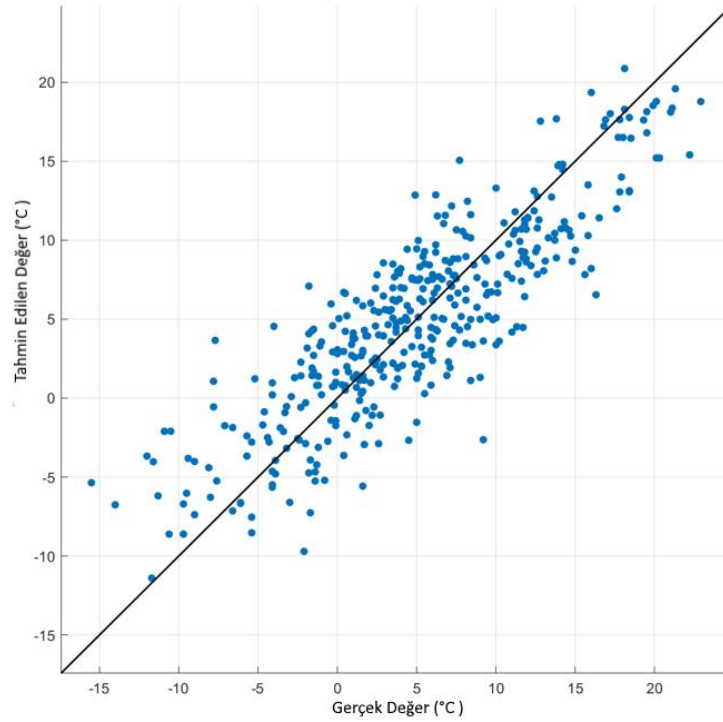
	K=5	K=10	K=15
Algoritma Adı	Başarı (RMSE)	Başarı (RMSE)	Başarı (RMSE)
Linear Regression (Linear)	32,066	32,672	30,687
Tree (Fine Tree)	6,0767	5,6188	5,8034
Tree (Medium Tree)	5,5364	5,3615	5,2326
Tree (Coarse Tree)	4,4615	5,1468	5,1283
SVM (Linear SVM)	4,3650	4,3075	4,2014
SVM (Quadratic SVM)	3,9639	3,9331	3,8999
SVM (Cubic SVM)	4,2489	4,3075	5,7393
SVM (Fine Gaussian SVM)	7,1905	3,9331	7,1704
SVM (Medium Gaussian SVM)	4,0880	4,5371	3,9895
SVM (Coarse Gaussian SVM)	4,9592	7,1660	4,7102
Ensemble (Boosted Trees)	4,2211	4,0494	4,2846
Ensemble (Bagged Trees)	4,2840	4,2545	4,3465
Gaussian Process Regression (Squared Exponential GPR)	7,1912	7,1717	7,1781
Gaussian Process Regression (Matern 5/2 GPR)	7,2149	7,2013	7,2078
Gaussian Process Regression (Exponential GPR)	6,3589	6,3608	6,7796
Gaussian Process Regression (Rational Quadratic GPR)	3,6670	3,7186	3,6580

Şekil 6.28’te Eğitim-10 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait tepki grafiği görülmektedir.



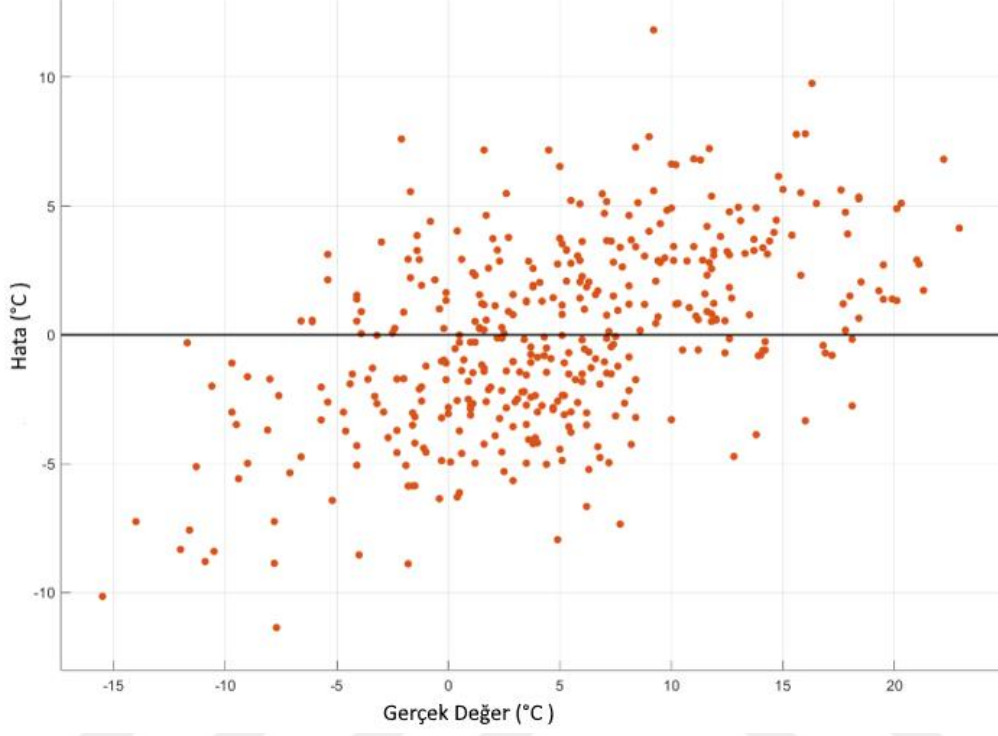
Şekil 6.28 Eğitim-10 Sonucu Elde Edilen Tepki Grafiği

Şekil 6.29’da Eğitim-10 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait sapma grafiği görülmektedir.



Şekil 6.29 Eğitim-10 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği

Şekil 6.30'da Eğitim-10 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait hata grafiği görülmektedir.



Şekil 6.30 Eğitim-10 Sonucu Elde Edilen Hata Grafiği

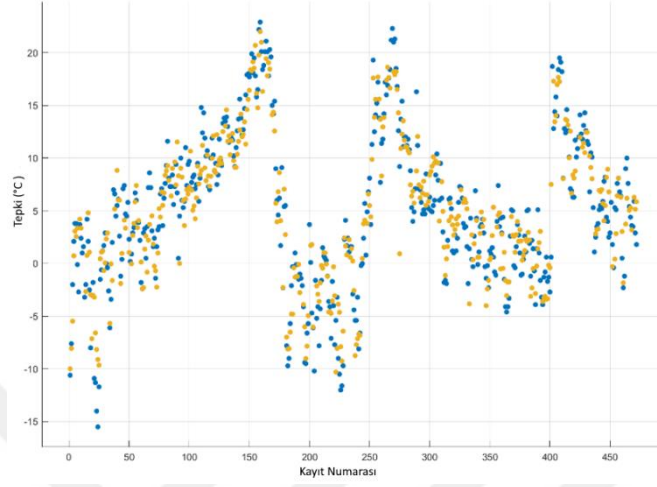
6.11 Eğitim-11:

Bu eğitimde, üç yıllık akşam toplanmış rasat verileri kullanılarak gün içindeki en düşük sıcaklık kestirimi yapılmıştır. Tablo 13’de kullanılan veriler normalizasyon yapılmamış verilerdir. Gün içindeki en düşük sıcaklık kestirimi yapılmıştır. Bunun sebebi en yüksek ve en düşük sıcaklık tahminini ayırarak eğitim 11’deki gibi başarıyı yükseltmektir. K Çaprazlama değeri farklı değerler seçilmiştir. Tablo 13’de görüldüğü üzere eğitimdeki en yüksek başarı Gauss Süreç Regresyonu ile sağlanmıştır. Bu eğitim ile çalışmadaki en yüksek başarı sağlanmıştır. Eğitim 11’de bir önceki Eğitim 10’dan farklı olarak veriler normalize edilerek eğitime dâhil edilmiştir.

Tablo 13 Eğitim-11 Sonucu Elde Edilen Başarım Sonuçları Tablosu

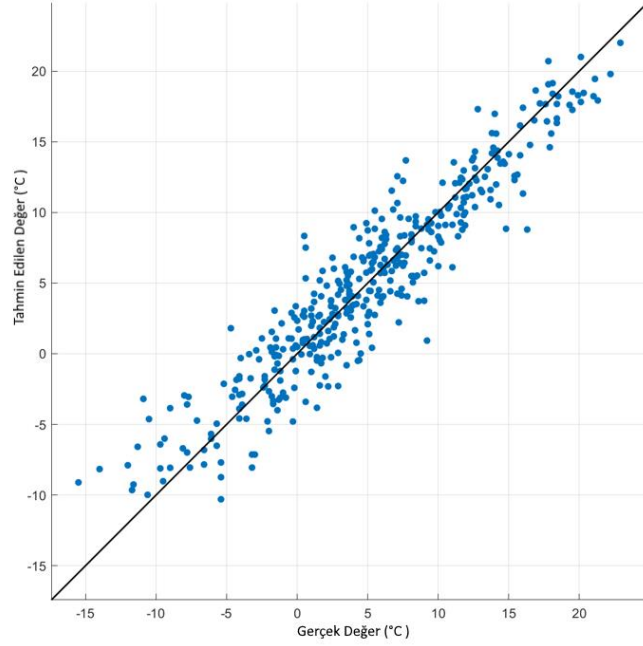
	K=5	K=10	K=15
Algoritma Adı	Başarı (RMSE)	Başarı (RMSE)	Başarı (RMSE)
Linear Regression (Linear)	20,142	20,774	20,162
Tree (Fine Tree)	3,4145	3,4151	3,3641
Tree (Medium Tree)	3,2480	3,1684	3,1645
Tree (Coarse Tree)	3,2294	3,3234	3,2157
SVM (Linear SVM)	2,6632	2,6074	2,6440
SVM (Quadratic SVM)	3,1233	3,0602	2,9535
SVM (Cubic SVM)	3,0328	2,8283	2,8865
SVM (Fine Gaussian SVM)	7,1481	7,1208	7,1264
SVM (Medium Gaussian SVM)	2,7364	2,7270	2,7138
SVM (Coarse Gaussian SVM)	2,6435	2,6251	2,6175
Ensemble (Boosted Trees)	2,5389	2,5455	2,4930
Ensemble (Bagged Trees)	2,6022	2,5496	2,5394
Gaussian Process Regression (Squared Exponential GPR)	7,1912	7,1779	7,1790
Gaussian Process Regression (Matern 5/2 GPR)	7,2209	7,2056	7,2116
Gaussian Process Regression (Exponential GPR)	5,9836	6,4264	5,8707
Gaussian Process Regression (Rational Quadratic GPR)	2,5207	2,4801	2,5092

Şekil 6.31’de Eğitim-11 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait tepki grafiği görülmektedir.



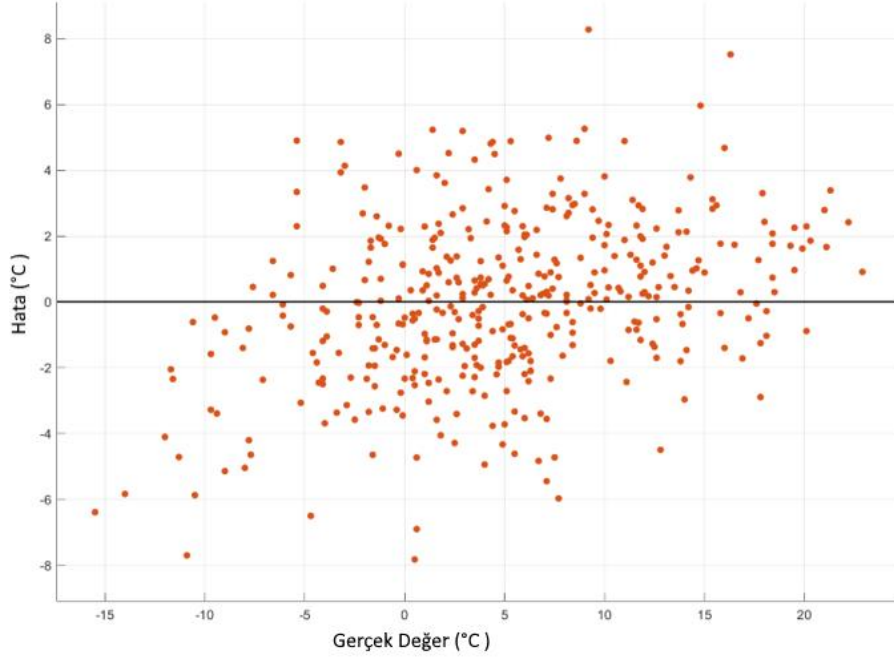
Şekil 6.31 Eğitim-11 Sonucu Elde Edilen Tepki Grafiği

Şekil 6.32’de Eğitim-11 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait sapma grafiği görülmektedir.



Şekil 6.32 Eğitim-11 Sonucu Elde Edilen Tahmin Değerleri ile Gerçek Değerlere Ait Sapma Grafiği

Şekil 6.33'te Eğitim-11 sonucu elde edilen en başarılı algoritma olan Gauss Süreç Regresyonu algoritmasına ait hata grafiği görülmektedir.



Şekil 6.33 Eğitim-11 Sonucu Elde Edilen Hata Grafiği

Tablo 14Tablo 2 Gerçekleştirilen Eğitim Sonuçları Özet Tablosu

	Veriler	Parametreler	Hangi Yöntem	Başarı (RMSE)	Başarı (R-Squared)	Başarı (MSE)	Başarı (MAE)
Eğitim 1	3 yıllık gündüz ve akşam rasat verileri. Veriler normalize edilmiştir.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.	Gauss Süreçi Regresyon	3,5101	0,88	12,321	2,6833
Eğitim 2	3 yıllık gündüz ve akşam rasat verileri. Veriler normalize edilmiştir.	PCA değerinin başarıya etkileri incelenmiştir. PCA 97 de en iyisidir.	Gauss Süreçi Regresyon	5,5601	0,71	30,914	4,3935
Eğitim 3	3 yıllık gündüz ve akşam rasat verileri. Veriler normalize edilmemiştir.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.	Ensemble (Boosted Trees)	2,7496	0,93	7,5601	2,0921
Eğitim 4	3 yıllık gündüz ve akşam rasat verileri. Veriler normalize edilmemiştir.	PCA değerinin başarıya etkileri incelenmiştir. PCA 97 de en iyisidir.	Gauss Süreçi Regresyon	4,5963	0,80	21,126	3,7225
Eğitim 5	3 yıllık sabah ve akşam rasat verileri kullanılmıştır. Veriler normalize edilmiştir. Bu rasat verileri yer seviyesinden ilk 400mbar kadar olan bölge alınmıştır.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.	Gauss Süreçi Regresyon	4,2084	0,83	17,711	3,3011
Eğitim 6	3 yıllık sabah ve akşam rasat verileri kullanılmıştır. Veriler normalize edilmiştir. Bu rasat verileri yer seviyesinden ilk 500mbar kadar olan bölge alınmıştır.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.	Gauss Süreçi Regresyon	4,6212	0,80	21,356	3,6409
Eğitim 7	3 yıllık sabah ve akşam rasat verileri kullanılmıştır. Veriler normalize edilmiştir. Bu rasat verileri yer seviyesinden ilk 600mbar kadar olan bölge alınmıştır.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.	Gauss Süreçi Regresyon	5,1316	0,75	26,333	4,0355
Eğitim 8	3 yıllık sabah datası ile en yüksek sıcaklık kestirimi yapılmıştır. Veriler normalize edilmiştir.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.	Gauss Süreçi Regresyon	2,6187	0,93	6,8576	2,0455
Eğitim 9	3 yıllık sabah datası ile en yüksek sıcaklık kestirimi yapılmıştır. Veriler normalize edilmemiştir.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.	Gauss Süreçi Regresyon	1,2893	0,98	1,6623	0,9145
Eğitim 10	3 yıllık akşam datası ile en düşük sıcaklık kestirimi yapılmıştır. Veriler normalize edilmiştir.	K Çaprazlama Değeri 15 Seçilmiştir. PCA yok.	Gauss Süreçi Regresyon	3,6580	0,74	13,381	2,9322
Eğitim 11	3 yıllık akşam datası ile en düşük sıcaklık kestirimi yapılmıştır. Veriler normalize edilmemiştir.	K Çaprazlama Değeri 10 Seçilmiştir. PCA yok.	Gauss Süreçi Regresyon	2,4801	0,88	6,1509	1,9758

Tablo 14'te bu çalışma kapsamında gerçekleştirilen eğitimlere ait başarımların sonuçları görülmektedir. Toplamda 11 farklı eğitim gerçekleştirilmiş olup, her eğitimde 16 farklı makine öğrenme algoritması $K=5$, $K=10$ ve $K=15$ çaprazlama değerleri için Radyosonde rasatları sonucunda elde edilen veri seti kullanılarak eğitilmiştir. Gerçekleştirilen eğitimlerde Eğitim-3 hariç en iyi başarı Gauss Süresi Regresyon algoritması ile elde edilmiştir. Eğitim-3'te ise en iyi başarı Ensemble (Boosted Trees) algoritması ile elde edilmiştir.

Gerçekleştirilen eğitimlerde Radyosonde rasatları sonucu elde edilen veri setlerinde öz nitelik çıkarımı çalışmaları gerçekleştirilmiştir. Eğitim-1, Eğitim-2 ve Eğitim-3'te 3 yıllık gündüz ve akşam rasatlarına ait veriler kullanılmıştır. Eğitim-5, Eğitim-6 ve Eğitim-7'de ise gerçekleştirilen ilk 3 eğitimden farklı olarak farklı irtifalara kadar veriler kullanılarak eğitimler gerçekleştirilmiştir. Ancak, verilerin kırılması eğitim sonuçlarının kötüleşmesine neden olmuştur.

Eğitim-8, Eğitim-9, Eğitim-10 ve Eğitim-11'de ise 3 yıllık rasatlar sabah ve akşam rasatları olarak ayrılmış. Sabah rasatlarına ait veriler kullanılarak ertesi güne ait en yüksek sıcaklık, akşam rasatlarına ait veriler kullanılarak ertesi güne ait en düşük hava sıcaklığı kestirilmiştir. Bu sayede daha önce gerçekleştirilen eğitimlere göre başarımların oranı artmıştır.

Bunların yanı sıra gerçekleştirilen eğitimlerde veriler normalize edilerek normalize edilmemiş halleriyle karşılaştırılmıştır. Verilerin normalize edilmesi eğitimlerin sonuçlarını kötüleştirmiştir.

6.12 Literatür Karşılaştırması

Bu çalışma daha önce yapılmış 4 farklı çalışma ile karşılaştırılmıştır.

Tablo 15 Hava Durumu Kestirim Literatür Karşılaştırma Tablosu

Çalışma	Data Büyüklüğü	Ölçüm Metodu	Kestirilen Veri	Yöntem	Başarı Oranı (%)	RMSE (%)
Nandar [4]	6 yıllık	Yer Gözlem İstasyonu	Yağacak Yağmur Oranı	SVM		3
Rahul [5]	10 yıllık	Yer Gözlem İstasyonu	Sıcaklık	LSTM	84	
Anbarasi [24]	1 yıllık	Yer Gözlem İstasyonu	Deniz Yüzeyindeki Sıcaklık	LSTM	89	
Wang [25]	1 gün	Yer Gözlem İstasyonu	Sıcaklık	Büyük Veri Kapsül Ağlar (MICaps)		5
Bu Çalışma	3 yıllık	Radyosonde	Maksimum Minimum Sıcaklık	GSR		1.2

Tablo 13’ de önceden gerçekleştirilmiş literatürde bulunan 4 farklı çalışma ile bu çalışmalar süresince oluşan hava durumu kestirim başarı oranları görülmektedir.

Nandar [4] çalışmasında, Mynmardaki yağmur oranlarını tahmin etmeye çalışmıştır. Kullandığı veri 1990’dan 2006’ya kadarki hava durumu verileridir. Bu çalışmada hava sıcaklık tahmini yerine Mynmarda oluşan selleri kestirmek için yağmur yağma miktarını hesaplamaya çalışılmıştır. SVM algoritmasını kullanmış ve RMSE %3 hata ile hesaplayabilmiştir.

Rahul [5] çalışmasında, Yer Gözlem İstasyonlarından topladığı verileri sırası ile sıcaklık, nem, çiy noktası, basınç, görünürlük ve yağış datalarını kullanarak eğitimlerini yapmıştır. Veri seti 2007’den 2017’e kadar olan verilerdir. Çalışmasında ulaştığı başarı oranı %84’tür.

Anbarasi [24] çalışmasında, büyük veri yöntemi ile hava durumu kestirimi yapmaya çalışmıştır. Kestirmeye çalıştığı veri denizciler için denizlerdeki hava durumudur. Büyük veriyle eğitim yapmak için 3B Okyanus verisi kullanmış. Bu veri ile deniz yüzeyindeki

sıcaklık datasını tahmin etmeye çalışmıştır. Çalışmalarında ulaştığı hata oranı %11'dir. Bu çalışmanın asıl amacı büyük verilerde daha hızlı tahmin yapmayı sağlamaktır.

Wang [25] çalışmasında, kendi yaptığı hava durumu veri toplama sistemi ile topladığı sıcaklık, nem, rüzgâr hızı , rüzgâr yönü datalarını kullanarak kendini eğitip bir sonraki gün için tahmin yürütmüştür . RMSE %5'tir.

Gerçekleştirilen bu çalışmada Nandar [4], Rahul [5], Anbarasi [24] ve Wang'dan [25] farklı olarak Radyosonde rasatlarından elde edilen 3-boyutlu 3 yıllık veriler kullanılarak makine öğrenme algoritmaları eğitilerek hava durumu kestirimi gerçekleştirilmiştir. Bu sayede daha önce literatürde gerçekleştirilen çalışmalardan daha düşük hata oranı ile hava durumu kestirimi gerçekleştirilmiştir.



7. SONUÇ

Dünya genelinde hava tahminlerinin yapılp, bu verilere göre önlemler alınması çok önemlidir. Bu verilerin ölçülememesi veya yanlış ölçülmeleri, ülke ekonomilerine çok büyük zararlar verebilmektedir. Uçakların piste inememesinden, savaş uçakları/tanklar/topçu atışlarının hedefi tutturamamalarına veya şehir içi bölgelerde oluşacak sel ve dolu gibi doğal afetlere önlem alınamamalarına kadar birçok farklı problem yaratabilmektedir. Bu etkiler, ülkemizin ekonomisine milyarlarca dolarlık zarara uğratabilme ihtimali yaratmaktadır.

Hava tahmininin gerçekleştirilebilmesi için atmosferin tüm değişkenlerinin saptanması ve atmosfer dönemeçlerinin zaman içerisinde nasıl değiştiğinin anlaşılması gerekmektedir. Atmosferin tüm değişkenlerinin saptanması ve tüm değişimlerinin bilinmesi mümkün değildir. Gelişen teknoloji sayesinde ortaya çıkan farklı rasat yöntemleri sayesinde hassas hava tahmini mümkün olmaya başlamıştır.

Bu çalışmada kullanılan Radyosonde Meteoroloji Ölçüm cihazı, içerisinde sıcaklık, basınç, nem ve GPS sensörleri barındırır. Üzerindeki sensörleri kullanarak o anlık sıcaklık, yüzde bağıl nem, rüzgâr hızı ve rüzgâr yönü değerlerini ölçerek, üzerinde bulunan anteni ve göndermeç devresi vasıtasıyla yer istasyonuna gönderir. Yer İstasyonuna iletilen meteorolojik veriler işlenerek, Dünya Meteoroloji Organizasyonu (DMO) veri formatları olan TEMP35 ve BUFR kod gibi formatlara çevrilir. Ardından elde edilen rasatlar DMO'ya gönderilerek yaklaşık atmosferik model oluşturulur. Bu model kullanılarak istatistiksel tabanlı yöntemlerle hava tahmini gerçekleştirilir. Ancak, atmosfer kaotik bir yapıda olduğu için istatistiksel yöntemler yeterli olmadığından bu çalışma kapsamında Makine Öğrenmesi Algoritmaları tabanlı kestirimler kullanılmıştır.

Literatürde daha önce Makine Öğrenmesi Algoritmaları tabanlı kestirimler [2] bulunuyor olsa da gerçekleştirilen çalışmalar yer gözlem istasyonlarından elde edilmiş tek boyutlu ve sabit irtifaya ait ölçümlere dayanmaktadır.

Bu çalışmada literatürdeki diğer çalışmalardan farklı olarak 3 yıllık Radyosonde rasatları kullanılarak oluşturulan 3-boyutlu atmosfer modeli ile Makine Öğrenmesi Algoritmaları eğitilmiş ve eğitilen bu modeller ile ertesi güne ait en yüksek ve en düşük sıcaklık kestirilmeye çalışılmıştır. Kestirim algoritmalarının başarıları incelenmiş içlerinden bu veri yapısına en uygun regresyon yöntemi olarak Gauss Süreci Regresyon belirlenmiştir.

Ayrıca Radyosonde verilerine, normalizasyonun etkisi incelenmiş olup normalize edilmesinin başarıyı düşürdüğü gözlenmiştir. Bu verideki bazı özneliklerin daha önemli olduğu sonucunu ortaya çıkarmıştır. Basınç değeri diğer öz niteliklere göre yüksek değerde olup sıcaklık tahmininde daha etkin olduğu gözlemlenmektedir. Bunların yanı sıra en yüksek sıcaklık ve en düşük sıcaklık tahmini için ayrı modellerin kullanılmasının da başarıyı arttırdığı tespit edilmiştir. Tüm analizler sonucunda önerilen yöntem ile 1,2 Ortalama Karekök Sapması değeriyle ertesi güne ait en yüksek sıcaklık 2,4 Ortalama Karekök Sapması değeri ile en düşük sıcaklık için kestirilmiştir. Gündüz sıcaklık tahminindeki hata oranı ile gece sıcaklıktaki hata oranı arasındaki farkın nedeni gündüz rasat veri sayısının gece rasat veri sayısına göre fazla olmasından kaynaklanmaktadır.

İlerleyen çalışmalarda daha fazla gün içeren Radyosonde verileri kullanılarak derin öğrenme algoritmalarıyla kestirim yapılarak hava tahmini hata oranı azaltılabilir. Ayrıca, bu çalışmada Ankara ilinde gerçekleştirilen 3 yıllık Radyosonde rasatları kullanılmıştır. İlerleyen çalışmalarda tüm ülke genelinde veya Dünya genelindeki geçmişe yönelik bütün Radyosonde rasatları birleştirilerek eğitimler gerçekleştirilebilir. Böylelikle hava tahminindeki hata oranı daha da azaltılabilir.

8. KAYNAKLAR

- [1] K. L. M. D. Sobrevilla, A. G. Quiñones, K. V. S. Lopez, and V. T. Azaña, “Daily weather forecast in Tiwi, Albay, Philippines using Artificial Neural Network with missing values Imputation,” *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, pp. 2981–2985, 2017, doi: 10.1109/TENCON.2016.7848592.
- [2] F. Luis and G. Moncayo, “Meteorology today aN INtroDUCTIoN to WeatHer, CLIMate, aNd tHe eNViRoNMeNt,” *Meteorol. today aN Introd. to Weather. Clim. Environ.*, p. 662, 2014.
- [3] D. N. Fente and D. Kumar Singh, “Weather Forecasting Using Artificial Neural Network,” *Proc. Int. Conf. Inven. Commun. Comput. Technol. ICICCT 2018*, no. Icicct, pp. 1757–1761, 2018, doi: 10.1109/ICICCT.2018.8473167.
- [4] A. Nandar, “Bayesian network probability model for weather prediction,” *Proc. 2009 Int. Conf. Curr. Trends Inf. Technol. CTIT 2009*, pp. 120–124, 2009, doi: 10.1109/CTIT.2009.5423132.
- [5] G. K. Rahul, S. Singh, and S. Dubey, “Weather Forecasting Using Artificial Neural Networks,” *ICRITO 2020 - IEEE 8th Int. Conf. Reliab. Infocom Technol. Optim. (Trends Futur. Dir.*, no. Icicct, pp. 21–26, 2020, doi: 10.1109/ICRITO48877.2020.9197993.
- [6] “Sıkça Sorulan Sorular - Meteoroloji Genel Müdürlüğü.” mgm.gov.tr/genel/sss.aspx?s=havatahmini (accessed Jul. 27, 2022).
- [7] G. K. Rahul, S. Singh, and S. Dubey, “Weather Forecasting Using Artificial Neural Networks,” *ICRITO 2020 - IEEE 8th Int. Conf. Reliab. Infocom Technol. Optim. (Trends Futur. Dir.*, pp. 21–26, 2020, doi: 10.1109/ICRITO48877.2020.9197993.
- [8] Vaisala, “Vaisala - a global leader in environmental and industrial measurement |.” vaisala.com/en (accessed Aug. 11, 2022).
- [9] “Home - Meteomodem.” meteomodem.com/ (accessed Aug. 11, 2022).
- [10] “Meisei Electric.” meisei.co.jp/english/ (accessed Aug. 11, 2022).
- [11] “Graw Radiosondes | NORIS Group GmbH.” graw.de/ (accessed Aug. 11, 2022).
- [12] “Magnetron Savunma ve Uzay A.S.” magnetrontr.com/ (accessed Aug. 12, 2022).
- [13] S. E. Haupt, J. Cowie, S. Linden, T. McCandless, B. Kosovic, and S. Alessandrini, “Machine learning for applied weather prediction,” *Proc. - IEEE 14th Int. Conf. eScience, e-Science 2018*, pp. 276–277, 2018, doi: 10.1109/eScience.2018.00047.

- [14] L. Vinet and A. Zhedanov, "A 'missing' family of classical orthogonal polynomials," *J. Phys. A Math. Theor.*, vol. 44, no. 8, p. 7728, 2011, doi: 10.1088/1751-8113/44/8/085201.
- [15] S. Kavitha, S. Varuna, and R. Ramya, "A comparative analysis on linear regression and support vector regression," *Proc. 2016 Online Int. Conf. Green Eng. Technol. IC-GET 2016*, 2017, doi: 10.1109/GET.2016.7916627.
- [16] H. Yang, L. Chan, and I. King, "Support vector machine regression for volatile stock market prediction," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 2412, pp. 391–396, 2002, doi: 10.1007/3-540-45675-9_58.
- [17] L. J. Cao and F. E. H. Tay, "Support vector machine with adaptive parameters in financial time series forecasting," *IEEE Trans. Neural Networks*, vol. 14, no. 6, pp. 1506–1518, 2003, doi: 10.1109/TNN.2003.820556.
- [18] U. Jain, K. Nathani, N. Ruban, A. N. J. Raj, Z. Zhuang, and V. G. V. Mahesh, "Cubic SVM classifier based feature extraction and emotion detection from speech signals," *Proc. - 2018 Int. Conf. Sens. Networks Signal Process. SNSP 2018*, pp. 386–391, 2019, doi: 10.1109/SNSP.2018.00081.
- [19] P. Viridi, Y. Narayan, P. Kumari, and L. Mathew, "Discrete Wavelet Packet based Elbow Movement classification using Fine Gaussian SVM," *1st IEEE Int. Conf. Power Electron. Intell. Control Energy Syst. ICPEICES 2016*, 2017, doi: 10.1109/ICPEICES.2016.7853657.
- [20] K. W. Yan, "Study on the forecast of air passenger flow based on SVM regression algorithm," *Proc. - 2009 1st Int. Work. Database Technol. Appl. DBTA 2009*, pp. 325–328, 2009, doi: 10.1109/DBTA.2009.33.
- [21] M. H. . R. Hadi Asheri , Abdolkhalegh Bayati , Hamid R . Rabiee, "MOTION VECTOR RECOVERY WITH GAUSSIAN PROCESS REGRESSION," pp. 953–956, 2011.
- [22] J. Eberhard and V. Geissbuhler, *Gaussian Processes for Machine Learning*, vol. 7, no. 5. 2006.
- [23] Y. Yağan, "Ravinsonde Rasatları Kurs Notları," 2007.
- [24] J. D. Anbarasi and V. Radha, "Review on Marine Weather Forecasting with Big data," no. Icaaic, pp. 783–790, 2022, doi: 10.1109/icaaic53929.2022.9792644.

- [25] Y. Wang, K. Ni, X. Wang, and J. Zhu, "Design of Automatic Weather Monitoring and Forecasting System based on Internet of Things and Big Data," *Int. Conf. Sustain. Comput. Data Commun. Syst. ICSCDS 2022 - Proc.*, no. 978, pp. 979–982, 2022, doi: 10.1109/ICSCDS53736.2022.9761041.
- [26] S. Ackerman and J. Knox, "Meteorology: Understanding the Atmosphere, Fourth Edition," pp. 244–270, 2015.



EK 1 Örnek Tam bir TEMP KODU

USTT66 ANKA 010000

TTAA 51001 17130 99915 09740 04003 00195 // // 92802

// // 85456 11949 03008 70904 23344 01016 50530 37561

34011 40680 47163 30010 30871 47774 30019 25992 45779 28532

20142 43981 28532 15334 46181 27037 10600 49781 28524 88395

47564 30511 77999 31313 73508 80015=

UKTT66 ANKA 010000

TTBB 51008 17130 00915 09740 11914 09356 22906 08350 33830

12950 44752 19523 55696 23746 66645 28139 77593 30559 88524

34762 99483 39360 11443 42164 22395 47564 33394 47565 44363

46773 55317 48372 66235 44779 77187 43182 88130 48781 99100

49781 21212 00915 04003 11694 00516 22520 34512 33461 30509

44413 29010 55383 31512 66353 32512 77282 29524 88136 27039

99134 27039 11100 28524 31313 73508 80015 41414 00900 51515

11803 02512 22800 02012 33600 35516=

ULTT66 ANKA 010000

TTCC 51001 17130 70834 53782 27015 50047 56981 29008 30370

56182 05507 20630 51384 00509 10085 43386 26511 88608 57381

28017 77999 31313 73508 80015=

UETT66 ANKA 010000

TTDD 5100/ 17130 11756 50982 22608 57381 33505 56781 44471

58981 55379 58182 66369 54782 77302 55982 88260 52583 99223

53783 11205 51784 22168 51984 33136 46985 44110 45785 55070

38387 21212 11760 28022 22682 26515 33558 29516 44517 30011

55447 26006 66428 25008 77387 27007 88356 30504 99337 36004

11315 04007 22303 05507 33268 05509 44234 02009 55204 00509

66182 01011 77163 35015 88143 34014 99127 34508 11123 34006

22115 28505 33111 25506 44110 25006 55106 24508 66094 29012

77086 30517 88080 31021 99070 29521 31313 73508 80015=