

**SEYAHAT SÜRESİ DEĞİŞKENLİĞİNİN ANALİZİ VE
YAPAY SİNİR AĞLARI İLE SEYAHAT SÜRESİNİN
TAHMİNİ**

**ANALYSIS OF TRAVEL TIME VARIABILITY AND
PREDICTION OF TRAVEL TIME USING ARTIFICIAL
NEURAL NETWORKS**

SİBEL SERTOK

PROF. DR ÖZGÜR YENİAY

Yardımcı Tez Danışmanı: DR. ÖĞRETİM ÜYESİ BERK ANBAROĞLU

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

İstatistik Anabilim Dalı için Öngördüğü

YÜKSEK LİSANS TEZİ olarak hazırlanmıştır.

2022

ÖZET

SEYAHAT SÜRESİ DEĞİŞKENLİĞİNİN ANALİZİ VE YAPAY SİNİR AĞLARI İLE SEYAHAT SÜRESİNİN TAHMİNİ

Sibel Sertok

Yüksek Lisans, İstatistik Bölümü

Tez Danışmanı: Prof. Dr. Özgür YENİAY

Eş Danışman: Dr. Öğr. Üyesi Berk ANBAROĞLU

Haziran 2022, 64 sayfa

Seyahat sürelerinin analizi, insanların seyahat davranışlarını belirlemesine yardımcı olmaktadır. Bir yerden bir yere giderken daha önce yapılan seyahat sürelerinin değişkenliğini ve güvenilirliğini bilmek zamanın iyi kullanılmasını sağlayabilir. Bu çalışmada New York şehrine ait La Guardia havalimanı ile John F. Kennedy havalimanları arasında taksi ile yapılan yolculukların seyahat sürelerinin değişkenliği ve güvenilirliği araştırılmıştır. Çalışmada ayrıca seyahat süresini tahmin etmek için iki yöntem kullanılmıştır. Bunlardan biri sadece seyahat süresini kullanan doğrusal olmayan otoregresif model (nonlinear autoregressive model-NAR) yöntemi, diğeri ise seyahat süresini etkileyen faktörlerden sıcaklık, görüş mesafesi ve rüzgar değişkenleri de kullanan doğrusal olmayan otoregresif dışsa model (nonlinear autoregressive model with external input-NARX) yöntemidir. İki yöntem içinde farklı eğitim seçenekleri uygulanmış ve her iki yöntem için de en iyi model 'sgdm' eğitim seçeneğinde bulunmuştur.

Anahtar Kelimeler: Seyahat süresi değişkenliği, Seyahat süresi güvenilirliği, Güvenilirlik ölçütleri, Yapay sinir ağları, NAR, NARX.

ABSTRACT

ANALAYSIS OF TRAVEL TIME VARIABILITY AND PREDICTION OF TRAVEL TIME USING ARTIFICIAL NEURAL NETWORKS

Sibel Sertok

Post graduate, Department of Statistics

Supervisor: Prof. Dr. Özgür Yeniay

Co- Supervisor: Dr. Öğr. Üyesi Berk Anbaroğlu

June 2022, 64 pages

Analysis of travel times helps people determine their travel behavior. Knowing the variability and reliability of previous travel times when going from one place to another can make good use of time. In this study, the variability and reliability of the travel times of the journeys made by taxi between the La Guardia airport of New York City and the John F. Kennedy airports were investigated. Two methods were also used in the study to estimate the travel time. One of them is the nonlinear autoregressive model -NAR method, which uses only the travel time, and the other is the nonlinear autoregressive model with external input -NARX, which also uses temperature, visibility and wind variables, which are the factors affecting the travel time. Different training options were applied in the two methods and the best model for both methods was found in the 'sgdm' training option.

Keywords: Travel time variability, Travel time reliability, Reliability metrics, Artificial neural networks, NAR, NARX.

TEŐEKKÜR

Bu tezde arařtırmalarımın her ařamasında ilgi ve önerileriyle beni yönlendiren, alıřmaktan mutluluk duyduğum değerli danışman hocam Sayın Prof. Dr. Özgür YENİAY'a, eş danışmanlığımı kabul edip benden desteğini esirgemeyen, değerli hocam Sayın Dr. Öğretim Üyesi Berk ANBAROĞLU'na en içten saygı ve minnetlerimi sunarım. Bana her konuda destek olan, her zaman koşulsuz inanan, beni bugünlere kadar getiren canım aileme en içten teşekkürlerimi sunarım.

İÇİNDEKİLER

İçindekiler

ÖZET	i
ABSTRACT.....	ii
TEŞEKKÜR.....	iii
İÇİNDEKİLER	iv
ŞEKİLLER DİZİNİ	vii
ÇİZELGELER DİZİNİ	ix
KISALTMALAR.....	x
1. GİRİŞ.....	1
2. SEYAHAT SÜRELERİNİN DEĞİŞKENLİĞİ VE GÜVENİLİRLİĞİ	3
2.1. Seyahat Süresi Dağılımları	3
2.1.1. Log-Normal Dağılım	3
2.1.2. Log-Normal (3P) Dağılımı	4
2.1.3. Weibull Dağılımı	4
2.1.4. Log-Lojistik Dağılımı	5
2.1.5. Log-Lojistik (3P) Dağılımı	5
2.1.6. Gamma Dağılımı.....	6
2.1.6. Burr Dağılımı	6
2.2. Seyahat Süresi Güvenilirliği	7
2.2.1. Serbest Akış Seyahat Süresi	8
2.3. Seyahat Süresi Dağılımı İle İlgili Literatür İncelemesi	9
2.4. Seyahat Süresi Güvenilirliği İle İlgili Literatür İncelemesi.....	10
3. SEYAHAT SÜRESİ TAHMİN YÖNTEMLERİ	12
3.1. Model Tabanlı Yöntemler.....	12
3.2. Veri Tabanlı Yöntemler	13
3.2.1. Parametrik Yöntemler.....	13
3.2.2. Parametrik Olmayan Yöntemler	13

4. YAPAY SİNİR AĞLARI	15
4.1. YSA ve Biyolojik Sinir Hücresinin İlişkisi	15
4.2. YSA'nın Katmanları	16
4.3. YSA ve Geleneksel Algoritmalar	19
4.4. YSA'ların Sınıflandırılması	20
4.4.1. YSA Öğrenme Algoritmaları:	20
4.5. Öğrenme Zamanına Göre Sinir Ağları	20
4.6. Öğrenme Kuralları	20
4.7. Yapısına Göre Yapay Sinir Ağları	22
4.8. Tek Katmanlı Algılayıcılar	23
4.9. Basit Algılayıcı Model	24
4.10. Adaline/Madeline Modeli	24
4.11. Çok Katmanlı Algılayıcılar	25
4.12. ÇKA'da Dikkat Edilmesi Gereken Noktalar	27
4.13. YSA'nın Avantajları	27
4.14. YSA'nın Dezavantajları	28
4.17. Doğrusal Olmayan Otoregresif Ağ (Nonlinear Autoregressive Model-NAR)..	28
4.18. Doğrusal Olmayan Dışsal Girdili Otoregresif Ağ (Nonlinear Autoregressive Model with External Input-NARX)	29
4.19. NARX Ağ Tasarımı	30
4.19.1. Ara Katman ve Nöron Sayıları	31
4.19.2. Aktivasyon Fonksiyonunun Seçimi	31
4.19.3. Verilerin Normalizasyonu	31
4.19.4. Eğitim Algoritması	32
4.19.5. Performans Ölçütleri	33
4.20. NARX ve NAR Ağlarının Uygulama Alanları	34
5. Seyahat süresi güvenilirlik analizi ve NAR ve NARX sinir ağları ile tahmini	37
5.1. Çalışmanın Amacı	37
5.2. Verinin Hazırlanması	37
5.3. Seyahat Sürelerinin Dağılım Analizleri	40
5.4. Seyahat Sürelerinin Güvenilirlik Analizleri	42
5.4.1. Seyahat Sürelerinin 95. Yüzdelik Sonucu	42

5.4.2. Tampon Süre İndeksi Sonucu.....	43
5.4.3. Planlanan Süre İndeksi Sonucu.....	43
5.4.4. Başarısızlık ve Zamanında Oranı Sonucu.....	45
5.5. Yapay Sinir Ağları ile Seyahat Süresi Tahmini.....	47
5.5.1. NAR-LM.....	47
5.5.2. NARX-LM.....	55
6. SONUÇLAR VE TARTIŞMA	63
KAYNAKLAR	65
EKLER.....	70
EK1: Uyum İyiliği Testleri	70

ŞEKİLLER DİZİNİ

Şekil 2.1 Seyahat sürelerinin dağılımını belirleme aşamaları.....	3
Şekil 3.1. Seyahat süresi tahmin yöntemleri	12
Şekil 4.1. Biyolojik sinir hücresinin gösterimi	16
Şekil 4.2. Yapay sinir ağı yapısı	17
Şekil 4.3. İleri beslemeli yapay sinir ağı.....	22
Şekil 4.4. Geri beslemeli yapay sinir ağları	23
Şekil 4.5. And, or ve xor problemlerinin eksen üzerinde gösterimi	25
Şekil 4.6. İleri doğru hesaplama	26
Şekil 4.7. Geriye doğru hesaplama	26
Şekil 4.8. NARX ağının yapısı	29
Şekil 4.9. Seri paralel mimari	30
Şekil 4.10. Paralel mimari.....	30
Şekil 5.1. JFK ve LaGuardia ve 23 Aralık 2015 günü yapılan yolculukların başlangıç noktaları	39
Şekil 5.1. Veri tabanı görüntüsü 1	39
Şekil 5.2. Veri tabanı görüntüsü 2	39
Şekil 5.3. Seyahat süresi grafiği.....	40
Şekil 5.4. Pazartesi günü 12:00-12:15 verisi için Easy fit programının görüntüsü	41
Şekil 5.5. Pazartesi günü dağılımı	41
Şekil 5.6. Haftanın her günü için dağılımların yüzdeleri.....	42
Şekil 5.7. Özel günlerin yolculuk sayıları.....	44
Şekil 5.8. Başarısızlık oranı grafiği	45
Şekil 5.9. Tampon süresi ve planlanan süre seyahat süresi ile gösteren grafik	46
Şekil 5.10. LM algoritmasının nöron sayısını gösteren grafik	47
Şekil 5.11. NAR YSA açık döngü mimarisi	49
Şekil 5.12. NAR YSA kapalı döngü mimarisi.....	49
Şekil 5.13. NAR ağı hataların otokorelasyonunu gösteren grafik	52
Şekil 5.14. NAR ağına ait zaman serisi için çıktı elemanlarının yanıtlarını gösteren grafik	52
Şekil 5.15. NAR regresyon katsayısını gösteren grafikler.....	53
Şekil 5.16. NAR ağı tahmini değerleri gösteren grafik	54

Şekil 5.17. Sıcaklık, görüş mesafesi ve rüzgâr değişkenlerinin seyahat süresi ile korelasyonu.....	55
Şekil 5.18. NARX YSA açık döngü mimarisi	56
Şekil 5.19. NARX YSA kapalı döngü mimarisi.....	56
Şekil 5.20. NARX ağı hataların otokorelasyonunu gösteren grafik	60
Şekil 5.21. NARX ağına ait zaman serisi için çıktı elemanlarının yanıtlarını gösteren grafik.....	60
Şekil 5.22. NARX regresyon katsayısını gösteren grafikler.....	61
Şekil 5.23. NARX ağı tahmini değerleri gösteren grafik	62

ÇİZELGELER DİZİNİ

Çizelge 2.1. Çeşitli ulaşım kurumlarının kullandığı seyahat süresi güvenilirliği ölçütleri	9
Çizelge 4.1. Sinir hücresi ile yapay sinir ağlarının bölümlerinin karşılaştırılması	16
Çizelge 4.2. Aktivasyon fonksiyonu türleri	18
Çizelge 4.3. Normalizasyon Yöntemleri.....	32
Çizelge 4.4. Performans ölçütleri	34
Çizelge 5.1. NAR ağına ait adam eğitim seçeneğinin sonuçları.....	49
Çizelge 5.2. NAR ağına ait rmsprop eğitim seçeneğinin sonuçları	50
Çizelge 5.3. NAR ağına ait sgd eğitim seçeneğinin sonuçları.....	51
Çizelge 5.4. NARX ağına ait adam eğitim seçeneğinin sonuçları.....	57
Çizelge 5.5. NARX ağına ait rmsprop eğitim seçeneğinin sonuçları	58
Çizelge 5.6. NARX ağına ait sgd eğitim seçeneğinin sonuçları.....	59

KISALTMALAR

BTI: Tampon Süre İndeksi

FR: Başarısızlık Oranı

GPS: Küresel Konumlandırma Sistemi

JFK: John F. Kennedy

MI: Sefalet İndeksi

NAR: Doğrusal Olmayan Otoregresif Model

NARX: Doğrusal Olmayan Otoregresif Dışsal Model

MAE: Hatanın Mutlak Ortalaması

RMSE: Hata Kareler Ortalamasının Karekökü

MSE: Hata Kareler Ortalaması

OTR: Zamanında Oranı

PI: Planlanan Süre İndeksi

TT: Seyahat Süresi

\overline{TT} : Ortalama Seyahat Süresi

TT_f : Serbest Akış Seyahat Süresi

TTI: Seyahat Süresi İndeksi

TT^{95} :Seyahat süresinin 95. yüzdelik değeri

USDOT: Amerika Birleşik Devletleri Ulaştırma Bakanlığı

YSA: Yapay Sinir Ağları

1. GİRİŞ

Dünya nüfusunun artmasıyla birlikte trafik problemleri de çoğalmıştır. Trafik, ülkelerin ekonomisini, insanların refah düzeyini etkilemektedir. Dolayısıyla trafiğin iyi yönetilmesi ekonomik kazanç sağlarken insanların daha huzurlu bir şekilde yaşamasına da yardımcı olur. Trafiğin iyi yönetilmemesi bazen düzeltilebilecek sebepler barındırırken bazen de beklenmedik olaylar içerebilir. Trafikteki tıkanıklık birçok nedenden kaynaklanabilir: yolların dar olması, trafik çalışmaları sebebiyle şerit kapanmalarının olması, kar yağışı, özel bir organizasyonun olması gibi. Tüm bu nedenler seyahat süresini etkilemektedir. Seyahat süresi, zamana ve mekana bağlı olarak değişkenlik göstermektedir. Seyahat süresinin analiz edilmesi hem ülkeler hem de seyahat eden insanlar için önemlidir. Literatürde seyahat süresinin analizinde olasılık dağılımlarının belirlendiği ve güvenilirliğin ölçüldüğü çok sayıda çalışma bulunmaktadır. Seyahat süresi güvenilirliğinin birden fazla tanımı vardır. Örneğin toplu ulaşım araçlarının planlanan sürede hedeflerine ulaşması durumu ya da bir yoldaki trafik sıkışıklığının belirsizlik düzeyi olarak tanımlanabilir. Diğer bir ifadeyle seyahat süresi güvenilirliği, seyahat süresinin zaman içinde ne kadar değiştiğini gösterir. Sözü edilen seyahat süresini etkileyen faktörler dolayısıyla seyahat süresi güvenilirliğini de etkilemektedir. Bu faktörler her gün gerçekleşmediği için güvenilir olmayan seyahat süresi üretmeye de elverişlidir. Seyahat eden insanlar, bu tür belirsizliklerin seyahat sürelerine etkilerini bildiklerinde, ulaşım aracını, çıkış saatini ve rota seçimini daha etkin bir şekilde belirleyebilirler.

Seyahat süresinin güvenilirliği doğrudan ölçülemediğinden, seyahat süresi verisine ihtiyaç duyulur. Seyahat süresi bir yolculuğun başladığı ve bittiği nokta arasında geçen süredir [1]. Seyahat süresinin güvenilirliğinin analizi için büyük ulaşım verisine ihtiyaç duyulmaktadır. Seyahat süresini ölçme araçları arasında Küresel Konumlandırma Sistemi (Global Positioning System-GPS), Bluetooth sensörler, araç plakası eşleştirme, hava fotogrametresi ile donatılmış araçlar yer alır. GPS, seyahat süresi verilerini toplamada en yaygın kullanılan yöntemdir. Birçok aracın anlık konumu ve hızı GPS aracılığıyla kolaylıkla toplanabilmektedir. Ancak bazı durumlarda maliyetli olabilir. Örneğin her gün yüz binlerce taksi yolculuğunun yapıldığı bir şehirde, GPS alıcısı yerleştirilmiş her aracın (ör. taksi) sabit aralıklarla (ör. dakikada bir) izlenmesi çok maliyetli bir işlemdir.

Dolayısıyla, arařtırmacılar daha kısıtlı veri setleri ile alıřabilecekleri yöntemler üzerine yoğunlařmıřlardır. Bu tür veri setlerinde yolculukların bařlangı ve bitiř noktalarına ek olarak yolcu sayısı, yolculuğun uzunluęu ve bařlangı/bitiř zamanı gibi öznitelikler de bulunmaktadır. New York Őehrine ait taksi veri seti de bu alanda bilinen en deęerli aık veri setlerinden biridir [2].

Bu alıřmada, New York Őehrindeki John F. Kennedy (JFK) ve La Guardia havalimanları arasındaki seyahat sürelerinin önce güvenilirlięi ve daęılımı arařtırılmıř daha sonra da yapay sinir aęlarıyla (YSA) seyahat süresi tahmini yapılmıřtır. Bu iki havalimanı arasında aktarmalı uuřu olan yolcuların trafik düzensizlięinden sıkıntı ekmemeleri iin farklı gün ve saatlerdeki seyahat süresi güvenilirlięi dört farklı indeks yardımıyla deęerlendirilmiřtir. Amerika Birleřik Devletleri Ulařtırma Bakanlıęının (USDOT) önerdięi bu indeksler: i) Seyahat Süresi İndeksi (TTI= Travel Time Index) ii) Tampon Süre İndeksi (BTI= Buffer Time Index) iii) Planlanan Süre İndeksi (PI= Planning Time Index) ve iv) Bařarısızlık ve zamanında oranıdır (FR= Failure Rate and OTR= On Time Rate). alıřmada, haftanın yedi günü 15 dakikalık zaman dilimlerine bölünmüř olup, bu süreçlerde gerekleřtirilen yolculukların tampon süre endeksi, planlanan süre endeksi, bařarısızlık ve zamanında oranı ile seyahat süresi güvenilirlięi analiz edilmiřtir. Seyahat süresi güvenilirlięi ile doęrudan iliřkili bir kavram olan seyahat süresi daęılımı tespit edilmiř ve güvenilirlięi arařtırılmıřtır.

Tezin ikinci bölümünde, seyahat süresinin tahmin yöntemlerinden YSA kullanılmıřtır. YSA iinde zaman serisi ve dıřsal girdi bulunduran doęrusal olmayan otoregresif dıřsal model (nonlinear autoregressive model with external input-NARX) yöntemi ve yine zaman serisi bulunduran tek deęiřkenden yararlanan doęrusal olmayan otoregresif model (nonlinear autoregressive model-NAR) yöntemleri kullanılarak daha sonraki dönemlerdeki seyahat süresinin tahmini iin iki havalimanı arasında yapılan yolculuk süreleri, yolculukların yapıldıęı zamanlardaki hava durumu ve görüř mesafesi bilgileriyle bir model oluřturulmuřtur. Model oluřturulurken sinir aęlarının eęitimi iin farklı opsiyonlar, gecikme sayıları ve eęitim, test doęrulama oranları kullanılmıřtır.

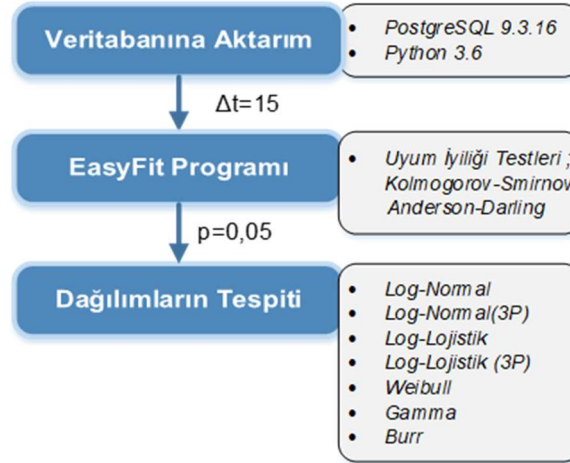
2. SEYAHAT SÜRELERİNİN DEĞİŞKENLİĞİ VE GÜVENİLİRLİĞİ

Bu bölümde, iki havalimanı arasındaki seyahat sürelerinin olasılık dağılımlarından ve seyahat sürelerinin güvenilirliğinin ölçümünde kullanılan seyahat sürelerinin 95. yüzdelikleri, tampon süre indeksi, planlanan süre indeksi ve başarısızlık oranlarından bahsedilecektir.

2.1. Seyahat Süresi Dağılımları

Seyahat süresinin dağılımı değişkenliğin istatistiksel olarak yapısını ve düzenini ortaya koyar. Bu nedenle bu kısımda haftanın her günü için on beş dakikalık zaman aralıklarında, JFK havalimanından başlayıp, LaGuardia havalimanına gitmiş tüm taksi yolculuklarının seyahat süreleri toplanmış ve bu seyahat sürelerinin dağılımı araştırılmıştır.

Seyahat sürelerinin dağılımının nasıl tespit edildiğini gösteren süreç Şekil 2.1.'de verilmiştir.



Şekil 2.1. Seyahat sürelerinin dağılımını belirleme aşamaları

Seyahat sürelerinin dağılımı için literatürde sıklıkla kullanılan dağılımların bazıları aşağıda verilmiştir.

2.1.1. Log-Normal Dağılım

Normal dağılımın dönüştürülmesinden elde edilen Log-Normal dağılım sağa çarpık bir dağılımdır ve pozitif değerler alır. Dağılıma ilişkin olasılık yoğunluk fonksiyonu ile birikimli dağılım fonksiyonu sırasıyla eşitlik (2.1) ve eşitlik (2.2)'de verilmiştir.

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2\right), \sigma>0, -\infty<\mu<\infty \quad (2.1)$$

$$F(x) = \Phi\left(\frac{\ln(x)-\mu}{\sigma}\right) \quad (2.2)$$

Log-Normal dağılımda, Φ standart normal birikimli dağılım fonksiyonu, μ ölçüm parametresi ve σ şekil parametresidir.

2.1.2. Log-Normal (3P) Dağılımı

Log-Normal (3P) dağılım, Log-Normal dağılımın üç parametrelidir. Log-Normal (3P)'nin olasılık yoğunluk fonksiyonu ile birikimli dağılım fonksiyonu sırasıyla eşitlik (2.3) ve eşitlik (2.4)'te verilmiştir:

$$f(x) = \frac{1}{(x-\gamma)\sigma\sqrt{2\pi}} \exp\left\{-\frac{[\ln(x-\gamma)-\mu]^2}{2\sigma^2}\right\}, \sigma>0, -\infty<\mu<\infty \quad (2.3)$$

$$F(x) = \Phi\left(\frac{[\ln(x-\gamma)-\mu]}{\sigma}\right) \quad (2.4)$$

Burada, μ ölçüm parametresi, σ şekil parametresi, γ konum parametresi ve Φ standart normal bir değişkenin birikimli dağılım fonksiyonunu göstermektedir.

2.1.3. Weibull Dağılımı

Sürekli bir dağılım olan Weibull dağılımı, pozitif değerler alır. Weibull dağılımının olasılık yoğunluk fonksiyonu ile birikimli dağılım fonksiyonu sırasıyla eşitlik (2.5) ve (2.6)'da belirtilmiştir.

$$f(x) = \frac{kx^{k-1}}{\lambda^k} \left(\exp\left(-\frac{x^k}{\lambda}\right) \right), \lambda>0, k>0 \quad (2.5)$$

$$F(x) = 1 - \exp\left(-\frac{x}{\lambda}\right)^k \quad (2.6)$$

Burada, k şekil parametresi, λ ölçek parametresidir.

2.1.4. Log-Lojistik Dağılımı

Log-Lojistik dağılımı, Weibull dağılımına alternatif bir dağılımdır. İki parametrelili olan Log-Lojistik dağılımı sürekli ve çarpık bir dağılımdır. Log-Lojistik dağılımının olasılık yoğunluk fonksiyonu ve birikimli dağılım fonksiyonu sırasıyla eşitlik (2.7) ve (2.8)'de verilmiştir:

$$f(x) = \frac{\left(\frac{\mu}{\sigma}\right)\left(\frac{x}{\sigma}\right)^{\mu-1}}{\left(1+\left(\frac{x}{\sigma}\right)^\mu\right)^2}, \sigma > 0 \quad (2.7)$$

$$F(x) = \frac{x^\mu}{\sigma^\mu + x^\mu} \quad (2.8)$$

Burada, μ şekil parametresi, σ ölçek parametresidir.

2.1.5. Log-Lojistik (3P) Dağılımı

Log- Normal (3P)'de olduğu gibi bu dağılımda da konum parametresi (γ) bulunmaktadır. Örnek olarak pazartesi 12:30-12:45 arası da bu dağılıma uymaktadır. Log-Lojistik (3P)'nin olasılık yoğunluk fonksiyonu ile birikimli dağılım fonksiyonu sırasıyla eşitlik (2.9) ve (2.10)'da verilmiştir:

$$f(x) = \frac{\alpha}{\beta} \left(\frac{x-\gamma}{\beta}\right)^{\alpha-1} \left(1 + \left(\frac{x-\gamma}{\beta}\right)^\alpha\right)^{-2}, \beta > 0, x > \gamma, \alpha \geq 1 \quad (2.9)$$

$$F(x) = \left(1 + \frac{\beta}{x-\gamma}\right)^{-1} \quad (2.10)$$

Burada, α ölçüm parametresi, β şekil parametresi, γ ise konum parametresidir.

2.1.6. Gamma Dağılımı

Gamma dağılımı, pozitif değerlerin analizinde kullanılan sürekli bir dağılımdır. Gamma dağılımının olasılık yoğunluk fonksiyonu ile birikimli dağılım fonksiyonu sırasıyla eşitlik (2.11) ve (2.12)'de verilmiştir:

$$f(x) = \left(\frac{x}{\theta}\right)^{k-1} \frac{\exp\left(-\frac{x}{\theta}\right)}{\theta \Gamma(k)}, k > 0, \theta > 0 \quad (2.11)$$

$$F(x) = \frac{1}{\theta^k \Gamma(k)} \int_{-\infty}^x t^{k-1} \exp\left(-t/\theta\right) dt \quad (2.12)$$

Burada, k şekil parametresi, θ ölçek parametresi ve $\Gamma(k)$ Gamma fonksiyonudur.

2.1.6. Burr Dağılımı

Burr dağılımı, negatif olmayan değerler için sürekli bir dağılımdır. Burr dağılımının olasılık yoğunluk fonksiyonu ile birikimli dağılım fonksiyonu sırasıyla eşitlik (2.13) ve (2.14)'te verilmiştir:

$$f(x) = \frac{\frac{kc}{\alpha} \left(\frac{x}{\alpha}\right)^{c-1}}{\left(1 + \left(\frac{x}{\alpha}\right)^c\right)^{k+1}}, c > 0, k > 0, \alpha > 0 \quad (2.13)$$

$$F(x) = 1 - \frac{1}{\left(1 + \left(\frac{x}{\alpha}\right)^c\right)^k} \quad (2.14)$$

Burada, c ve k şekil parametreleri ve α ölçek parametresidir.

2.2. Seyahat Süresi Güvenilirliği

ABD Ulaştırma Bakanlığı (USDOT), seyahat süresi güvenilirliğini “ulaşım sistemindeki seyahat sürelerinin kesinlik ve öngörülebilirlik derecesi” olarak tanımlamaktadır. Güvenilmez ulaşım sistemleri kullanıcılar için maliyeti artırır ve gecikmelere sebep olur [3].

Seyahat sürelerini etkileyen bazı faktörler şunlardır:

- Dar yollar,
- Trafik kontrol cihazları,
- Hava durumu,
- Trafik çalışmaları,
- Trafik kazaları,
- Özel gün kutlamaları.

Literatürde çeşitli seyahat süresi güvenilirlik ölçütleri bulunmaktadır. ABD’de bazı ulaşım kurumlarının kullandığı seyahat süresi güvenilirlik ölçütleri Çizelge 2.1’de verilmiştir. Bu çalışmada, BTI, PTI, FR ve OTR güvenilirlik ölçütleri kullanılmıştır.

Tampon süre indeksi, seyahat edenlerin gitmek istedikleri yere zamanında varmaları için ortalamaya eklenmesi gereken ek süreyi verir. Bu ek süre beklenmeyen gecikmeleri ifade etmektedir. Tampon süre indeksi, eşitlik (2.15) yardımıyla hesaplanmaktadır:

$$BTI = \frac{TT^{95} - \overline{TT}}{\overline{TT}} \quad (2.15)$$

Eşitlik 2.15 de, TT seyahat süresini, TT^{95} seyahat süresinin 95. yüzdeleri değeri, \overline{TT} ise ortalama seyahat süresini göstermektedir. Bu indeksin süreye çevrilmesi için ortalama değeri ile çarpılması gerekmektedir. Örneğin, ortalama seyahat süresinin 20 dakika olduğu bir yolculukta tampon süre indeksi 0.78 ise, ortalama değeri eklenmesi gereken süre 15.6 olur.

Planlanan süre indeksi, tampon süresi dahil edildiğinde planlanması gereken toplam seyahat süresini belirtir. Tampon süre indeksi, gerekli ek süreyi gösterirken, planlanan süre indeksi gereken toplam seyahat süresini içermektedir. Ayrıca planlanan süre indeksi hem beklenmeyen gecikmeleri hem de tipik gecikmeleri içermektedir.

Bu indeksin hesaplanabilmesi için serbest akış süresinin (TT_f) hesaplanması gerekmektedir. Planlanan süre indeksi eşitlik (2.16) ile hesaplanmaktadır:

$$PTI = \frac{TT^{95}}{TT_f} \quad (2.16)$$

Bu indeksin de süreye çevrilebilmesi için serbest akış seyahat sürelerinin planlanan süre indeksiyle çarpılması gerekmektedir. Örneğin serbest akış süresi 16 dakika, planlanan süre indeksi de 2.68 çıkarsa planlanması gereken süre 43 dakika olmaktadır.

2.2.1. Serbest Akış Seyahat Süresi

New York'ta bazı özel günlerde taksi kullanım sayısı azalmaktadır. Bu günlerde serbest-akış süresi ölçülebilir. Serbest akış seyahat süresi (free-flow travel time), yolların daha boş olduğu zamanlardaki seyahat sürelerini temsil etmektedir.

Bu ölçütlerin dışında bir de FR ve OTR de kullanılmaktadır. Örneğin yapılan 100 seyahatin ortalaması 20 dakika ve 100 seyahatten 30'u 20 dakikanın üzerinde ise başarısızlık oranı:

$$FR = \frac{\text{Ortalamadan uzun süren seyahatlerin sayısı}}{\text{Tüm seyahatlerin sayısı}} = \frac{30}{100} = 0,3' \text{ dir.}$$

zamanında oranı:

$$OTR = \frac{\text{Ortalamadan uzun sürmeyen seyahatlerin sayısı}}{\text{Tüm seyahatlerin sayısı}} = \frac{70}{100} = 0,7' \text{ dir}$$

Çizelge 2.1. Çeşitli ulaşım kurumlarının kullandığı seyahat süresi güvenilirliği ölçütleri [4]

Ulaşım Kurumları	Kullanılan seyahat süresi güvenilirliği ölçütleri
Georgia Bölgesel Ulaştırma Dairesi	BTI ve PI
Florida Ulaştırma Bakanlığı	BTI ve OTR
Washington Eyaleti Ulaştırma Bakanlığı	TT^{95}
Maryland Eyaleti Otoyol İdaresi	TTI ve PI

2.3 . Seyahat Süresi Dağılımı İle İlgili Literatür İncelemesi

Seyahat sürelerinin dağılımının bulunması seyahat sürelerindeki değişimi daha iyi anlaşılmasını ve ona göre analiz yapılmasını sağlar. Bunun için araştırmacılar farklı ülkelerin farklı şehirlerinde otobüs, tren, taksi gibi ulaşım araçları üzerinde çalışmalar yapmaktadır.

Taylor [5], seyahat sürelerini ifade etmek için normal dağılım gibi simetrik dağılımların daha uygun olacağını belirtmektedirler. Örneğin; Taylor Fransa'daki metro ve otobüslerin seyahat süresi dağılımını normal dağılım olarak bulunması buna örnektir.

Mazloui ve arkadaşlarının çalışmasında [6], Avustralya'nın Melbourne şehrindeki otobüs seyahat süreleri analiz edilmiştir ve normal dağılımın seyahat sürelerini modellemek için değerlendirilebileceği belirtilmiştir. Ancak normal dağılımın simetrik yapısı değerlendirildiğinde eksi değerler de içerebileceği ve bunun gerçekte bağdaşmayacağı görülmektedir.

Kieu ve arkadaşları çalışmalarında [7], seyahat sürelerinin pozitif değerli olması gerektiği yönüne dikkat çekerek, Log-Normal dağılım gibi çarpık dağılımları incelemişlerdir ve

seyahat süresi dağılımına en uygun dağılımın Log-Normal dağılım olduğunu iddia etmişlerdir. Bu iddialarını da merkezi limit teoremine dayandırmışlardır.

Chen ve arkadaşlarının çalışmasında [8], Beijing’de dört farklı yol tipiyle yaptıkları çalışmada da Log-Normal dağılımın en iyi dağılım olduğu bulunmuştur.

Diğer bir çarpık dağılım olan Log-Lojistik dağılımı da seyahat süresi dağılımı için uygun bir dağılımdır. Log-Normal dağılımın Normal Dağılımdan dönüştürülmesi gibi Log-Lojistik dağılımı da Lojistik Dağılımdan dönüştürülmektedir. Log-Normal dağılımla karşılaştırıldığında esasen her iki dağılımın da şeklinin birbirine benzediği ancak Log-Lojistik dağılımın analiz edilmesinin daha kolay olduğu belirtilmektedir [9]. Örneğin Şili Santiago’daki metro otobüs ve arabaların seyahat süresi değişkenliğini tahmin etmek için Log-Lojistik dağılımından yararlanılmıştır [10].

Literatürde diğer dağılımlar kadar yaygın olmasa da yeni bir istatistiksel dağılım olan Burr Dağılımı da seyahat süresi dağılımı için önerilmektedir. Taylor ve arkadaşı [11] bu dağılımın esnek bir şekle ve uzun üst kuyrukları tanımlama yeteneğine sahip olduğunu belirtmişlerdir. Bu da seyahat süresi güvenilirliğini çok daha detaylı ölçmeyi sağlayabildiği ifade edilmiştir.

2.4 . Seyahat Süresi Güvenilirliği İle İlgili Literatür İncelemesi

Seyahat süresi güvenilirliği ile seyahat süresi dağılımı birbiriyle ilişkili iki konudur. Seyahat süresi güvenilirliğinin araştırılması için dağılımının belirlenmesi faydalıdır [8]. Trafikte hangi zaman diliminde yola çıkılırsa beklenen zamanda gidilebileceğini bilmek zaman yönetimi açısından önemlidir. Trafiğin seyahat süresi güvenilirliği anlaşılabilirse, zaman yönetimi de daha etkin bir şekilde yapılabilir.

Zamanı iyi kullanmanın önemi herkes tarafından bilindiği için seyahat süresi güvenilirliği uzun süredir araştırmalara konu olmuştur. Seyahat süresi ölçütlerinden standart sapma ve varyasyon katsayısı sık kullanılan ölçütlere aittir. Bunların dışında çarpıklık katsayısı, tampon süre indeksi, planlanan süre indeksi, zamanında oranı gibi ölçütler de kullanılmaktadır.

Dağılım literatüründe de bahsedilen Chen ve arkadaşlarının çalışmasında [8] güvenilirlik ölçütü olarak değişim katsayısı, tampon süre indeksi ve dakiklik oranını kullanmışlardır. Yine dağılım literatüründe bahsedilen bir başka makale olan Tirachini ve arkadaşının

çalışmasında [10] Şili'deki otobüs ve arabaların seyahat süresi güvenilirliği için seyahat süresinin çarpıklık ve basıklıklarını kullanmışlardır.

New York şehrindeki taksi veri setini kullanarak seyahat süresi güvenilirliği, çarpıklık, varyasyon katsayısı ve seyahat süresinin enliği yöntemlerini kullanarak belirlenmiştir [12].

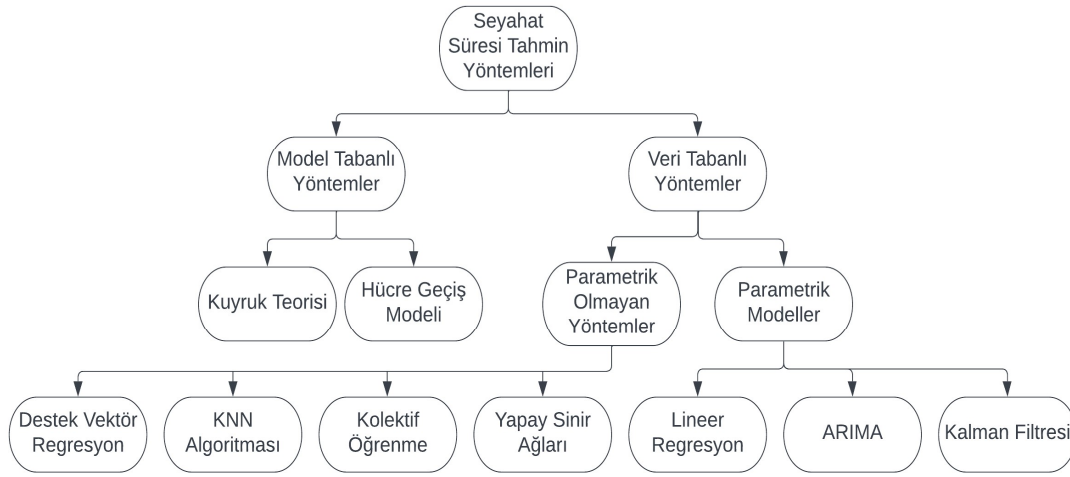
Higatani ve arkadaşlarının çalışmasında [13] ise Japonya'nın Hanshin Otoyolu verileri kullanılarak seyahat süresi güvenilirliği araştırılmıştır. Kullandıkları güvenilirlik ölçütleri ise seyahat süresinin 95'lik yüzdelikleri, standart sapma, varyasyon katsayısı ve tampon süre indeksidir.

Pu çalışmasında [14], farklı seyahat süresi güvenilirlik ölçüt yöntemleri karşılaştırılmıştır. Bu yöntemler arasında seyahat süresinin 95'lik yüzdesi, standart sapma, varyasyon katsayısı, varyasyon yüzdesi, tampon indeksi, planlanan süre indeksi, seyahat süresi indeksi, çarpıklık katsayısı ve kötü yolculukların gecikme zamanlarını gösteren sefalet (misery) indeksidir. Çalışmada, varyasyon katsayısının seyahat süresi güvenilirliğini standart sapmaya göre daha iyi temsil ettiği tespit edilmiştir .Ayrıca ortalama bazlı tampon indeksinin ve ortalama bazlı sefalet indeksinin çarpık dağılımlı seyahat süreleri güvenilirlik ölçütü için daha uygun olduğu belirtilmiştir.

Seyahat süresi güvenilirliğini ölçen genel ölçütlerin dışında, seyahat sürelerinin çarpıklık ve genişliğini kullanarak farklı bir ölçüt de önerilmiştir. Bu ölçüt güvenilir olmayan indeks (unreliability index) olarak adlandırılmıştır. Ancak güvenilir olmayan indeks kullanabilmek için rota uzunluğunun bilinmesi gerekmektedir [15].

3. SEYAHAT SÜRESİ TAHMİN YÖNTEMLERİ

Bir olayın geçmişteki veya şu anki verilerinden yararlanarak oluşumu hakkında iddiada bulunmaya tahmin denir. Seyahat süresi tahmini ulaşım için önemlidir. Seyahat süresi bilgisi tıkanıklık kontrolü, kaza hali gibi durumlar için giriş veya yardımcı veriler olarak kullanılabilir. Seyahat süresini tahmin etmek karmaşıktır. Bunun önemli bir nedeni, hava durumu, yol koşulları nedeniyle farklı dönemlerde değişiklik göstermesidir. Seyahat süresindeki bu dalgalanmaları anlamak ve iyi tahminler yapmak için doğru modeli bulmak gerekir. Şekil 3.1’de seyahat sürelerinin tahmini için literatürde kullanılan bazı yöntemler verilmiştir.



Şekil 3.1. Seyahat süresi tahmin yöntemleri [16]

3.1. Model Tabanlı Yöntemler

Model tabanlı yöntemler, trafik yoğunluğu ve trafik akışı gibi trafik parametrelerini kullanarak seyahat sürelerinin modellerini oluşturur. Birçok aracın toplu ya da bireysel olarak davranışını tanımlayabilir [17].

1. Kuyruk Teorisi (Queuing Theory): Seyahat sürelerinin gecikme sürelerini inceleyerek kuyruk uzunluğunu veya seyahat sürelerini tahmin etmek için kullanılır. Seyahat süresi tahmininin yanında trafik akımının tahmininin yapıldığı çalışmalar da vardır [18].
2. Hücre Geçiş Modeli (Cell Transmission Model): Bu model her biri eşit parçaya bölünmüş hücreler yardımıyla seyahat süresini tahmin eder. Hücreler ağları temsil

eder. Bir hücrede bulunan araçların diğer bir hücreye seyahati, hücrenin yoğunluğu, hücredeki araç sayıları incelenir [19].

3.2. Veri Tabanlı Yöntemler

Veri tabanlı yöntemler kendi içinde parametrik ve parametrik olmayan yöntemler olarak ikiye ayrılır.

3.2.1. Parametrik Yöntemler

Parametrik yöntemler, genellikle önceden bilinen bir dağılıma uyum sağlar. Öncül olarak verinin normal dağıldığı varsayılır.

1. Doğrusal Regresyon: Bağımsız değişken ile bağımlı değişken arasında doğrusal bir ilişki olduğunda bağımsız değişken yardımıyla bağımlı değişkeni tahmin etmek için kullanılan yöntemdir. Bağımlı değişkenin süreklilik varsayımı bulunmaktadır. Seyahat sürelerini tahmin etmede de kullanılmaktadır [20].
2. ARIMA: En bilinen tahmin yöntemlerinden biridir. Seyahat süresi tahmini dışında finansal veriler, üretim verileri gibi başka alanlarda da tahminler yapar. Durağan olmayan zaman serilerinin fark alma işlemiyle durağan bir hale dönüştürür [21].
3. Kalman Filtresi: Bir önceki tahmin değeri, ölçüt değeri ve tahmin ile gözlem arasındaki fark olan kalman kazancı ile şimdiki tahmin değerini hesaplayabilen bir sistemdir [22].

3.2.2. Parametrik Olmayan Yöntemler

Herhangi bir parametre varsayımına gerek olmayan yöntemlerdir.

1. *K- En Yakın Komşu Algoritması (KNN)*: Tahmin edilmesi istenen yeni bir örneğin öznitelik özelliklerine göre ona en yakın k tane noktanın değerlerinin ortalaması yeni örnek için tahmin değeridir [23]. Burada iki önemli unsur vardır. Biri mesafe için hangi uzaklık ölçeğinin kullanılacağı diğeri ise kaç tane k noktasının olacağıdır [24].
2. *Kolektif Öğrenme (Ensemble Learning)*: Veri alt kümelere bölünerek her bir alt küme temel öğrenici olarak görülüp makine öğrenmesi ile eğitilirler. Eğitilen alt kümelerin sonuçları çokluğuna göre tahmin yapan bir yöntemdir [25].
3. *Yapay Sinir Ağları (Artificial Neural Network)*: İnsan beyni örnek alınarak oluşturulan matematiksel sistemlerdir. Yapay sinir hücrelerinin birleşiminden

oluşmaktadır. Girdi, ara ve çıktı katmanlarından oluşur. Eğitim verileri sayesinde öğrenerek tahmin yapmaktadır. İleriki bölümlerde ayrıntılı olarak açıklanacaktır.

4. *Destek Vektör Regresyon (Support Vector Regression)*: Destek vektör makineleri hem sınıflama hem de regresyon için kullanılabilir. Verileri en uygun olacak şekilde n boyutlu düzlemle iki kategoriye ayırır. Doğrusal ayrılabilir ve doğrusal ayrılamama durumu vardır [26].

Verilerin dağılımı normal dağılıma uymadığı zaman tahmin yapılırken parametrik olmayan yöntemlerin kullanılması daha doğrudur. Bu çalışmada seyahat sürelerinin tahmini yapay sinir ağları yöntemleriyle yapılacaktır.

4. YAPAY SİNİR AĞLARI

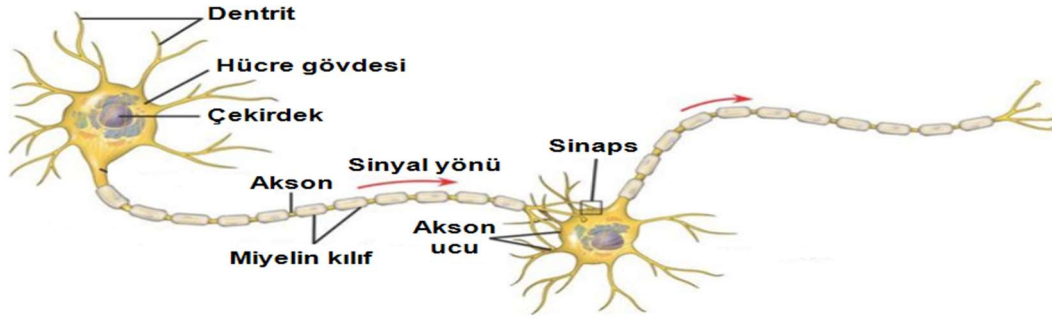
Bu bölümde yapay sinir ağlarının yapısı, elemanları, öğrenme algoritması, avantajları ve dezavantajları açıklanmıştır.

İnsan beyniyle ilgili bugüne kadar birçok çalışma yapılmıştır. Yapılan çalışmalar sonucu beynin sinir sistemimizin önemli bir parçası olduğu bulunmuştur. Beyinde yaklaşık 10 milyar sinir hücresi vardır. Sinir sistemimiz nöron adı verilen sinir hücrelerinden oluşur ve bu hücreler sayesinde bilgi alışverişi yapılmaktadır. Bu da bilim insanları farklı alanlarda da çalışmaya yöneltmiştir. Günümüzde ilerleyen teknolojiyle birlikte biyoloji ve veri biliminin ortak çalışmasıyla yeni alanlar keşfedilmeye başlanmıştır. Bunlardan biri de yapay sinir ağlarıdır. Yapay sinir ağları insan beyninin çalışma şeklini örnek alarak oluşturulmuş sistemlerdir. İnsan beynindeki gibi öğrenme, öğrendiklerini uygulama, yeni bilgiler keşfetme gibi fonksiyonları vardır. Birden çok yapay sinir hücresinin birleşimiyle yapay sinir ağı oluşur. Yapay sinir ağları birçok sorunun çözümünde kullanılmaktadır. Örneğin; sınıflandırma, kümeleme, örüntü tanımlama, sinyal filtreleme, tahmin yapma gibi problemlerin çözülmesine yardımcı olmaktadır. İlk defa 1943 yılında Warren McCulloch ve Walter Pitts tarafından çalışılmaya başlanmıştır. Bu konu ile ilgili çalışmalar devam ederken 1960'lı yıllarda Pappert ve Minsky tarafından yazılan "Algılayıcılar" isimli kitapta yapay sinir ağları ile ilgili olumsuz görüşler bildirilmiştir. Bu kitapta YSA'nın bilimsel açıdan bir değerinin olmadığı ve doğrusal olmayan problemleri çözemediğini kanıtlamaya çalıştılar. Bu kitabın yayınlanmasından sonra yapay sinir ağlarının gelişimi duraklamaya girdi. 1980 yıllarına gelindiğinde ise XOR (exclusive or) probleminin (doğrusal olmayan problemlerin) çözülmesiyle ilerleme yeniden başladı [27]. XOR problemi ilerleyen kısımlarda açıklanacaktır. Özetle YSA biyolojik sinir hücresinin örnek alındığı, yeni bir örnek çıktığında daha önce eğitildiği örneklerdeki bilgilerden faydalanarak yeni çıkarımlar yapabilen sistemlerdir. YSA'nın yapısının örnek alındığı bir sinir hücresi Şekil 4.1'de gösterilmiştir.

4.1. YSA ve Biyolojik Sinir Hücresinin İlişkisi

Genel olarak bir sinir hücresini hücre gövdesinden çıkan kısa ve ağaç dallarına benzeyen yapıdaki dendrit, hücrenin ana işlevlerini belirleyen çekirdek, aksonun etrafını saran miyelin kılıf ve uzun ve uç kısımlara doğru saçaklanan aksondan oluşur. Dendrit diğer sinir hücresinden gelen sinyalleri hücre çekirdeğine gönderir. Hücre çekirdeği dendritten gelenleri toplar ve aksonlara gönderir. Toplanan sinyaller akson tarafından işlenerek

sinapslara gönderilir. Sinapslar bir sinir hücresinin aksonu ile diğer sinir hücresinin dendritine bağlanmasını sağlar. Miyelin kılıf ise sinyallerin daha hızlı biçimde gönderilmesini sağlar (Şekil 4.1.).



Şekil 4.1. Biyolojik sinir hücresinin gösterimi [28]

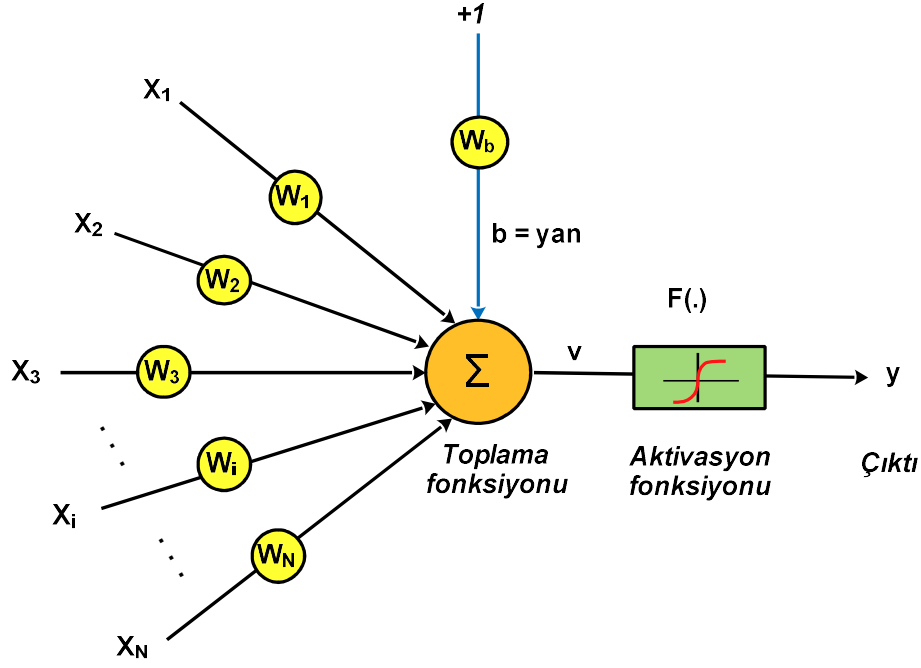
Sinir hücresinde bulunan yapıların yapay sinir hücresindeki karşılıkları Çizelge 4.1’de gösterilmiştir.

Çizelge 4.1. Sinir hücresi ile yapay sinir ağlarının bölümlerinin karşılaştırılması [29]

Biyolojik Sinir Sistemi	Yapay Sinir Ağı
Dendrit	Toplama fonksiyonu
Hücre gövdesi	Aktivasyon fonksiyonu
Sinaps	Ağırlıklar
Akson	Çıktı

4.2. YSA’nın Katmanları

YSA giriş katmanı, gizli katmanlar ve çıktı katmanı olmak üzere üç katmandan oluşmaktadır. Giriş katmanında dışarıdan bilgiler gelir. Bu katmanda herhangi bir işlem gerçekleşmez. Girdi katmanı sayısı girdilerin sayısına eşittir [30]. Girdi katmanındaki her eleman kendinden sonraki ara katmanlara bağlıdır. Gizli (Ara) katmanlarda bu bilgiler işlenir ve çıkış katmanına gönderilir. Ara katmandaki her eleman da çıkış katmanındaki bütün elemanlara bağlıdır. Gizli katmanının sayısını belirlemek için herhangi bir kural yoktur. Çıkış katmanına gelen bilgiler ise giriş katmanına uygun olacak şekilde bilgileri işleyerek dışarıya gönderir. Her çıkış katmanının bir çıktısı vardır. Şekil 4.2’de yapay sinir ağının genel yapısı verilmiştir.



Şekil 4.2. Yapay sinir ağı yapısı

Burada X_1, X_2, \dots, X_N girdileri göstermektedir. Girdiler dışarıdan gelebileceği gibi başka bir sinir hücresinden de gelebilirler. W_1, W_2, \dots, W_N ağırlıkları göstermektedir. Ağırlıklar, girdilerin yapay nöron üzerindeki etkisini belirleyen katsayılardır. b , yan değeridir. Sıfırdan farklı bir yan değerinin seçimi, tüm girdilerin sıfır olduğu durumda çıktının sıfır değerini almasını önler. Toplama fonksiyonu, net girdi değerini verir:

$$v = net = \sum_{i=1}^N X_i W_i + b \quad (4.1)$$

Toplama fonksiyonu

Ağırlıklarla girdilerin çarpılıp toplanmasıyla bulunduğu gibi farklı yollarla da elde edilebilir. Örneğin; ağırlıklarla girdilerin çarpılıp tekrar birbirleriyle çarpılması gibi, ağırlıkla girdi çarpıldıktan sonra değerlerin en büyüğünün ya da en küçüğünün alınması gibi diğer toplama fonksiyonları da bulunmaktadır [31]. Bunlar probleme hangi toplama fonksiyonunun daha uygun olabileceğine göre belirlenmektedirler.

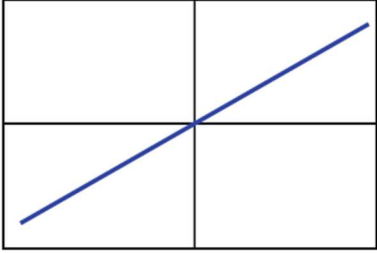
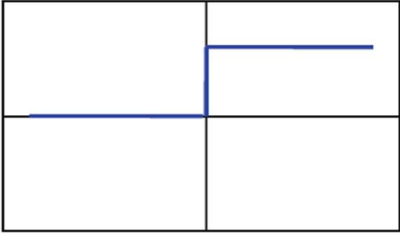
Aktivasyon fonksiyonu

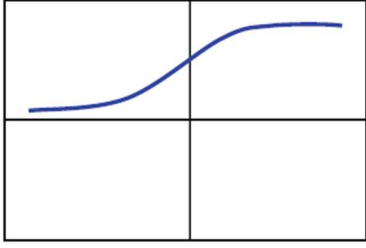
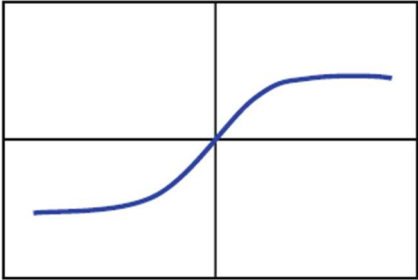
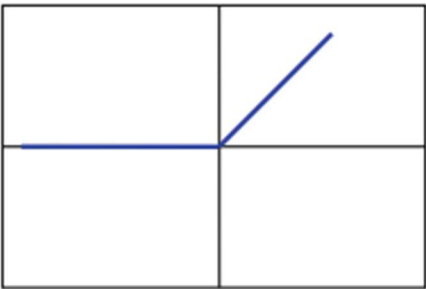
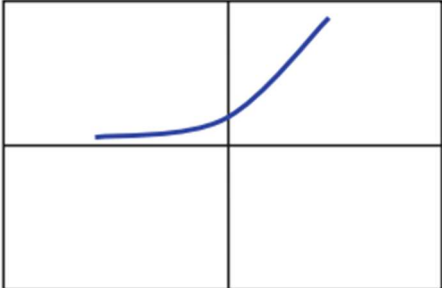
Aktivasyon fonksiyonu $F(.)$ ise net girdiyi işleyerek girdiye uygun çıktılar üretir. Birden fazla aktivasyon fonksiyonu türü vardır. Aktivasyon fonksiyonunun seçilmesinde dikkat edilmesi gereken nokta türevinin kolay hesaplanabiliyor olmasıdır çünkü geri beslemeli

ağlarda türevinin kolay alınırsa işlem yavaşlaması azalır. Geri beslemeli ağlarda türevinin alınması da hatayı azaltmak içindir. Aktivasyon fonksiyonu türlerinin grafiği ve formülleri Çizelge 4.2’de gösterilmiştir.

- Doğrusal aktivasyon fonksiyonu: Grafiği doğrusaldır. Bu fonksiyonun türevi sabit bir sayıdır. Bundan dolayı pek işlevsel değildir.
- Adım aktivasyon fonksiyonu: Çift değer alabilen bir fonksiyondur. Eşik değerine göre alacağı değer değişir.
- Sigmoid aktivasyon fonksiyonu: En çok kullanılan aktivasyon fonksiyonudur [32]. Doğrusal bir fonksiyon değildir.
- Tanjant hiperbolik aktivasyon fonksiyonu: Sigmoid fonksiyonuyla benzer yapıdadır ama tanım aralığı farklıdır. Bu fonksiyon daha çok sınıflandırma yapmak için kullanılır. Sigmoid ve tanjant hiperbolik fonksiyonları ileri beslemeli sınıf ağlarında kullanılır [33].

Çizelge 4.2. Aktivasyon fonksiyonları [34]

Aktivasyon fonksiyonu	Grafik	Eşitlik	Aralık
Doğrusal aktivasyon fonksiyonu		$F(x)=x$	$(-\infty, \infty)$
Adım aktivasyon fonksiyonu		$F(x)=0, x<0$ $F(x)=1, x\geq 0$	$(0,1)$

Sigmoid aktivasyon fonksiyonu		$F(x) = \frac{1}{1 + e^{-x}}$	(0,1)
Tanjant hiperbolik aktivasyon fonksiyonu		$F(x) = \tanh(x)$	(-1,1)
ReLU		$F(x) = 0, x < 0$ $F(x) = x, x \geq 0$	$[0, \infty)$
SoftPlus		$F(x) = \ln(1 + e^x)$	$(0, \infty)$

4.3. YSA ve Geleneksel Algoritmalar

Yapay sinir ağlarını geleneksel algoritmalarla karşılaştırıldığında bazı farklılıkların olduğu görülür. YSA'nın genel algoritmalarından ayrılan bir özelliği birden çok işlemciyle çalışmasıdır. Geleneksel algoritmalarda tek işlemci vardır. Ayrıca YSA'da geleneksel algoritmalarındaki gibi bilgiler net ve kesin değildir, deneyimlerden faydalanılır. YSA'nın bir diğer özelliği ise hata toleransına sahip olmasıdır. Geleneksel algoritmaların hata toleransı

yoktur. Bunların yanı sıra geleneksel algoritmaların YSA'ya göre daha hızlı olduğu söylenebilir.

4.4. YSA'ların Sınıflandırılması

YSA öğrenme algoritmaları ve yapısına göre sınıflandırılabilir.

4.4.1. YSA Öğrenme Algoritmaları:

Öğrenme, yapay sinir ağlarının en önemli özelliğidir. Öğrenmenin iyi bir şekilde olması içinse ağırlıkların en uygun halde olması gerekir. Dışarıdan bilgi geldiğinde bunu en uygun ağırlıkla çıktı olarak sunduğunda ağ için eğitimin iyi olduğu söylenebilir. Yapay sinir ağlarında birkaç öğrenme yöntemi vardır. Bunlar:

- *Danışmanlı Öğrenme:* Bu öğrenme türünde girdi ile istenilen çıktı bilgisi aynı anda verilir. Ağın ürettiği çıktı bilgisiyle başta verilen çıktı bilgisi karşılaştırılarak hata oranı ölçülür. Hata oranına göre de ağırlıklar yeniden düzenlenir. Bu işlem hata oranı istenilen seviyeye gelene kadar devam eder. Başlangıçta ağa verilen veri setine ise eğitim veri seti denir. Geri yayılım ağı, Perceptron (algılayıcı), delta öğrenme ve eğimli iniş danışmanlı öğrenmedir [35].
- *Danışmansız Öğrenme:* Danışmanlı öğrenmedeki gibi hem giriş hem de çıktı bilgisi verilmez. Sadece giriş bilgisi verilir ve ağın bunu kendi öğrenmesi istenir. Hebbian öğrenme, rekabete dayalı öğrenme ve kendi kendine öğrenme danışmansız öğrenmedir [35].
- *Destekleyici Öğrenme:* Ağın hem öğrenerek hem de sonuç çıkararak işlem yaptığı öğrenme türüdür [32]. Yaptığı işlemin sonucuna göre başarılı olduğuna ya da başarısız olduğuna karar verilir.

4.5. Öğrenme Zamanına Göre Sinir Ağları

- *Dinamik Sinir Ağları:* Öğrenmenin devam ettiği, zaman gecikmesine ve geri besleme gecikmesine sahip ağlardır.
- *Statik Sinir Ağları:* Öğrenmenin ağın aktif olmadığı zaman devam etmediği, zaman faktörünün ve geri besleme gecikmesinin olmadığı ağlardır.

4.6. Öğrenme Kuralları

Ağırlık değerleri belli kurallara göre değiştirilmektedir. Yapay sinir ağlarında kullanılan bazı kurallar şunlardır.

- *Hebb kuralı*: En eski ve en basit öğrenme kuralıdır. 1949 yılında Donald Hebb tarafından geliştirilmiştir. Bu kurala göre eğer bir hücre aktif ise, komşu hücrelerini de aktive eder. Aktif değilse komşu hücrelerini de devre dışı bırakır mantığını içermektedir. Hücreler birbiriyle aynı fazda (her ikisi de pozitif veya negatif olduğunda) ise birbirini bağlayan ağırlık değerinin artmasını ters fazda (biri negatif biri pozitif olduğunda) ise azaltılması gerektiğini söyler.
- *Perceptron kuralı*: Tek katmanlı algılayıcılarda kullanılan bir kuraldır. 1957 yılında Frank Rosenblatt tarafından geliştirilmiştir. Bu kurala göre öğrenme katsayısı adı verilen bir sabit ile ağırlıklar çarpılıp ağırlıklara ekleme veya çıkartma işlemi yapılır.
- *Hopfield kuralı*: Hopfield kuralı Hebb kuralını baz alarak oluşturulmuştur. Bu öğrenmede öğrenme katsayısına bağlı olarak ağırlıklarda değişiklik yapılır. Hücrelerin ikisi de aktifse ağırlık öğrenme katsayısı kadar artırılır. Hücreler pasifse de öğrenme kuralı kadar azaltılır.
- *Delta kuralı*: Hebb kuralının başka bir çeşidi olan delta kuralında, ağa girişte verilen beklenen çıktılar ile ağırlık çıktısının farkının en aza inmesi amaçlanmıştır. Beklenen çıktı ile ağırlık çıktısının farkına hata denilirse bu hatayı geriye doğru ağırlıklara dağıtarak doğru sonuca yaklaşılar. Delta kuralı en küçük kareler yöntemine dayanmaktadır. Delta kuralı en çok kullanılan kurallardan biridir. Adaline/madeline ağları ve çok katmanlı sinir ağlarında kullanılmaktadır.
- *Kohonen kuralı*: Teuvo Kohonen tarafından geliştirilmiştir [36]. Kuralda hücreler arasında bir rekabet söz konusudur. En büyük çıktıya sahip hücre kazanan olmaktadır ve yalnızca kazanan hücre ile komşu hücrelerinin ağırlıkları değiştirilir.

Öğrenme algoritmaları ve kurallarının yanı sıra yapay sinir ağlarının öğrenmesi için öğrenme katsayısı ve momentum katsayısı da önemlidir. Bu değerlerin büyüklüğüne göre de öğrenme etkilenmektedir.

Öğrenme katsayısı

Ağırlıkların ne kadar değiştiğini gösteren katsayıdır. Katsayı eğer çok küçük olursa ağın öğrenmesi zaman alır. Büyük olursa da optimal çözüm yerine yerel çözümlere takılabilir. Öğrenme katsayısı 0 ile 1 arasında sabit bir sayıdır. Ağı tasarlayan kişi bu aralıkta kendi belirlediği bir sayıyı alabilir ama Ercan Öztemel “Yapay Sinir Ağları”

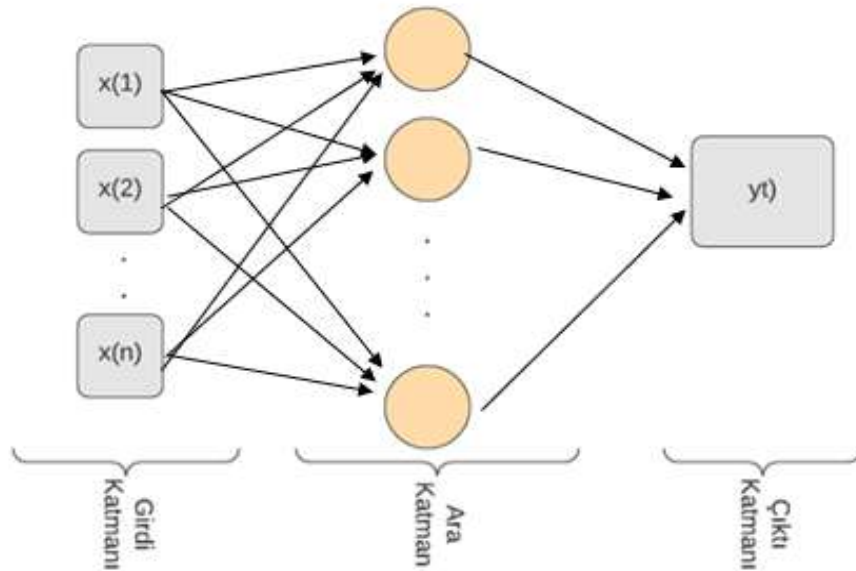
adlı kitabında öğrenme katsayısı için 0.2 ile 0.4 arasındaki değerin çalışmalar için daha uygun olacağını söylemektedir [37]. Yine de deneme yanılma yoluyla en doğru öğrenme katsayısı bulunabilir.

Momentum katsayısı

Ağın öğrenmesindeki bir diğer önemli katsayı, momentum katsayısıdır. Bir önceki yinelemedeki ağırlık değişiminin bir kısmını yeni değişim miktarına ekler [37]. Her yapay sinir ağında bulunmayabilir. Problemin yerel sonuçları optimal sonuçlar gibi görmemesi için momentum katsayısı seçimi önemlidir. Momentum katsayısı da öğrenme katsayısı gibi 0 ile 1 arasında bir değer alır. En doğru momentum katsayısı yine deneme yanılma yoluyla bulunabilir. Yapılan çalışmalarda 0.6 ile 0.8 aralığının tercih edilmesinin daha uygun olduğu söylenmektedir [37].

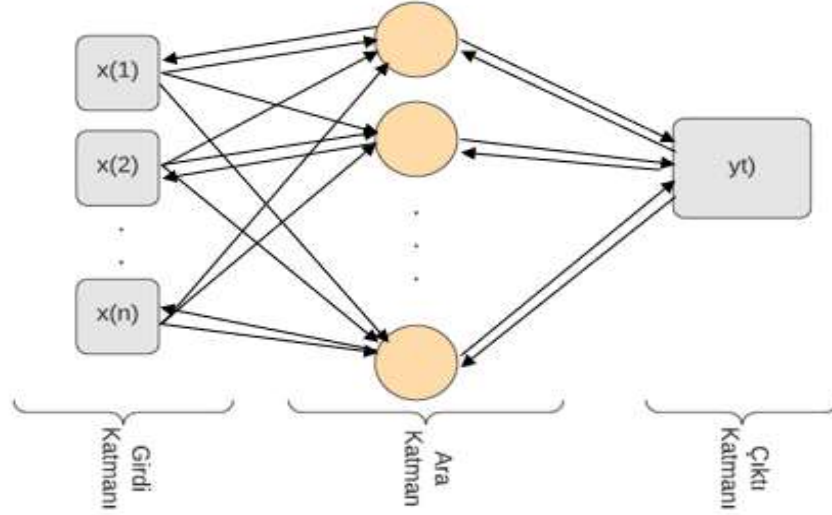
4.7. Yapısına Göre Yapay Sinir Ağları

- İleri Beslemeli Yapay Sinir Ağları: Dışarıdan gelen bilginin girdi tabakasından başlayarak sırasıyla ara ve çıktı tabakalarından tek yönlü olarak geçtiği ağ yapısıdır. İleri beslemeli yapay sinir ağları dinamik ağlar değildir. Doğrusal olmayan statik bir ağ olduğu söylenebilir [38]. İleri beslemeli yapay sinir ağının yapısı Şekil 4.3'te verilmiştir.



Şekil 4.3. İleri beslemeli yapay sinir ağı

- Geri Beslemeli Yapay Sinir Ağları: İleri beslemeli yapay sinir ağları gibi statik bir yapısı yoktur. Doğrusal olmayan dinamik bir yapıya sahiptir [30]. Bu tip yapay sinir ağında bilgi sadece tek yönlü değildir. Çıktılar da girdi olarak kullanılmaktadır. Bilgi akışı kendinden bir önceki katmana ya da kendi katmanları arasında da olabilmektedir. Geri beslemeli yapay sinir ağının yapısı Şekil 4.4'te verilmiştir.



Şekil 4.4. Geri beslemeli yapay sinir ağları

4.8. Tek Katmanlı Algılayıcılar

Tek katmanlı algılayıcılar doğrusal problemleri çözmeye başarılıdır. Doğrusal olmayan problemlerin çözümünde kullanılamazlar. Doğrusal olmaktan kasıt bir doğru veya düzlemle sınıflara ayrılabilir olmasıdır. Bu ağlar sadece girdi ve çıktı katmanından oluşmaktadır. Ara katmanlar bulunmamaktadır. Giriş katmanındaki değerler ile ağırlıklar çarpılıp eşik değerinin eklenmesiyle çıktı değeri oluşturulur. Eşik değerinin eklenmesinin sebebi ise çıktının sıfır olmasını önlemektir. Eşik değeri burada 1'dir. Tek katmanlı algılayıcılar 1 ve 0 (veya 1 ve -1) olarak çıktı değerleri almaktadır. Bu da sınıflamanın olduğunu gösterir. Amaç en iyi ayıran doğruyu veya düzlemi bulmaktır. Bu algılayıcıların diğer bir problemi ise ağın tekrar etmesiyle ağırlıklar değiştirildiğinde öğrendiklerini unutma ihtimalinin olmasıdır [37]. Bu algılayıcılarda önemli iki model bulunmaktadır. Bunlar, basit algılayıcı model ve adaline/madaline modelleridir.

4.9. Basit Algılayıcı Model

Bir yapay sinir hücresinden oluşur. 1958 yılında Rosenblat tarafından geliştirilmiştir [37]. Ağırlık değerleri ile girdi değerlerinin çarpılmasıyla net değeri aşağıdaki formülle bulunur.

$$Net = \sum_{t=1}^n \text{Ağırlık değeri} * \text{girdi değeri} \quad (4.2)$$

Basit algılayıcı ağların çıktıları (Ç) mantıksal çıktılardır, 0 ve 1 değerlerinden oluşur. Eşik değeri de çıktıların hesaplanmasına yardımcı olmaktadır.

$$\zeta = \begin{cases} 1, & Net > \text{Eşik değeri} \\ 0, & Net \leq \text{Eşik değeri} \end{cases} \quad (4.3)$$

Girdilerle birlikte beklenen çıktılar da ağa gönderilir. Ağın çıktı değeri ile beklenen çıktı değerleri arasında fark varsa ağırlıklar ve eşik değeri değiştirilerek yani eğitilerek tekrar hesaplanır. Aralarında fark yoksa ağırlıklar değiştirilmez. Ağın beklenen çıktısı çıktı değerinden küçükse ağırlıklar azaltılır, eğer büyükse de ağırlıklar artırılır. Ağırlıklar değiştirilirken öğrenme katsayısından yararlanılır. Dengeye gelinceye kadar da ağırlıkların değiştirilmesi devam eder. Örnekler doğru sınıflandırıldığında ağın eğitimi tamamlanmış olur.

4.10. Adaline/Madaline Modeli

Adaline (Adaptive Linear Neuron) ve Madaline (Multiple Adaptive Linear Neuron) 1959 yılında Bernard Widrow ve Marcian Hoff tarafından geliştirilmiştir. Adaline ağı, en küçük kareler yöntemini kullanarak öğrenir. Bu öğrenme kuralı aynı zamanda Delta kuralı olarak da bilinmektedir [37]. Ağda girdiler ve ağırlıklar çarpılıp eşik değeri eklendikten sonra net girdi hesaplanır:

$$Net = \sum_{t=1}^n \text{Ağırlık değeri} * \text{girdi değeri} + \text{eşik değeri} \quad (4.4)$$

Bu net girdi değeri ile çıktı değeri karşılaştırılır. Çıktı değerleri için 1 ve -1 denilirse eğer net girdi 0'a eşit veya sıfırdan büyükse "1", sıfırdan küçükse -1 ağı çıktısı olarak kabul edilir.

$$\zeta = \begin{cases} 1, & Net \geq 0 \\ -1, & Net < 0 \end{cases} \quad (4.5)$$

Beklenen çıktıdan ağın çıktısının çıkarılmasıyla hata hesaplanır. Bu hata değerine göre ağın ağırlıkları yeniden düzenlenir. Amaç hatayı minimum yapmaktır.

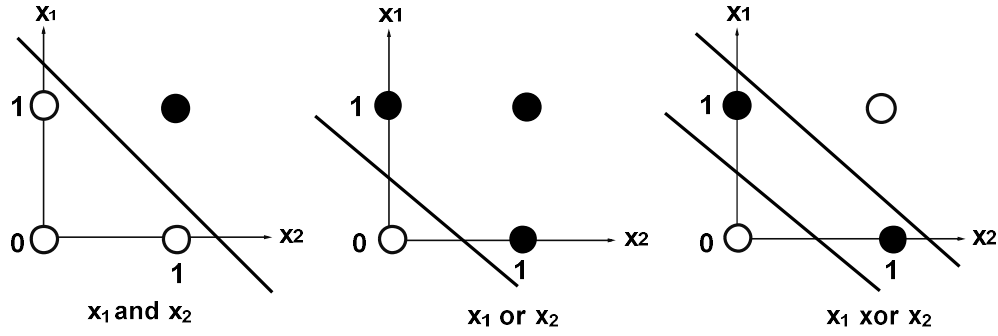
Madaline ağı ise birden çok adaline ağının birleşmesiyle oluşmaktadır. Madaline gerçek dünya problemlerine uygulanan ilk yapay sinir ağıdır. Telefon hatlarındaki yankıları

kaldırmak için adaptif bir filtre olarak kullanılmıştır [36]. Öğrenmesi adaline ağının öğrenmesiyle aynıdır.

Tek katmanlı ağların birbirinden farkı öğrenme şekilleridir. Basit algılayıcı ağlarda öğrenme katsayısı ile adaline/madeline ağlarda da hatayı en aza indirerek öğrenme sağlanmaktadır.

4.11. Çok Katmanlı Algılayıcılar

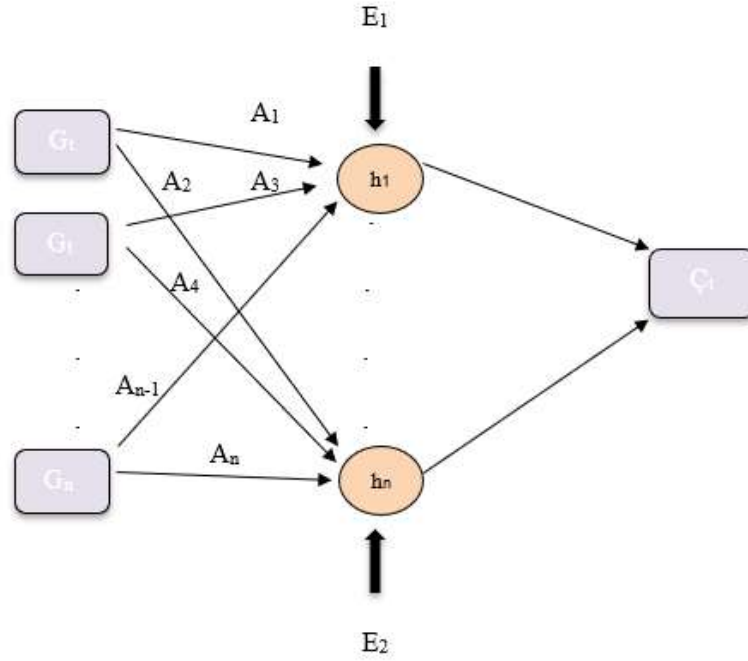
Karmaşık problemlerin çözümüne yardımcı olması amacıyla Rumelhart tarafından bulunmuştur. XOR (exclusive or) problemi gibi doğrusal olmayan problemler için çok katmanlı algılayıcılara ihtiyaç duyulmuştur. XOR problemi bir doğru ya da düzlemle ayrılabilen bir problemdir. Ayrılabilmesi için giriş ve çıkış katmanının yanında bir de gizli (ara) katmana ihtiyaç duymaktadır.(Şekil 4.5)



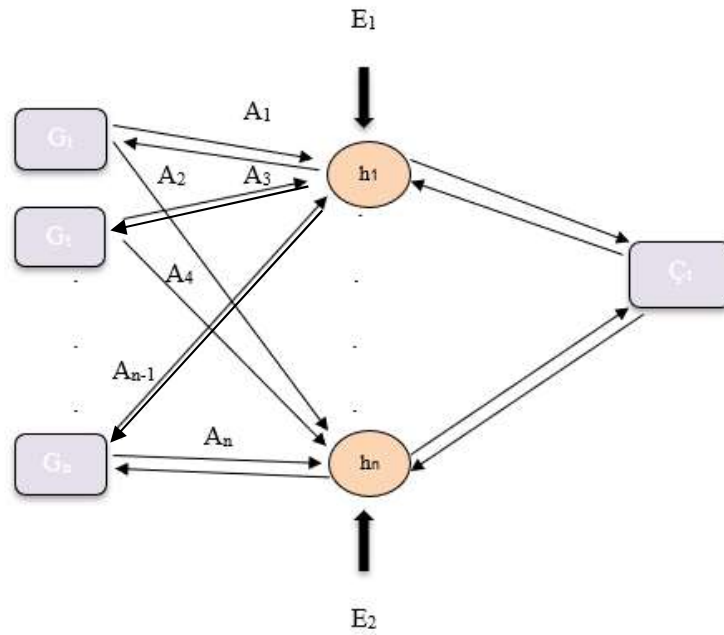
Şekil 4.5. AND, OR ve XOR problemlerinin eksen üzerinde gösterimleri [39]

XOR problemini çözmek için iki eşik değeri verildiğinde eşik değerlerinin üstünde kalanlar “0”, arasında kalanlar ise “1” olarak sınıflandırılır. Ara katmanların sınıflandırmaya yardımcı olduğu söylenebilir. Ara katmanların problemin karmaşıklığına göre de sayısı artmaktadır. Yukarıda bahsedildiği üzere bilgi ilk olarak girdi katmanına gelmektedir. Burada herhangi bir işlem gerçekleşmez. Girdi katmanından sonra ara katmana gelen bilgi burada işlenerek kendinden sonraki katmana gönderilir. Çıktı katmanında da ara katmandan gelen bilgi işlenir ve çıktı olarak sunulur. Çok katmanlı algılayıcılarda danışmanlı öğrenme yöntemi kullanılmaktadır. Yani girdilerle birlikte çıktılarda verilip ağın öğrenmesi istenmektedir. Öğrenme kuralı olarak da adaline/madeline ağının kullandığı en küçük kareler yöntemine dayanan delta kuralının geliştirilmiş hali kullanılmaktadır. Geliştirilmiş delta kuralına göre önce ileriye doğru hesaplama (feed forward) yaparak ağın çıktısı hesaplanır. Şekil 4.6’da ileri beslemeli hesaplama yapısı gösterilmiştir. Daha sonra geriye doğru hesaplama (back propagation) ile ağın ağırlıklarının değiştirilmesi ve

hatanın geriye doğru yayılması yapılır. Şekil 4.7'de geriye doğru hesaplama yapısı gösterilmiştir.



Şekil 4.6. İleri doğru hesaplama



Şekil 4.7. Geriye doğru hesaplama

Kısaca çok katmanlı algılayıcılar için ilk olarak örneklerin toplanması gerekmektedir. Örneklere uygun ağı kaç tane girdisi olmalı, kaç tane gizli katmanı ve kaç tane çıktı katmanı içermeli gibi ağın topolojik yapısını oluşturan etmenler belirlenmeli ancak bu etmenlerin belirlenmesi için belirli bir yöntem bulunmamaktadır. Deneme yanılma yoluyla ya da daha önceki çalışmalar temel alınarak oluşturulabilir. Ağı yapısı oluşturulduktan sonra öğrenme katsayısı, momentum katsayısı, toplama ve aktivasyon fonksiyonlarından hangisinin kullanılacağı, ağırlıklar ve eşik değerleri belirlenmelidir. Yine bunları belirlemek için belirli bir yöntem bulunmamaktadır. Bu elemanların belirlenmesinden sonra ağı eğitilmesi işlemine geçilebilir. İlk olarak ileri doğru hesaplama ile daha sonra da geriye doğru hesaplama ile çok katmanlı algılayıcıların ilk döngüsü tamamlanmış olur. Bu döngüler en uygun çözümü buluncaya kadar devam eder.

4.12. ÇKA'da Dikkat Edilmesi Gereken Noktalar

- Ağa örnekleri sunarken, örneklerin problem uzayını doğru temsil edebilecek örnekler olması gerekmektedir. Yani ağdan istenilenin ağı daha önceden görmüş olması gerekir. Farklı bir örnek verildiğinde istenilen çözümü ağ sunamaz.
- Ağa sunulan girdilerin sayısal girdiler olması gerekmektedir. Eğer girdiler sayısal değilse uygun bir şekilde sayısal hale çevrilmeleri gerekir.
- Ağı ağırlıklarının seçimi için belirli bir yöntem yoktur. Yapılan çalışmalar ağırlıkların -1.0 ile 0,1 arasında olmasının daha uygun olduğunu söylemektedir [37].
- Ağa sunulan örnekler farklı ölçeklerde toplanmışsalar bu örneklerin aynı ölçüğe gelmeleri (belirli bir aralıkta olmaları ya da normalize edilmiş olmaları gibi) gerekir. Aynı ölçekte olmaları, hatalı girilen çok büyük ve çok küçük değer varsa bunların ağa etkisini de azaltmış olur.

4.13. YSA'nın Avantajları

- Kendi başlarına öğrenebilmektedir.
- Eksik bilgi ile çalışabilir. Eksik bilgi olsa dahi doğru sonuçlar üretebilir.
- Hata toleransları vardır. İşlemler yapılırken oluşabilecek bir hata yapay sinir ağlarının çalışmasını bozamaz.
- Doğrusal olmayan problemleri çözmede başarılıdır.
- Aynı anda tüm parametreleri çalışabilir. Bu da paralellik özelliği olduğunu gösterir.
- Öğrenmesini tamamladıktan sonra hiç görmediği örnekler üzerinde doğru tahminler bulunabilir.

- Bilgiler ađın kendisinde saklanır. Bunun için bir veri tabanına ihtiyaç duymaz. Bilgileri ađda saklaması veri kaybı olduđunda çalışmasına devam etmesini açıklamaktadır.
- Yapay sinir ađları verilen örneklerle genelleme yaparak girdi ve beklenen çıktıya uygun olarak bir çözüm üretir.

4.14. YSA'nın Dezavantajları

- Sadece sayısal bilgilerle çalışabilmektedirler. Eğer veri setinde sayısal olmayan bilgiler varsa bunları problemin tanımına göre nümerik veriye çevirmek gerekmektedir.
- Ađın nasıl oluşturulması gerektiđine dair bir kural yoktur.
- Ađın ađırlığı, öğrenme katsayısı, moment katsayısı gibi parametrelerinin nasıl seçileceđine dair bir kural yoktur.
- Ađın girdilerinin aynı ölçekte olması gerekmektedir. Belirli bir ölçekte girilmezse çok büyük ve çok küçük deđerler ađı etkileyebilir. Aynı şekilde çıktıların da belirli bir ölçekte olması gerekmektedir.
- Ađın eğitilmesi için belirli bir sayı yoktur. Deneme yanılma yoluyla yeterli sayıya ulaşılmaktadır.
- Ađa verilen girdilerin nasıl çıktılara dönüştürüldüğü tam olarak açıklanamamaktadır. Bu da yapay sinir ađlarının kara kutuya benzetilmesine yol açmıştır.
- Bazen bulunan çözüm optimal çözüm olmayabilir.

4.17. Doğrusal Olmayan Oto regresif Ađ (Nonlinear Autoregressive Model-NAR)

NAR sinir ađları, zaman gecikmesine sahip tek bir deđişken üzerinden sorunu çözen kapalı döngülü bir ađdır. Tahmin edilmesi istenen deđişkenin geçmiş deđerleri kullanılır. Eşitliđi aşıđıda verilmiştir.

$$y_t = f(y_{(t-1)}, y_{(t-2)}, y_{(t-3)}, \dots, y_{(t-d)}) \quad (4.20)$$

Bu eşitlikte t zaman, d zamansal gecikme, $y_{(t-1)}, y_{(t-2)}, y_{(t-3)}, \dots, y_{(t-d)}$ ađ çıktılarını temsil etmektedir.

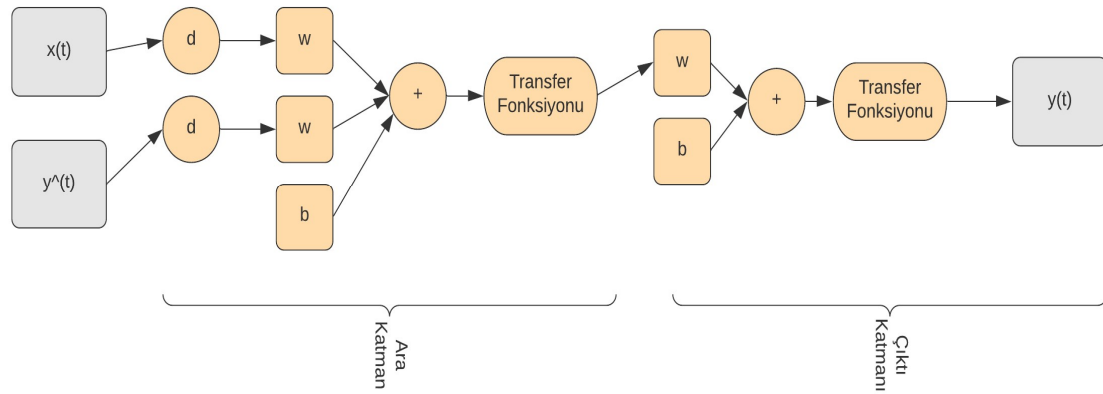
4.18. Doğrusal Olmayan Dışsal Girdili Otoregresif Ağ (Nonlinear Autoregressive Model with External Input-NARX)

NARX sinir ağları, dışarıdan girdili, geri beslemeli, zaman gecikmesine sahip döngü mekanizmalı dinamik sinir ağlarıdır. “X” dışsal girdileri temsil etmektedir. NARX doğrusal ve durağan olmayan serileri modelleyebilmektedir [40]. Yani herhangi bir varsayıma ihtiyaç duymamaktadır [41]. NARX tekrarlayan sinir ağları yapısının bir çeşididir. NARX in eşitliği aşağıdaki gibidir.

$$y_t = f(y_{(t-1)}, y_{(t-2)}, y_{(t-3)}, \dots, y_{(t-d)}, x_{(t-1)}, x_{(t-2)}, x_{(t-3)}, \dots, x_{(t-d)}) \quad (4.21)$$

Burada yine t zaman, d zamansal gecikme, $y_{(t-1)}, y_{(t-2)}, y_{(t-3)}, \dots, y_{(t-d)}$ ağ çıktılarını temsil etmektedir. NAR ağından farklı olarak $x_{(t-1)}, x_{(t-2)}, x_{(t-3)}, \dots, x_{(t-d)}$ ise ağ girdilerini temsil eder. Diğer geri beslemeli ağların aksine, NARX ağları bütün gizli katmanlardan değil sadece çıktı katmanındaki nöronlardan geribildirim almaktadır. Modelde çıkışa zaman gecikmesi uygulanarak tekrar girdi olarak kullanılmaktadır. Yani bağımlı çıktı değişken $y(t)$ 'nin tahmini için, çıktı değişkeninin önceki değerleri ve bağımsız harici (dışsa) girdi değişkenler birlikte analize girer.

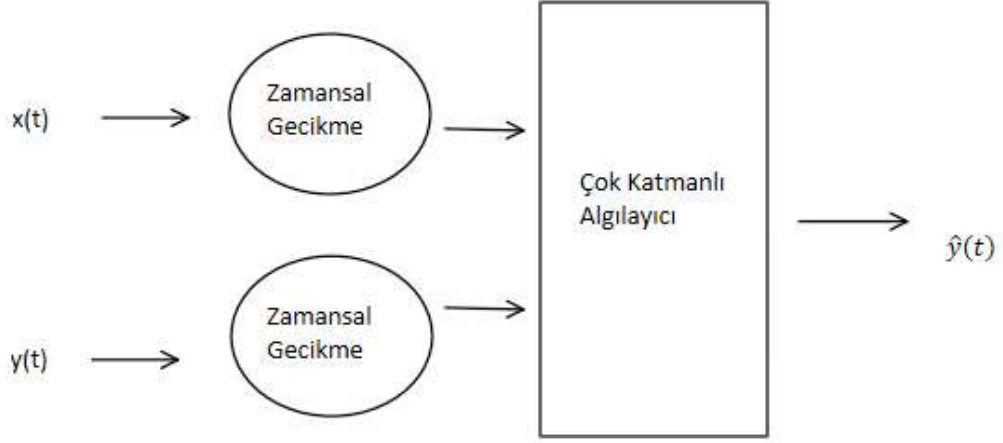
NARX ağının yapısı Şekil 4.8’de verilmiştir.



Şekil 4.8. NARX ağının yapısı

NAR ve NARX sinir ağlarının iki farklı mimarisi bulunmaktadır. Biri seri paralel mimari (açık döngü olarak da adlandırılır) diğeri de paralel mimari (kapalı döngü olarak da adlandırılır) dir.

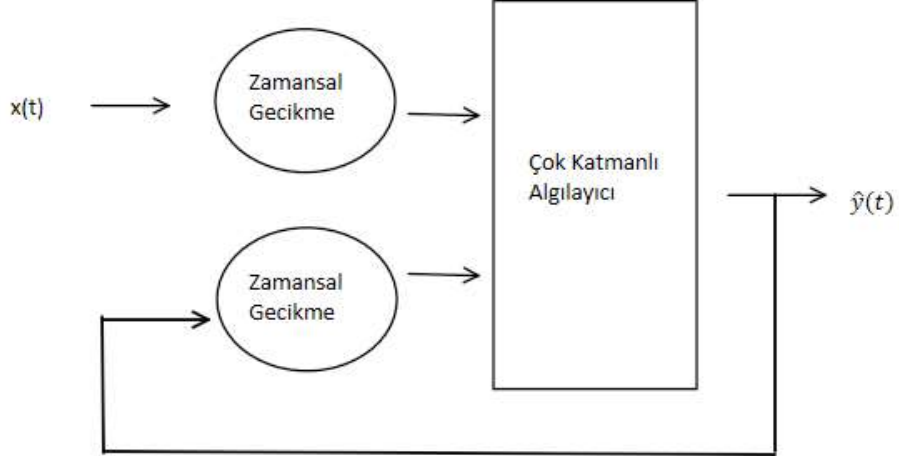
Seri paralel mimaride tahmin edilmesi istenen değişkeni etkileyen dışsal değişkenlerle birlikte tahmin edilmesi istenen değişkenin geçmiş değerleri Şekil 4.9’da görüldüğü gibi girdi olarak ağa gönderilir.



Şekil 4.9. Seri paralel mimari

Paralel mimaride ise tahmin edilecek değişkenin geçmiş değerleri girdi olarak kullanılmaz. Onun yerine tahmin edilen çıktı değeri tekrar girdi olarak ağa gönderilir. Şekil 4.10'da mimari yapısı gösterilmiştir.

Açık döngüyle eğitim gerçekleştirildikten sonra döngü kapatılarak bir adım sonrası tahmin edilir.



Şekil 4.10. Paralel mimari

4.19. NARX Ağ Tasarımı

Sinir ağlarında ara katman sayısı ve nöron sayısı, aktivasyon fonksiyonu seçimi, verilerin normalizasyonu, eğitim algoritmasının seçimi ağın performansında önemli bir etkiye sahiptir.

4.19.1. Ara Katman ve Nöron Sayıları

Katman ve nöron sayısının belirlenmesinde belirli bir yöntem bulunmamaktadır. Deneme yanılma yolu ile ideal katman ve nöron sayısı belirlenir. Genel olarak tek bir ara katmanın çoğu problemde yeterli olduğu ve daha az nöron sayısına sahip ağların daha iyi genelleme yeteneğine ve daha az uyum sorununa sahip oldukları için tercih edilir [42].

4.19.2. Aktivasyon Fonksiyonunun Seçimi

Bir ağın giriş ve çıkışları arasındaki ilişkiyi belirleyen aktivasyon fonksiyonu, ağın performansını etkileyen faktörlerden biridir. Farklı katmanlarda aynı aktivasyon fonksiyonları kullanılabileceği gibi farklı aktivasyon fonksiyonları da kullanılabilir. Çok katmanlı ağlarda fonksiyonun kolay türevinin alınabiliyor olması tercih edilir.

4.19.3. Verilerin Normalizasyonu

Veriler ham haldeyken çok büyük ya da çok küçük değerler içerebilir. Analiz yapılırken ağı yanlış yönlendirmemek için verileri aynı ölçeğe getirmek daha doğru bir analiz için gereklidir. Normalizasyon veri setindeki değerlerin dağılımını düzenli bir hale getirmektedir [43]. Ayrıca lineer olmayan aktivasyon fonksiyonları $[0,1]$ aralığında değerler üretmektedir. Bunun için ağ eğitime başlamadan önce bu aralığa getirilmelidir [41]. Verilerin belirli bir ölçekte olması, örneğin $[0,1]$ aralığında, ağın eğitiminin kolaylaşmasına ve hesaplama problemlerinin azaltılmasına katkı sağlamaktadır [42].

İstatistikte birden fazla normalizasyon yöntemi vardır. Bunlardan bazıları z-skor normalizasyonu, min-max normalizasyonu, medyan normalizasyonu, sigmoid normalizasyonudur [44]. Çizelge 4.3 de bu normalizasyon yöntemlerinin formülleri ve açıklamaları verilmiştir.

Çizelge 4.3. Normalizasyon yöntemleri

Normalizasyon Adı	Formülü	Formül Açıklaması
Z-skor Normalizasyon	$x' = \frac{x_i - \mu_i}{\sigma_i}$	x' = Normalleşmiş veri x_i = Girdi değeri μ_i =Girdilerin ortalama değeri σ_i = Girdilerin standart sapması
Min-Mak Normalizasyonu	$x' = \frac{x_i - x_{min}}{x_{mak} - x_{min}}$	x' = Normalleşmiş veri x_i = Girdi değeri x_{min} = Girdilerin en küçük değeri x_{mak} = Girdilerin en büyük değeri
Medyan Normalizasyonu	$x' = \frac{x_i}{Medyan(a_i)}$	x' = Normalleşmiş veri x_i = Girdi değeri a_i = Girdilerin medyanı
Sigmoid Normalizasyonu	$x' = \frac{e^{xi} - e^{-xi}}{e^{xi} + e^{-x}}$	x' = Normalleşmiş veri x_i = Girdi değeri e = logaritma değeri

4.19.4. Eğitim Algoritması

Problemin karmaşıklığına göre uygun eğitim algoritması seçmek ağın performansını arttırmaktadır. Çok katmanlı ileri beslemeli ağlarda kullanılan bazı eğitim algoritmaları şunlardır: Levenberg-Marquart, Bayesian Regularization, Scaled Conjugate Gradient, BFGS Quasi-Newton, Resilient Backpropagation, Conjugate Gradient with Powell/Beale Restarts, Fletcher-Powell Conjugate Gradient, Polak-Ribière Conjugate Gradient, One Step Secant, Variable Learning Rate Backpropagation. Bu algoritmalarından Levenberg-Marquart (LM) ve Bayesian Regularization (BR) algoritması farklı optimizasyon yöntemleri kullanarak performansı geliştiren yüksek performanslı geri yayılım eğitim algoritmalarıdır [45].

Levenberg-Marquart Algoritması: Doğrusal olmayan fonksiyonların karelerinin toplamı şeklinde ifade edilen, çok değişkenli bir fonksiyonların minimumunu belirlemeye çalışır [46]. LM algoritması Gradient-Descent (Gradyan İniş) ve Gauss-Newton algoritmalarının harmanlanmış halidir. Gauss-Newton algoritmasının hız avantajını, Gradyan iniş algoritmasının kararlılığını devralmıştır. Ayrıca Gauss-Newton algoritmasından daha iyi yakınsayabilir [47].

Bayesian Düzenleştirme Algoritması: LM algoritması ile ağırlık ve tahmin değerlerinin düzenlenerek güncelleme yapar [41]. BR algoritmasında işlemler daha uzun sürer ama ağın ezberleme olasılığı azalır ve küçük ya da gürültülü veriler için daha iyi sonuçlar verir.

Eğitim algoritmalarının yanında Matlab'ın sinir ağını eğitmek için sunduğu seçenekler de vardır:

1) 'sgdm' eğitim seçeneğinin matlab kodu:

```
options = trainingOptions('sgdm',...
    'LearnRateSchedule','piecewise',...
    'LearnRateDropFactor', 0.2, ...
    'LearnRateDropPeriod', 5, ...
    'MaxEpochs', 20, ...
    'MiniBatchSize', 64, ...
    'Plots','training-progress')
```

2) 'adam' eğitim seçeneği için matlab kodu:

```
options = trainingOptions('adam', ...
    'InitialLearnRate',3e-4, ...
    'GradientDecayFactor',0.90, ...
    'MaxEpochs',20, ...
    'MiniBatchSize',64, ...
    'Plots','training-progress')
```

3) 'rmsprop' eğitim seçeneği için matlab kodu:

```
options = trainingOptions('rmsprop', ...
    'InitialLearnRate',3e-4, ...
    'SquaredGradientDecayFactor',0.99, ...
    'MaxEpochs',20, ...
    'MiniBatchSize',64, ...
    'Plots','training-progress')
```

Bu üç eğitim seçenekleri ile de ağın iyileştirilmesi yapılabilir.

4.19.5. Performans Ölçütleri

Ağın başarısını ölçmek için bazı performans ölçüleri kullanılmaktadır. Bunlardan en çok kullanılanı Hata Kareler Ortalaması (MSE) dir. Hata Kareler Ortalamasının Karekökü

(RMSE), Hatanın Mutlak Ortalaması (MAE) ve Ortalama Mutlak Yüzde Hata (MAPE) de yine kullanılan diğer ölçütlerdir. Bu tezde MSE, MAPE ve MAE kullanılmıştır. Çizelge 4.4'te performans ölçütlerinin formülleri verilmiştir.

Çizelge 4.4. Performans ölçütleri ve formülleri

Performans Ölçütleri	Formülü
Hata Kareleri Ortalaması	$MSE = \frac{1}{n} \sum_{t=1}^n e_t^2$
Hatanın Mutlak Yüzde Ortalaması	$MAPE = \frac{\%100}{n} \sum_{t=1}^n \left \frac{e_t}{y_t} \right $
Hatanın Mutlak Ortalaması	$MAE = \frac{1}{n} \sum_{t=1}^n e_t $

4.20. NARX ve NAR Ağlarının Uygulama Alanları

NARX ve NAR ağları normal yapay sinir ağlarından farklı olarak içinde zaman serisi barındıran ve karmaşık problemlere çözüm getiren ağlardır. Bundan dolayı birçok konuya uygulanmış ve başarılı sonuçlar getirmiştir.

Enerji alanındaki bir çalışmada bulut örtüsünün (cloud cover) ışımsal soğutma modeli (radiative cooling) üzerindeki etkisi araştırılmıştır. Araştırma bulut örtüsü dikkate alınarak sistem sıcaklığını farklı iklim koşullarında tahmin etmek için NARX yöntemini kullanmıştır [48]. Bir başka çalışmada rüzgar enerjisi rüzgar hızından etkilendiğinden rüzgar hızları tahmin edilmek istenmiştir. Üç farklı bölge için rüzgar hızı verilerini bir saatlik olarak toplayıp NARX ve NAR yöntemleri ile 30 günü tahmin etmeye çalışılmıştır. Bu çalışmada ayrıca iki aktivasyon fonksiyonu olan tansig ve logsig'i de karşılaştırılmıştır. Sinir ağındaki nöron sayısını deneme yanılma yöntemi ile bulunmuş ve Levenberg Marquardt eğitim algoritması kullanılmıştır. Önerilen modelleri değerlendirmek için de MAE, MAPE ve RMSE ölçütleri kullanılmıştır. Sonuçlara göre tansig fonksiyonu logsig fonksiyonuna göre iki sinir ağı modeli için de daha iyi sonuç verdiği görülmüştür [49]. Rüzgar hızlarıyla ilgilenen başka bir çalışmada da Zonguldak ilindeki bir dakikalık zaman serileri kullanılarak farklı gizli katman ve nöron sayıları ile rüzgar hızlarının tahmin edilmek istenmiştir. Modelin

başarısı MAE, MSE ve RMSE ile ölçülmüştür. Tahmin için NAR ve filtreleme yöntemi ile pencere fonksiyonlarını kullanarak karşılaştırılmış ve diğer yöntemlerden daha başarılı bulunmuştur [50]. Rüzgar hızı tahmin çalışmasının dışında enerji alanındaki bir diğer çalışmada Ankara ili için güneş ışınım tahmini NARX yöntemi kullanılmıştır. Güneş ışınım verilerini etkileyen yağış, sıcaklık ve nem verileri de değişken olarak kullanılmıştır. Harici veri olarak kullanılan yağış, sıcaklık ve nem verileri ayrı ayrı değerlendirilmiş ve sonuçlara göre sıcaklık verisiyle olan analizin daha iyi sonuç verdiği bulunmuştur [51].

Çevre alanındaki bir çalışmada Puglia bölgesindeki 76 kuyunun yeraltı suyu seviyelerinin tahmininde NARX den yararlanılmıştır. Bayesian Regularization (BR), Levenberg Marquard (LM) ve Scaled Conjugate Gradient (SCG) eğitim algoritmaları kullanarak en başarılı NARX ağı bulunmaya çalışılmıştır. NARX-BR diğerlerine göre daha iyi performans göstermiştir [52]. Hava kalitesi tahmini yapan başka bir çalışmada Çanakkale'nin dört ilçesi için hava kalitesini NARX kullanarak tahmin etmeyi amaçlamışlar. Günlük olarak toplanan PM10 ve SO2 parametreleri NARX'in girdileri olarak modellenmiş ve yine günlük olarak toplanan NO, NO2, NOX ve O3 verileri ise sırasıyla NARX'in çıktıları olarak modellenmiştir. Tasarlanan model %90 ın üzerinde başarı göstermiştir [53].

Ekonomi alanında önemli kavramlardan biri olan ticari kredi faiz oranlarını aylık veriler şeklinde ve bu oranı etkileyen makroekonomik bazı değişkenlerle NARX ve Vektör Otoregresif (VAR) modelleri ile tahmin edilmeye çalışılmıştır. Elde edilen sonuçlara göre NARX yönteminin VAR yönteminden daha başarılı olduğu bulunmuştur [54].

Görüntü işleme alanında yapılan bir çalışmada ise kaotik bir hareketin yapay sinir ağları ile yeniden modellenmeye çalışılması amaçlanmıştır. Ardışık görüntülerdeki nesnelerin hareketleri görüntü işleme ile belirlenip konumları ile NAR da eğitilmiştir. Ağın eğitiminde Levenberg-Marquardt geri yayılım algoritması olarak kullanılmıştır. Ağ eğitildikten sonra test edilerek yeniden bir kaotik hareket elde edilmiştir. Sonuç olarak yeniden üretilen model gerçek modelle aynı özellikleri ve beklenen davranışları gösterdiği görülmüştür [55]. Son olarak, tezin konusuna benzer bir çalışma da Amerika da seçtikleri bir otoyol üzerindeki iki bağıntı (güney ve kuzey) için 15 dakikalık aralıklarla toplanan seyahat sürelerinin tahmini için NAR ve NARX yöntemlerini kullanarak tahmin yapmayı amaçlamışlardır. Seyahat sürelerini etkileyen faktörlerden haftanın günleri (Pazar"1" ve cumartesi "7") ve günün saatleri (00:00-01:00 "1" ve 11:00-12:00 "24") ve hafta sonu ve hafta içi (pazartesten cumaya "1", hafta sonu "0") şeklinde giriş verileri olarak kullanılmıştır. Modellemeler Matlab uygulaması üzerinden yapılmış ve ağların eğitilmesinde Levenberg-Marguardt

algoritması kullanılmıştır. NARX sonuçlarına göre kuzey bağıntısı toplam R^2 değeri %87, güney bağıntısı toplam R^2 değeri %79 olarak bulunmuştur. NAR sonuçlarına göre de kuzey bağıntısı toplam R^2 değeri %86, güney bağıntısı toplam R^2 değeri %84 olarak bulunmuştur. İki modelin karşılaştırılmasında ayrıca MSE değerlerine de bakılmış ve R^2 değerleri ile benzerlik göstermiştir [56].

5. SEYAHAT SÜRESİ GÜVENİRLİK ANALİZİ VE NAR VE NARX SİNİR AĞLARI İLE TAHMİNİ

5.1. Çalışmanın Amacı

Bu uygulamanın amacı, New York şehrinde bulunan La Guardia havalimanı ile JFK havalimanları arası taksi ile yapılan yolculukların seyahat sürelerinin dağılımını ve güvenilirlik durumunu belirlemek ve daha sonra sıcaklık, rüzgar ve görüş mesafesi değişkenlerini de ekleyerek seyahat süresinin tahminini yapmaktır. Şekil 5.1’de iki havalimanının harita görüntüsü verilmiştir.



Şekil 5.1. JFK ve LaGuardia ve 23 Aralık 2015 günü yapılan yolculukların başlangıç noktaları [57]

5.2. Verinin Hazırlanması

Kullanılan veri 2015 yılına ait New York ‘un sarı taksi yolculuk kayıtlarından oluşmaktadır. Veri seti (<https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page%20>) [58] internet adresinden alınmış olup aşağıdaki değişkenleri içermektedir:

- Vendor id: Hangi vericiden sinyal aldığı (1= Creative Mobile Technologies, 2= Verifone Inc)
- t_pickup: Taksimetrenin devreye girdiği saat ve tarih

- t_dropoff: Taksimetreinin sona erdiği saat ve tarih
- num_passenger: Yolcu sayısı
- trip_distance :Yolculuk mesafesi
- l_pickup_lon: Taksinin yolcuyu aldığı yerin boylamı
- l_pickup_lat: Taksinin yolcuyu aldığı yerin enlemi
- ratecodeid: Mesafeye göre ücretlendirme (1= Standart rate, 2=JFK, 3=Newark,4= Nassau veya Westchester, 5=Negatiated, 6=Group ride)
- flag_store: Yolcunun daha önce kayıtlı olup olmadığı (Y=Kayıtlı, N= Kayıtlı değil)
- l_dropoff_lon: Taksinin yolcuyu bıraktığı yerin boylamı
- l_dropoff_lat: Taksinin yolcuyu bıraktığı yerin enlemi
- payment_type: Ödeme tipi (1= Kredi kartı, 2= Nakit, 3= Ücretsiz, 4= Tartışma, 5=Bilinmiyor, 6= İptal edilmiş yolculuk)
- fare_amount: Sayaç tarafından hesaplanan zaman ve mesafe ücreti
- extra: Çeşitli ekstralar ve ücretler (gece tarifesi gibi)
- mta_tax: vergi
- surcharge: iyileştirme ek ücreti
- tip: bahşiş
- talls: Tüm geçiş ücretlerinin toplam tutarı
- total: Alınan toplam ücret

Çalışılan veri setinde yaklaşık 144 milyon veri bulunmaktadır. Bu veriler PostgreSQL veri tabanına aktarılmıştır. PostgreSQL veri tabanına eklenmesinin sebebi PostGIS ile konumsal olarak sorgu yapabilme imkanını vermesidir. Verilerin Ocak ayından alınan örnek tablosu Şekil 5.2 ve Şekil 5.3'te gösterilmiştir.

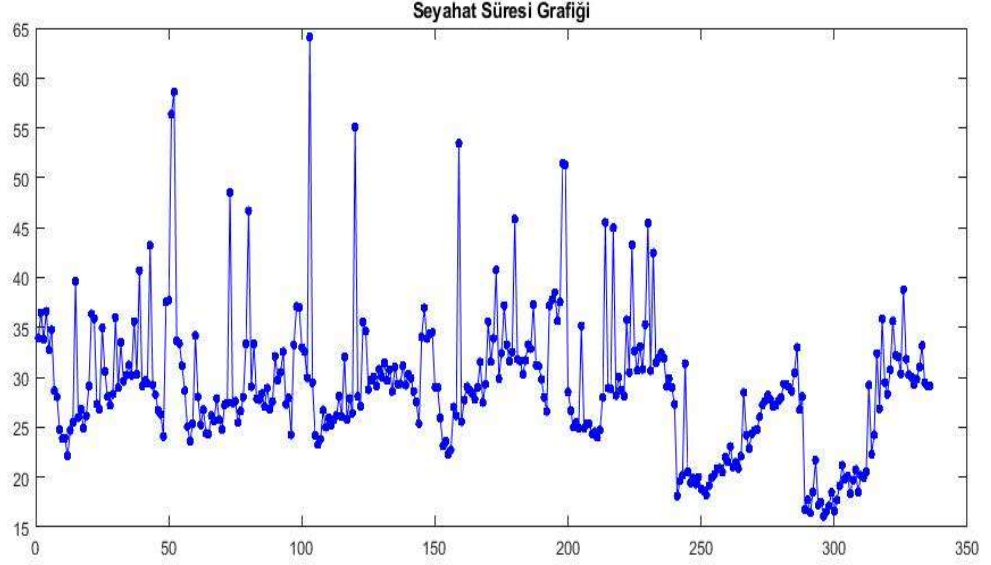
id integer	vendorid character	t_pickup timestamp without time	t_dropoff timestamp without time	num_passengers smallint	trip_distance real	l_pickup_lon double precision	l_pickup_lat double precision	ratecodeid character(2)	flag_store character(1)	l_dropoff_lon double precision	l_dropoff_lat double precision
1	2	2015-01-15 19:05:39	2015-01-15 19:23:42	1	1.59	-73.993896484375	40.7501106262207	1	N	-73.9747848510742	40.750617980957
2	1	2015-01-10 20:33:38	2015-01-10 20:53:28	1	3.3	-74.0016479492188	40.7242431640625	1	N	-73.9944152832031	40.7591094970703
3	1	2015-01-10 20:33:38	2015-01-10 20:43:41	1	1.8	-73.9633407592773	40.8027877807617	1	N	-73.9518203735352	40.8244132995605
4	1	2015-01-10 20:33:39	2015-01-10 20:35:31	1	0.5	-74.0090866088867	40.7138175964355	1	N	-74.0043258666992	40.7199859619141
5	1	2015-01-10 20:33:39	2015-01-10 20:52:58	1	3	-73.9711761474609	40.7624282836914	1	N	-74.0041809082031	40.7426528930664
6	1	2015-01-10 20:33:39	2015-01-10 20:53:52	1	9	-73.8743743896484	40.7740478515625	1	N	-73.9869766235352	40.7581939697266
7	1	2015-01-10 20:33:39	2015-01-10 20:58:31	1	2.2	-73.9832763671875	40.7260093688965	1	N	-73.9924697875977	40.7496337890625
8	1	2015-01-10 20:33:39	2015-01-10 20:42:20	3	0.8	-74.00262626586914	40.7341423034668	1	N	-73.9950103759766	40.726325987695
9	1	2015-01-10 20:33:39	2015-01-10 21:11:35	3	18.2	-73.7830429077148	40.6443557739258	2	N	-73.9875946044922	40.7593574523926
10	1	2015-01-10 20:33:40	2015-01-10 20:40:44	2	0.9	-73.9855890737305	40.7679481506348	1	N	-73.9859161376953	40.7593650817871
11	1	2015-01-10 20:33:40	2015-01-10 20:41:39	1	0.9	-73.9886169433594	40.7231025695801	1	N	-74.00439453125	40.7285842895508
12	1	2015-01-10 20:33:41	2015-01-10 20:43:26	1	1.1	-73.993782043457	40.7514190673828	1	N	-73.9674072265625	40.7572174072266
13	1	2015-01-10 20:33:41	2015-01-10 20:35:23	1	0.3	-74.0083618164063	40.7043762207031	1	N	-74.0097732543945	40.7077255249023
14	1	2015-01-10 20:33:41	2015-01-10 21:03:04	1	3.1	-73.9739456176758	40.7604484558105	1	N	-73.9973449707031	40.7352104187012
15	1	2015-01-10 20:33:41	2015-01-10 20:39:23	1	1.1	-74.006721496582	40.7317771911621	1	N	-73.995216366289	40.7398948669434

Şekil 5.2. Veri tabanı görüntüsü 1

payment_type character(1)	fare_amount real	extra real	mta_tax real	surcharge real	tip real	tolls real	total real	l_pickup geometry(Point,4326)	l_dropoff geometry(Point,4326)
1	12	1	0.5	3.25	0	0.3	17.05	0101000020E610000000000009C7F52C0000000A003604440	0101000020E6100000000000E0627E52C00000004014604440
1	14.5	0.5	0.5	2	0	0.3	17.8	0101000020E610000000000001B8052C00000000B45C4440	0101000020E610000000000080A47F52C0000000802A614440
2	9.5	0.5	0.5	0	0	0.3	10.8	0101000020E610000000000060A77D52C0000000C0C1664440	0101000020E6100000000000A0EA7C52C00000006086694440
2	3.5	0.5	0.5	0	0	0.3	4.8	0101000020E6100000000000E0948052C0000000605E5B4440	0101000020E6100000000000E0468052C000000080285C4440
2	15	0.5	0.5	0	0	0.3	16.3	0101000020E6100000000000C027F52C00000004097614440	0101000020E610000000000080448052C0000000400F5F4440
1	27	0.5	0.5	6.7	5.33	0.3	40.33	0101000020E6100000000000C0F57F52C00000000014634440	0101000020E6100000000000A02A7F52C0000000800C614440
2	14	0.5	0.5	0	0	0.3	15.3	0101000020E610000000000000E7F52C00000000E0ED5C4440	0101000020E6100000000000A0847F52C000000000F45F4440
1	7	0.5	0.5	1.66	0	0.3	9.96	0101000020E6100000000000A02B8052C000000060F85D4440	0101000020E610000000000040AE7F52C000000040F85C4440
2	52	0	0.5	0	5.33	0.3	58.13	0101000020E6100000000000601D7252C0000000407A524440	0101000020E6100000000000C0347F52C0000000A032614440
1	6.5	0.5	0.5	1.55	0	0.3	9.35	0101000020E6100000000000E0137F52C0000000204C624440	0101000020E610000000000040197F52C0000000E032614440
1	7	0.5	0.5	1.66	0	0.3	9.96	0101000020E610000000000080457F52C0000000A08E5C4440	0101000020E6100000000000408052C000000040425F4440
1	7.5	0.5	0.5	1	0	0.3	9.8	0101000020E610000000000209A7F52C0000000802E604440	0101000020E6100000000000EA7D52C000000080E604440
2	3	0.5	0.5	0	0	0.3	4.3	0101000020E61000000000008098052C00000000295A4440	0101000020E61000000000020A8052C0000000C0965A4440
1	19	0.5	0.5	3	0	0.3	23.3	0101000020E61000000000020557E52C00000006056614440	0101000020E61000000000080D47F52C0000000601B5E4440
2	6	0.5	0.5	0	0	0.3	7.3	0101000020E610000000000206E8052C0000000E0A5D4440	0101000020E610000000000A0B17F52C0000000E0B45E4440

Şekil 5.3. Veri tabanı görüntüsü 2

Bu çalışmada kullanılanlar değişkenler t_pickup, t_dropoff, l_dropoff_lon, l_dropoff_lat, l_pickup_lon, l_pickup_lat dır. Bu değişkenleri kullanarak Python aracılığıyla haftanın her günü için 15'er dakikalık olmak üzere veriler çekilmiştir. Verilerden hatalı olanlar (aşırı uzun süren seyahatler, boş olanlar, negatif değerli yolculuklar) analize dahil edilmemiştir. Birinci bölüm veri hazırlığından sonra tahmin için ayrıca veriler toplanmıştır. Seyahat süresini etkileyebilecek sıcaklık, görüş mesafesi ve rüzgar gibi değişkenler de (<https://www.timeanddate.com/weather/usa/new-york/historic?month=1&year=2015>) [59] siteden çekilerek eklenmiştir. Bu değişkenler de yine 15 dk'lık olacak şekilde alınmıştır. Seyahat süresinin grafiği Şekil 5.4'te verilmiştir. NARX ve NAR analizi için veriler ayrıca normalleştirilmiştir.



Şekil 5.4. Seyahat süresi grafiği

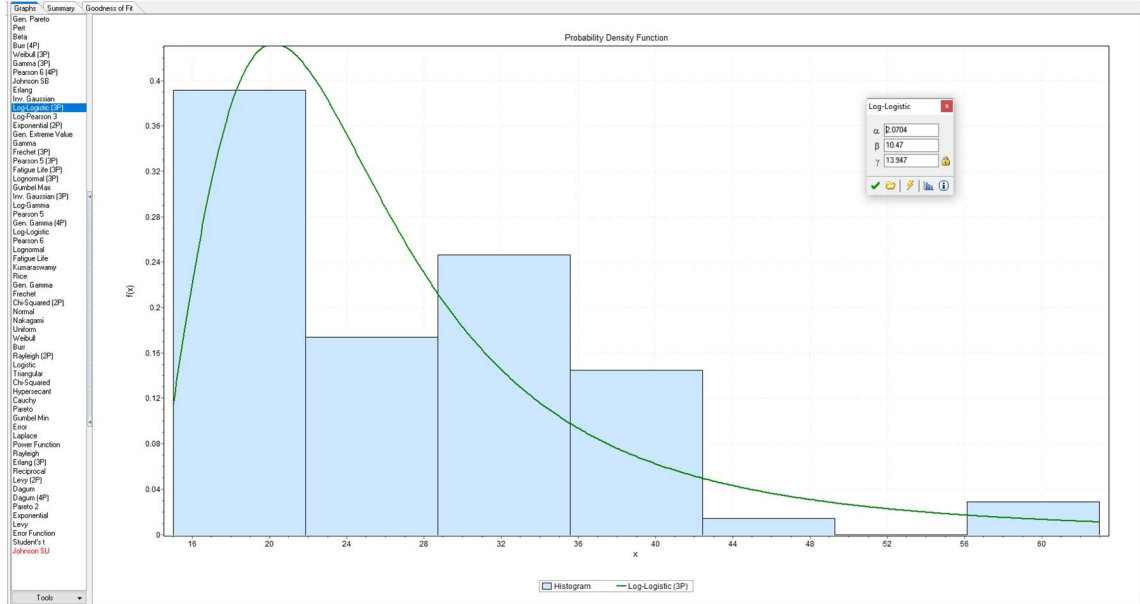
5.3. Seyahat Sürelerinin Dağılım Analizleri

Easy Fit programı yardımıyla önceki kısımda elde edilen yolculuk sürelerinin en çok hangi dağılıma uyum sağladığı tespit edilmiştir. Bu programın kullanılmasının sebebi diğer programlara göre daha geniş dağılım seçeneklerine sahip olması ve bu dağılımların içinde seyahat süreleri için ihtiyaç duyulan dağılımları barındırmasıdır. Şekil 5.5'te easy fit programına ait ekran görüntüsü verilmiştir.

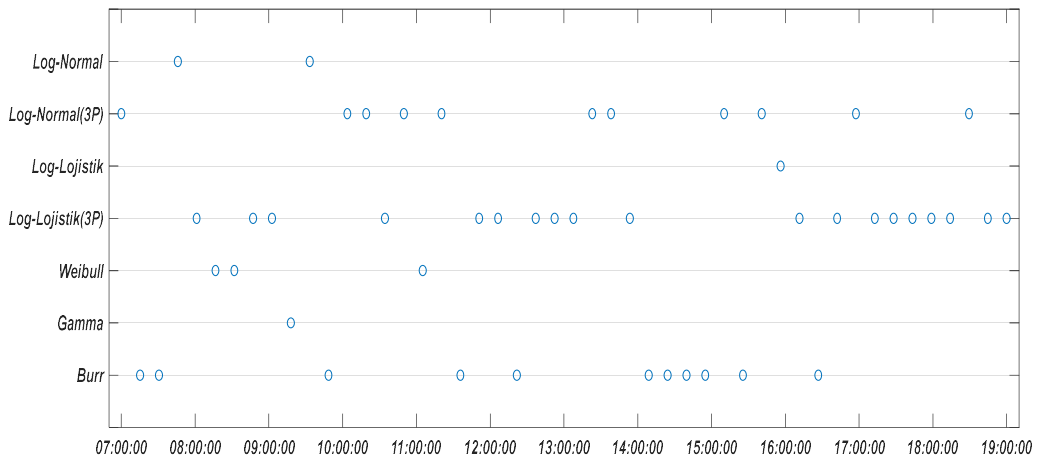
Verinin hangi dağılıma uyduğunu görmek için örneklem büyüklüğüne göre uyum iyiliği testlerinden Kolmogorov-Smirnov ve Anderson-Darling testlerinin %5 önemlilik düzeyi baz alınmıştır. Kolmogorov- Smirnov testi örneklem büyüklüğünün 35'ten büyük olduğu durumlarda kullanılması daha uygun olduğu için kullanılmıştır. Anderson-Darling testi de yirmi ile elli arası örneklem büyüklüğüne uygun olduğu için kullanılmıştır [60]. Haftanın her günü için ve belirlenen zaman aralıklarında elde edilen yolculuk sayıları (örneklem büyüklüğü), Ek 1'de belirtilmiştir. Her bir veri seti için her dağılımın olasılık değerine bakılmıştır. Eğer olasılık değeri 0.05'ten küçükse veri o dağılıma uymuyor, 0.05'ten büyükse de veri bahsedilen dağılımlardan birine uyuyor olarak kabul edilmiştir.

Bütün günlerin her zaman aralığı için dağılımı bulunduktan sonra yüzdelendirilmiştir. Örneğin Pazartesi günü on dokuz tane Log-Lojistik (3P) dağılımına uyan zaman dilimi varsa, $\frac{19}{48} * 100 = \%39,58$ ile Pazartesi gününün Log-Lojistik ile modellenebileceği kabul edilmiştir. Pazartesi günü için incelenen tüm zaman aralıklarında hangi dağılımın baskın

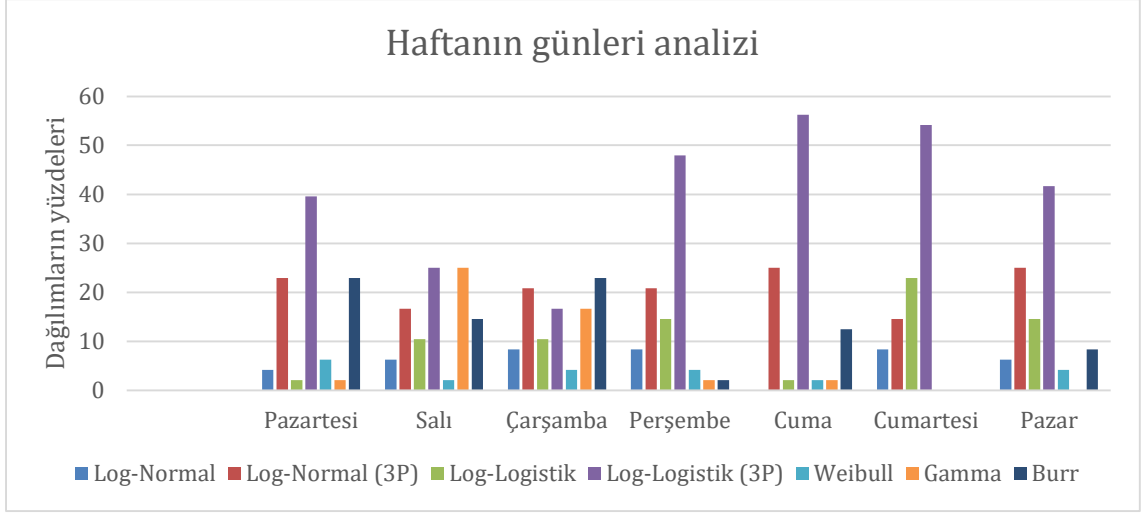
dağılım olduğu Şekil 5.6’da gösterilmiştir. Haftanın her günü için hangi dağılımın baskın olduğunu gösteren grafikte Şekil 5.7’de verilmiştir.



Şekil 5.5. Pazartesi günü 12:00-12:15 verisi için Easy fit programının görüntüsü



Şekil 5.6. Pazartesi günü dağılımı



Şekil 5.7. Haftanın her günü için dağılımların yüzdeleri

5.4. Seyahat Sürelerinin Güvenilirlik Analizleri

Bu kısımda ilk olarak seyahat sürelerinin 95.yüzdeliklerinin sonucu sonra tampon süre indeksi daha sonra planlanan süre indeksi ve en son başarısızlık ve zamanında oranının deneysel çalışmaları sunulacaktır.

5.4.1. Seyahat Sürelerinin 95. Yüzdelik Sonucu

Salı günü 9:45-10:00, Çarşamba günü 8:30-8:45, 12:45-13:00 ve Cuma günü 8:15-8:30 da yaşanabilecek en kötü seyahat süreleri fazla olan süre dilimleridir. Bu süre dilimleri daha detaylı incelendiğinde Salı günü 9:45-10:00 arası 95. yüzdelik 46 yolculuktan iki yolculuk aşmıştır ve bu iki yolculuğun aynı günlerde olmadığı görülmüştür.

Çarşamba 8:30-8:45 süre dilimi incelendiğinde ise 65 yolculuktan üç yolculuk 95. yüzdelik değerini aşmaktadır ve bu yolculukların ikisinin aynı ay içinde olduğu ama aynı gün içerisinde olmadığı görülmüştür.

Aynı ay içerisinde olan yolculuklardan biri 2 Aralık diğer ise 20 Aralık tarihindedir. Çarşamba 12:45-13:00 aralığında ise toplam 74 yolculuktan üçü 95'lik yüzdeyi aşmaktadır ve bu yolculuklardan ikisi aynı gün (28 Ekim) öğle saatlerinde çok yakın saatlerde gerçekleşmektedir.

Hafta içi bir diğer bakılan saat ise cuma 8:15-8:30 arasındır. Bu aralıkta 89 yolculuktan dördü 95'lik dilimi aşmaktadır ve bu dört yolculuktan ikisi 8 Mayıs sabah saatlerinde aralarında bir dakikalık süre farkıyla hareket etmektedirler.

Hafta sonu ise Cumartesi ve Pazar günlerinin öğleden sonraki zamanlarında kötü bir seyahat süresiyle karşılaşılması olasıdır. Cumartesi 18:15-18:30 arasında hafta sonu için en fazla olan süre bu süre dilimindedir. Bu süre dilimi de ayrıntılı incelendiğinde yirmi iki yolculuktan bir tanesi 95'lik dilimi aşmaktadır. Hafta sonu ikinci en kötü seyahat süresi ise cumartesi günü 13:30-13:45 arasındadır. Bu aralıktaki 75 yolculuğun dördü 95'lik süre dilimini aşmaktadır ve bu dört yolculuktan ikisi aynı güne denk gelmektedir. 16 Mayıs tarihine denk gelen iki yolculuk öğle saatlerinde aralarında yaklaşık on dakikalık farkla hareket etmektedirler. Aynı günlere denk gelen yolculukların hepsine bakıldığında denk geldiği günlerle ilgili herhangi özel bir şey bulunmamaktadır. Bu gecikmeler kazalardan, yol çalışmalarından ya da başka bir nedenden dolayı olabilir.

5.4.2. Tampon Süre İndeksi Sonucu

Hafta içi ortalamaya eklenmesi gereken en çok süre Salı günü 9:45-10:00 arasındadır. Orana göre gezi ortalama değerinden yaklaşık iki kat daha uzun sürebileceğini gösteriyor. Bu süre diliminden sonra ikinci olarak cuma günü 9:30-10:00 arası gelmektedir. Çarşamba ve Perşembe günleri öğle saatleri için de eklenmesi gereken sürenin fazla olduğu söylenebilir. Ortalamaya en az eklenmesi gereken süre dilimi de Cuma günü 16:45-17:00 saat dilimidir. Genel olarak da sabah saatleri akşam saatlerine göre daha az eklenmesi gereken süre olduğu görülmüştür. Hafta sonu değerlerine göre ise Cumartesi günü 13:30-13:45 arası en fazla ortalamaya eklenmesi gereken süredir. Bu değer hafta içi eklenecek değer neredeyse yarısı kadardır. En az eklenmesi gereken süre ise yine cumartesi günü 7:45-8:00 arasındadır.

5.4.3. Planlanan Süre İndeksi Sonucu

Planlanan süre indeksinin hesaplanabilmesi için serbest akış süresinin hesaplanması gerekmektedir. Bu ölçütün doğru olması için birden fazla gün kullanılmıştır. Bu günler:

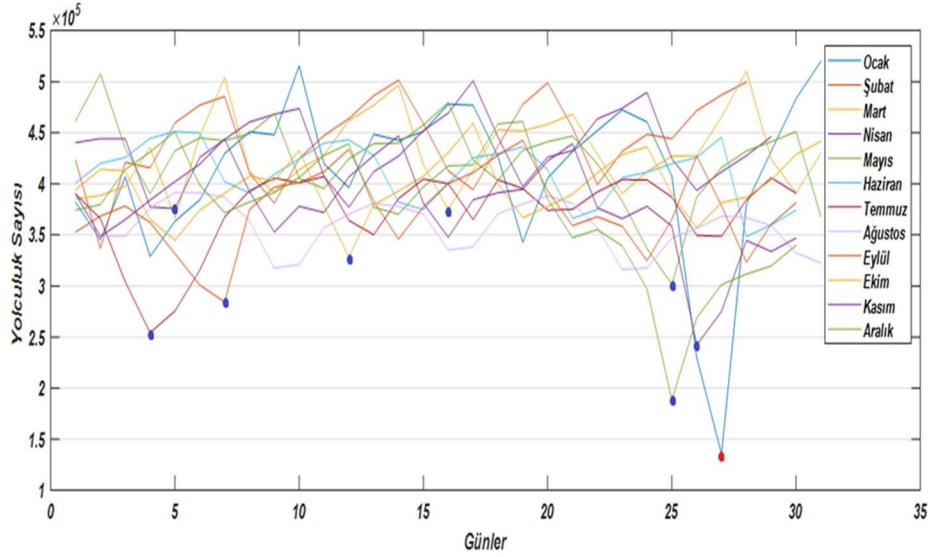
- 19 Ocak : Martin Luther King'in doğum günü
- 17 Mart : St.Patrick günü
- 5 Nisan : Paskalya
- 25 Mayıs : Anma günü
- 4 Temmuz : Bağımsızlık günü
- 7 Eylül : İşçi Bayramı

- 12 Ekim : Kolomb Günü
- 26 Kasım : Şükran günü
- 25 Aralık : Noel

Bu günler Şekil 5.8’de mavi noktayla işaretlenmiştir. Mavi noktaların dışındaki kırmızı nokta ise 27 Ocak gününü göstermektedir.

27 Ocak gününde de düşüş yaşanmıştır. Ancak bu düşüş o günlerde gerçekleşen kar fırtınasından kaynaklanmaktadır. Bir başka deyişle diğer özel günler gibi her sene aynı gün düşüş yaşanmayabilir. Bu nedenle 27 Ocak günü serbest akış süresi hesaplamaya katılmamıştır.

Serbest akış süresinin ölçmek için yukarıda bahsedilen günleri diğer günlerdeki gibi 15 dakikalık aralıklarla 7:00- 19:00 arasında olmak üzere seyahat süreleri toplanmış ve her bir süre aralığı için ortalama seyahat süresi bulunmuştur. Bu süreler serbest akış seyahat süreleri denilebilir.



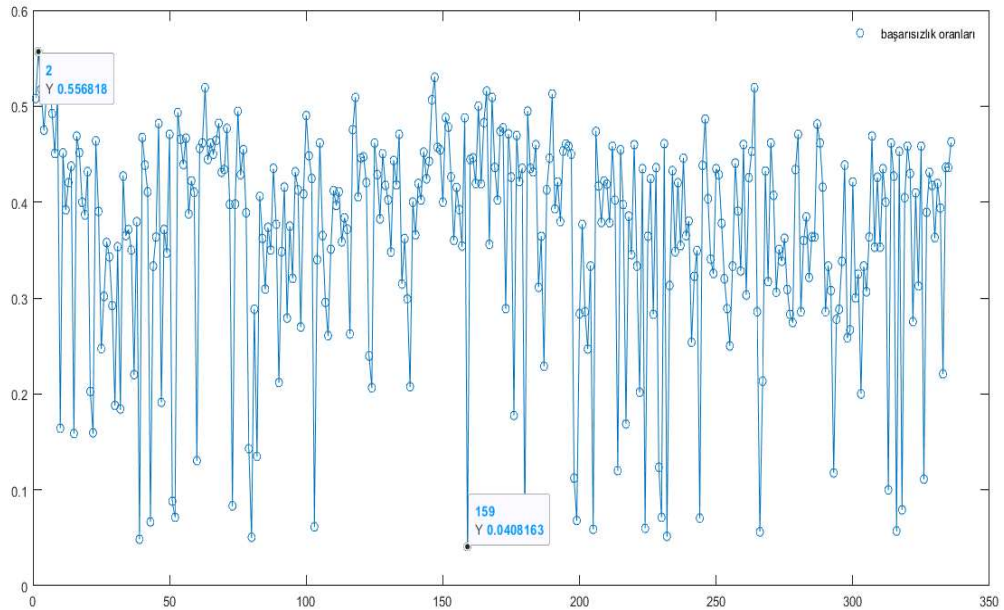
Şekil 5.8. Özel günlerin yolculuk sayıları

Planlanan süre indeksinin en fazla olduğu süre dilimi çarşamba günü 8:30-8:45 arasındadır. Tampon süre indeksinin en fazla olduğu gün ve süre dilimi planlanan sürede da yüksektir.

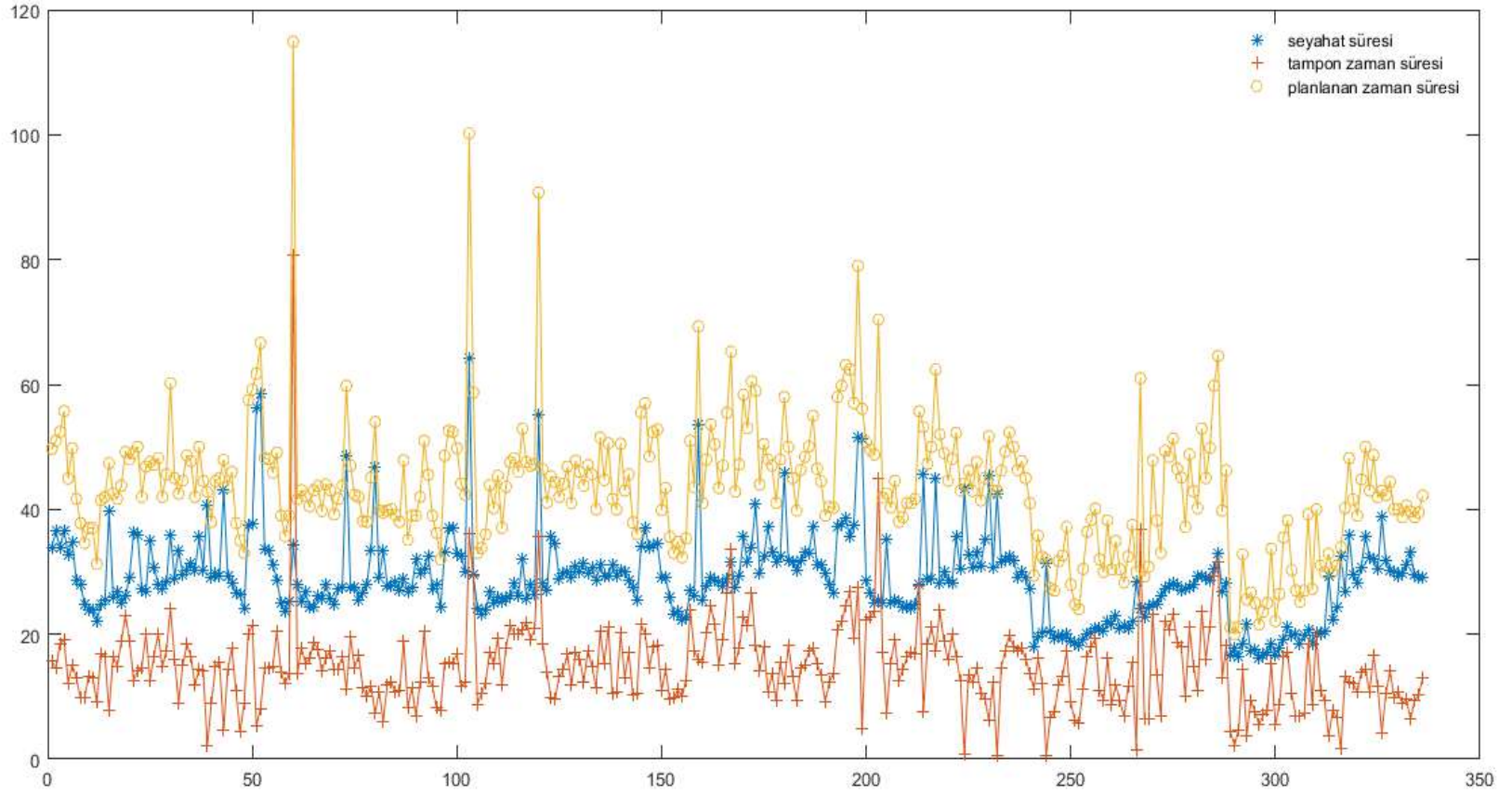
Bu da güvenilirlik ölçütlerinin birbirleriyle ilişkili olduklarını göstermektedir. Hafta içi en düşük değer ise Salı günü 16:45-17:00 arasındadır. Cumartesi günü 18:00-18:15 arası en fazla, Pazar da 7:15-7:30 arası en az değeri göstermektedir.

5.4.4. Başarısızlık ve Zamanında Oranı Sonucu

Başarısızlık oranı bir seyahatte ortalama seyahatten daha fazla olan seyahatlerin tüm seyahatlere bölümü, zamanında oranı ise ortalama altı yapılan seyahatlerin tüm seyahatlere bölümü olarak ifade edilmiştir. Başarısızlık ve zamanında oranı birbirini tamamlayan oranlardır. Haftanın her günü için başarısızlık oranlarına bakıldığında pazartesi günü sabah 07:15 ile 07:30 arasında yapılan yolculukların oranı daha fazla çıkmıştır. Bu zaman diliminden sonra yine aynı gün için 08:15-8:30 arası oran daha yüksektir. Genel olarak sabah saatlerinin başarısızlık oranının daha yüksek olduğu görülmektedir. Perşembe günü 10:30-10:45 arası ise başarısızlık oranının en düşük olduğu daha sonra ise pazartesi 16:30-16:45 arasının başarısızlık oranının düşük olduğu görülmektedir. Oranın yüksek olduğu zamanlar daha çok günlerin sabah saatlerine denk gelmektedir. Bu orana ait grafik Şekil 5.9’da verilmiş ve en yüksek ve düşük oranlar işaretlenmiştir. Şekil 5.10’da da seyahat süresi mavi, tampon süre süresi kırmızı ve planlanan süre süresi de sarı ile gösterilmiştir.



Şekil 5.9. Başarısızlık oranı grafiği



Şekil 5.10. Tampon süresi ve planlanan süre seyahat süresi ile gösteren grafik

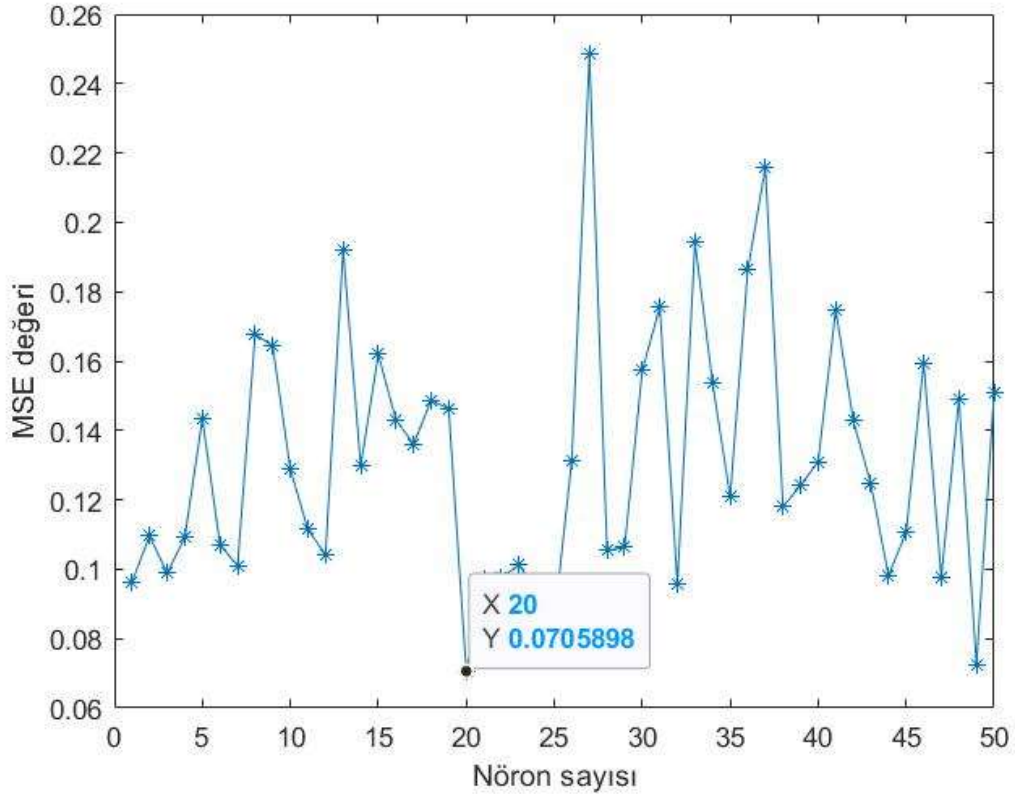
5.5. Yapay Sinir Ağları ile Seyahat Süresi Tahmini

Bölüm 5.3 ve 5.4'te seyahat süresi verisinin haftanın her günü için dağılımları ve güvenilirliği hesaplandı. Bu bölümde de Matlab 2020b programı yardımıyla NAR ve NARX sinir ağları kullanılarak tahmin çalışması yapılmıştır.

5.5.1. NAR-LM

NAR sinir ağı, dışsal girdiler olmadan sadece tahmin edilecek değişkenle tahmin yapan bir ağıdır. Burada sadece seyahat süresi kullanılarak tahmin yapılmaya çalışılmıştır.

Analize başlamadan önce veriler min-max yöntemi ile normalizasyonu yapılmıştır. Verilerin normalizasyon işlemi sonrası yeni girdi ve çıktı değerleri ile en iyi ara katman değerlerini bulmak için 1'den 60'a kadar bir döngü içerisinde performans ölçülerinden MSE değerlerine bakılarak ara katman nöron sayısı bulunmaya çalışılmıştır. Birkaç kez deneme yapıldıktan sonra Şekil 5.11'de gösterildiği gibi LM Algoritması için MSE değeri 0,0705 ile en düşük olan ara katman nöron sayısı 20 olarak bulunmuştur.



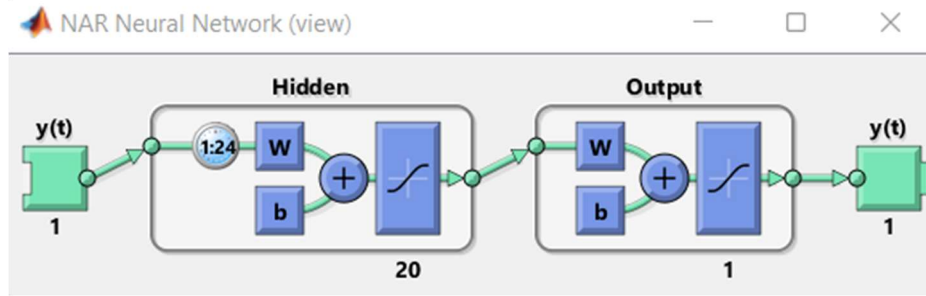
Şekil 5.11. LM algoritmasının nöron sayısını gösteren grafik

En iyi ara katman nöronu bulunduktan sonra diğer parametreler yapay sinir ağları ile 15'er dakikalık 48 adım tahmin bulunuldu. En iyi modeli bulabilmek için;

- Gecikme sayıları 3,6,12,24 olarak,
- Eğitim oranları % (Eğitim oranı 60, Doğrulama oranı 20, Test oranı 20), % (Eğitim oranı 70, Doğrulama oranı 15, Test oranı 15) ve % (Eğitim oranı 80, Doğrulama oranı 10, Test oranı 10) olarak,
- Eğitim seçenekleri 'adam', 'rmsprop' ve 'sgdm' olarak her biri tek tek denendi.

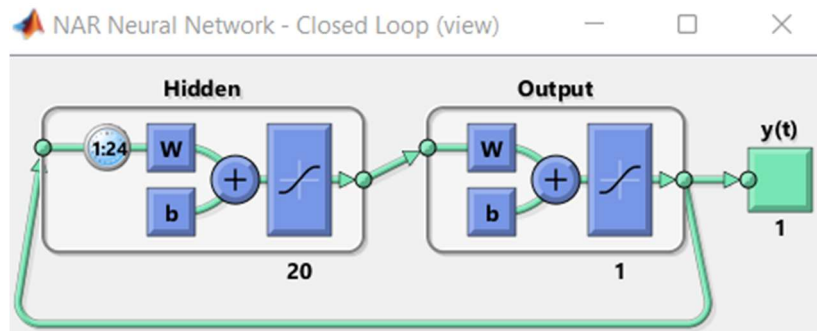
Çizelgeler de tüm denemeler verilmiştir. 36 denemenin mae eğitim, mae doğrulama, mae test, mape eğitim, mape doğrulama, mape test, mse eğitim, mse doğrulama, mse test ve R^2 eğitim, R^2 doğrulama ve R^2 test sonuçları gösterilmektedir.

Matlab'ın analiz sonuçlarına göre verdiği açık ve kapalı döngü şemaları Şekil 5.12 ve 5.13'te verilmiştir. Açık döngü yapısında bir değişken (seyahat süresi), 24 gecikme, 20 nöron sayısı ve her iki katmanda da tansig aktivasyon fonksiyonu olduğu görülmektedir.



Şekil 5.12. NAR YSA açık döngü mimarisi

Kapalı döngü ise tahmin sonuçlarının alındığı döngü olup, şeması açık döngünün şemasına benzerdir.



Şekil 5.13. NAR YSA kapalı döngü mimarisi

NAR ağının ‘adam’ eğitim seçeneğine ait çizelgesi aşağıda verilmiştir. Çizelge 5.1.’e göre en iyi model 24 gecikme ile eğitim oranı %80, doğrulama oranı %10 ve test oranı %10 olmaktadır. Eğitim, doğrulama ve test R^2 değerlerine bakıldığında değerlerin birbirine tutarlı olduğunu yani birinin çok yüksek diğerinin çok düşük çıkmadığı söylenebilir. Bu durumda aşırı öğrenme veya az öğrenme durumunun da olmadığı söylenebilir. MAE, MAPE ve MSE değerlerine bakıldığında da en küçük değer 24 gecikme, %(80,10,10) modelinde olmadığını ama bu modeldeki değerlerin de çok yüksek olmadığı söylenebilir. Oranlara göre bakılırsa sırasıyla ortalama başarı durumları en düşük % (60,20,20) sonra% (70,15,15) ve en yüksek % (80,10,10)’dir. Eğitim oranı arttıkça daha iyi modeller kurulabileceği söylenebilir.

Gecikme sayılarına göre bakıldığında ‘adam’ eğitim seçeneği için 24 gecikme sayısı ile ortalama olarak daha yüksek başarı elde edildiği sonra 6 gecikme ve 3 gecikmenin ortalama olarak birbirine yakın olduğunu daha sonra ise 12 gecikmenin geldiği söylenebilir. Gecikme sayılarının artması modelin başarısı üzerinde etkilidir denilemez.

Çizelge 5.1. NAR ağına ait adam eğitim seçeneğinin sonuçları

Eğitim seçenekleri	NAR - ADAM											
	3			6			12			24		
Gecikme	%60	%70	%80	%60	%70	%80	%60	%70	%80	%60	%70	%80
	%20	%15	%10	%20	%15	%10	%20	%15	%10	%20	%15	%10
	%20	%15	%10	%20	%15	%10	%20	%15	%10	%20	%15	%10
MAE/Test	0,1359	0,1308	0,1306	0,1477	0,1588	0,1588	0,1484	0,1538	0,1322	0,1439	0,1582	0,1328
MAE/Eğitim	0,1267	0,1075	0,1038	0,1112	0,1231	0,106	0,1138	0,1144	0,1106	0,1146	0,1198	0,1232
MAE/Doğrulama	0,1405	0,1174	0,1399	0,1502	0,1314	0,1227	0,0969	0,1138	0,1744	0,1067	0,1177	0,133
MAPE/Test	0,3023	0,2716	0,2944	0,2926	0,4147	0,2893	0,3189	0,3924	0,2377	0,2976	0,3962	0,2792
MAPE/Eğitim	0,0495	0,0479	0,0484	0,0421	0,0427	0,0386	0,0405	0,0417	0,0427	0,0471	0,0474	0,0424
MAPE/Doğrulama	0,3121	0,2438	0,2567	0,6416	0,4226	0,2844	0,0755	0,1136	0,6249	0,0542	0,1518	0,3182
MSE/Test	0,0396	0,036	0,0355	0,045	0,0527	0,0601	0,0651	0,0655	0,0328	0,0652	0,0455	0,0311
MSE/Eğitim	0,023	0,0452	0,0348	0,026	0,0405	0,024	0,0466	0,0357	0,0353	0,0365	0,038	0,0283
MSE/Doğrulama	0,0328	0,0382	0,0381	0,0512	0,0356	0,0382	0,0361	0,0351	0,0597	0,0333	0,0346	0,0314
Test R^2	0,6712	0,8593	0,8676	0,7378	0,7615	0,5406	0,6573	0,6143	0,8985	0,5301	0,7202	0,8035
Eğitim R^2	0,7278	0,6765	0,7535	0,7629	0,7246	0,7641	0,6877	0,7044	0,6973	0,734	0,7595	0,7704
Doğrulama R^2	0,7817	0,7876	0,8197	0,5983	0,7818	0,7554	0,8391	0,7891	0,5698	0,6724	0,8847	0,7449
Tüm R^2	0,7164	0,7328	0,766	0,7161	0,7397	0,7419	0,7013	0,7118	0,7214	0,6843	0,7727	0,7695

NAR ağının ‘rmsprop’ eğitim seçeneğine ait çizelgesi aşağıda verilmiştir. Çizelge 5.2 ye göre en iyi model 12 gecikme ile eğitim oranı %80, doğrulama oranı %10 ve test oranı %10 olmaktadır. Eğitim, doğrulama ve test R^2 değerlerine bakıldığında değerlerin birbirleriyle tutarlı olmadığı görülmektedir. Eğitim R^2 değeri test R^2 ve doğrulama R^2 değerlerinden büyüktür. Eğitim değerinin test değerinden büyük olması aşırı öğrenme olabileceği göstermektedir. Bundan dolayı 3 gecikme ve %(70,15,15) oranları ile ikinci R^2 değeri yüksek olan modele bakıldığında R^2 değerlerinin daha tutarlı olduğu söylenebilir. MAE, MAPE ve MSE değerlerine bakıldığında yine en küçük değerlere sahip olmadığı ve değerlerin çok yüksek olmadığı söylenebilir.

Oranlara göre bakıldığında ‘rmsprop’ için en düşük % (60,20,20) daha sonra % (80,10,10) ve % (70,15,15) oranlarının daha başarılı sonuçlar gösterdiği görülmüyor. Burada oranlar arttıkça ağın başarı durumunun arttığı söylenebilir

Gecikmelere bakıldığında ortalama en yüksek başarıyı 12 gecikme ile sağlamış ve sırasıyla 3 gecikme, 6 gecikme ve 24 gecikme ile devam etmektedir. Bu sonuçlardan gecikme sayısı arttıkça başarı oranının arttığını tam söylenemez.

Çizelge 5.2. NAR ağına ait rmsprop eğitim seçeneğinin sonuçları

Eğitim seçenekleri	NAR - RMSPROP											
	3			6			12			24		
Gecikme	%60	%70	%80	%60	%70	%80	%60	%70	%80	%60	%70	%80
Oran	%20	%15	%10	%20	%15	%10	%20	%15	%10	%20	%15	%10
	%20	%15	%10	%20	%15	%10	%20	%15	%10	%20	%15	%10
MAE/Test	0,1524	0,1589	0,1327	0,1463	0,1362	0,1587	0,1461	0,158	0,1584	0,147	0,1363	0,1548
MAE/Eğitim	0,1141	0,1115	0,1024	0,1115	0,1109	0,1049	0,1158	0,1193	0,1131	0,122	0,1187	0,1159
MAE/Doğrulama	0,1089	0,1248	0,1134	0,105	0,1072	0,1212	0,1053	0,1181	0,1438	0,105	0,1164	0,1169
MAPE/Test	0,308	0,402	0,2446	0,2921	0,2074	0,3229	0,2921	0,3992	0,3	0,2944	0,294	0,3955
MAPE/Eğitim	0,0312	0,0305	0,0317	0,0471	0,0451	0,0339	0,0458	0,0474	0,0464	0,0487	0,0471	0,04
MAPE/Doğrulama	0,1992	0,3549	0,2754	0,0735	0,2977	0,3328	0,0691	0,1809	0,3187	0,0583	0,306	0,1562
MSE/Test	0,0458	0,0658	0,0334	0,0654	0,0364	0,0458	0,0452	0,0459	0,0419	0,0452	0,0375	0,0455
MSE/Eğitim	0,0377	0,0272	0,0331	0,0358	0,0357	0,0339	0,0468	0,0371	0,0258	0,0482	0,0376	0,0466
MSE/Doğrulama	0,036	0,0335	0,035	0,0383	0,0339	0,0394	0,0317	0,0375	0,0586	0,0318	0,0349	0,0366
Test R^2	0,7323	0,6564	0,8853	0,6304	0,8146	0,6627	0,7302	0,7781	0,7177	0,6973	0,8062	0,7949
Eğitim R^2	0,7118	0,7723	0,6971	0,7113	0,6947	0,7018	0,6589	0,7277	0,8165	0,5769	0,745	0,6651
Doğrulama R^2	0,7312	0,7762	0,7363	0,801	0,7838	0,8123	0,7113	0,746	0,5857	0,7551	0,7144	0,7841
Tüm R^2	0,7149	0,7555	0,7124	0,6906	0,722	0,7203	0,6852	0,7359	0,7865	0,6337	0,744	0,6886

NAR ağının 'sgdm' eğitim seçeneğine ait çizelgesi aşağıda verilmiştir. Çizelge 5.3'e göre en iyi model 24 gecikme ile eğitim oranı %80, doğrulama oranı %10 ve test oranı %10 olmaktadır. Eğitim, doğrulama ve test R² değerlerine bakıldığında eğitim R² değeri test R² ve doğrulama R² değerlerinden büyüktür. Eğitim değeri test değerinden biraz büyüktür bu durumda aşırı öğrenme için net bir şey söylenemez ama doğrulama değerinin %63 lerede olması modelin bu eğitim seçeneği için çok iyi bir model olmadığı söylenebilir. Sgdm eğitim seçeneği için ikinci R² değeri yüksek olan model 12 gecikme %(80,10,10) modeldir. Bu modele bakıldığında da yine doğrulama değerinin eğitim ve test değerinden düşük olduğu görülüyor. NAR ağının en yüksek R² değeri sgdm eğitim seçeneğinde verdiği için en yüksek modele göre tahmin grafiği Şekil 5.16 da verilmiştir.

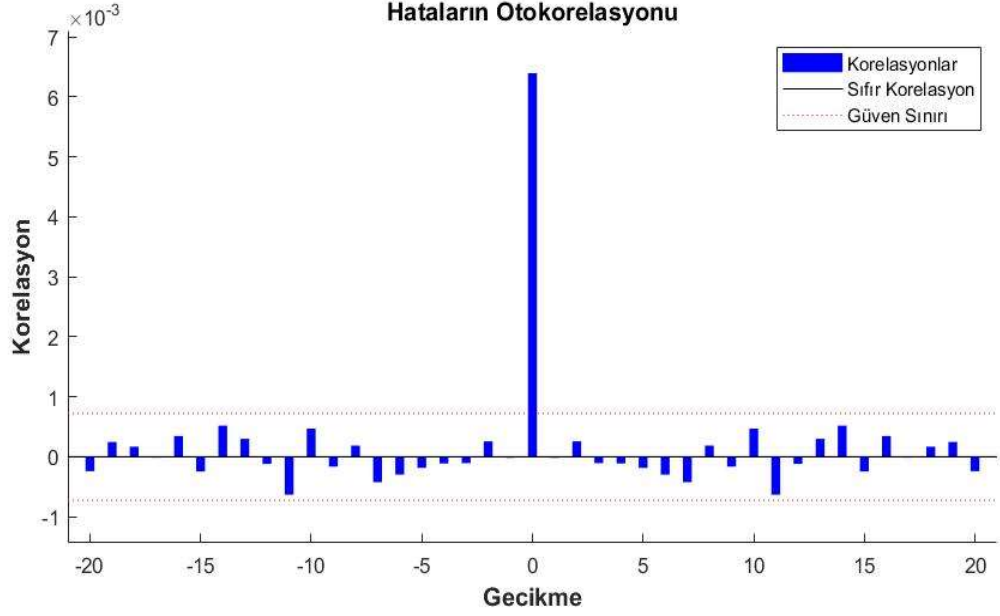
Oranlara göre değerlendirildiğinde sırasıyla ortalama başarı durumları en düşük % (60,20,20) sonra % (80,10,10) ve en yüksek % (70,15,15)'dir.

Gecikme sayılarına göre bakıldığında ise 24 gecikmenin olduğu durum ortalama olarak daha başarılı olup sırasıyla ise 12 gecikme, 3 gecikme daha sonra ise 6 gecikme gelmektedir.

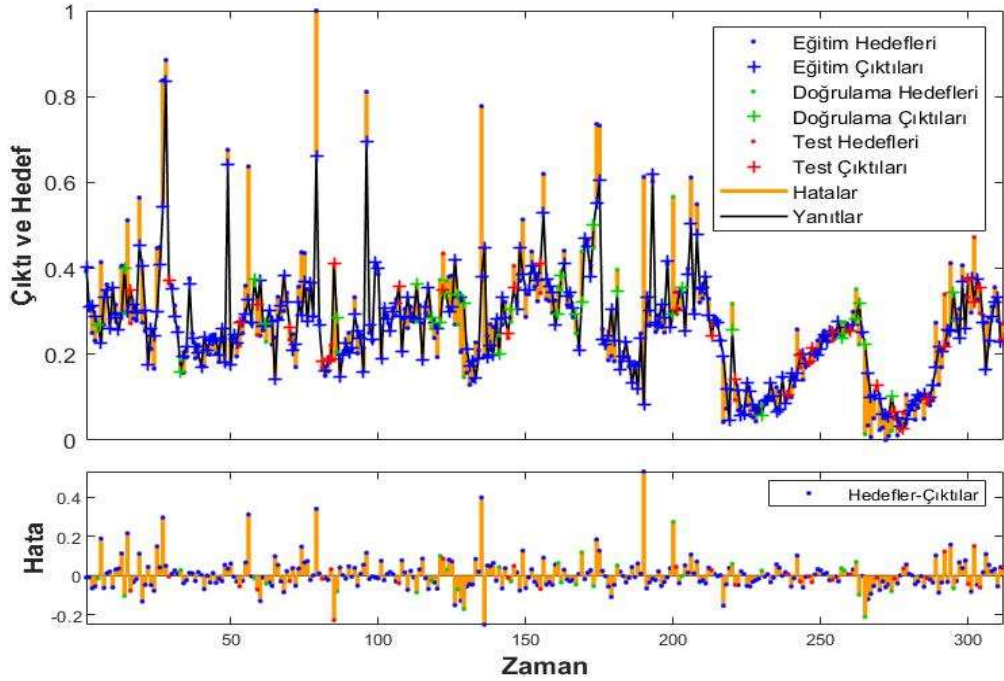
Çizelge 5.3. NAR ağına ait sgdm eğitim seçeneğinin sonuçları

Eğitim seçenekleri	NAR - SGDM											
	3			6			12			24		
Gecikme	%60	%70	%80	%60	%70	%80	%60	%70	%80	%60	%70	%80
Oran	%20	%15	%10	%20	%15	%10	%20	%15	%10	%20	%15	%10
	%20	%15	%10	%20	%15	%10	%20	%15	%10	%20	%15	%10
MAE/Test	0,1449	0,1586	0,1608	0,1428	0,1326	0,1323	0,1444	0,1527	0,1324	0,1436	0,1473	0,1326
MAE/Eğitim	0,1105	0,1104	0,1032	0,1076	0,1134	0,1113	0,1162	0,1146	0,122	0,121	0,1136	0,0986
MAE/Doğrulama	0,1361	0,1241	0,1287	0,096	0,1145	0,1359	0,1847	0,1911	0,1471	0,1045	0,107	0,1332
MAPE/Test	0,2928	0,4454	0,3804	0,2926	0,2911	0,2745	0,3824	0,3015	0,2937	0,2926	0,3822	0,2893
MAPE/Eğitim	0,0428	0,0477	0,0411	0,0462	0,0429	0,0477	0,0481	0,0354	0,0476	0,0482	0,0301	0,0294
MAPE/Doğrulama	0,309	0,4316	0,3696	0,0635	0,1172	0,2995	0,5753	0,6573	0,6262	0,07	0,0813	0,2983
MSE/Test	0,056	0,0358	0,0531	0,0589	0,0355	0,0331	0,0451	0,0655	0,0393	0,0451	0,0453	0,0334
MSE/Eğitim	0,0357	0,0367	0,0336	0,0451	0,0359	0,0363	0,0366	0,0367	0,0291	0,0484	0,0365	0,0211
MSE/Doğrulama	0,0411	0,0347	0,036	0,0305	0,0357	0,0346	0,0517	0,0545	0,0588	0,032	0,0331	0,0324
Test R ²	0,7542	0,7943	0,7841	0,7221	0,8239	0,8765	0,6664	0,6208	0,8392	0,794	0,7664	0,8008
Eğitim R ²	0,6998	0,7271	0,7244	0,6543	0,6927	0,7509	0,7413	0,7162	0,851	0,6842	0,7121	0,8572
Doğrulama R ²	0,6168	0,7343	0,8458	0,7	0,7383	0,6275	0,5451	0,5222	0,5515	0,7851	0,7848	0,6368
Tüm R ²	0,6897	0,7388	0,7495	0,6709	0,717	0,7548	0,6968	0,6798	0,81	0,7173	0,7303	0,8394

NAR-LM ağı için en iyi R^2 değerini alan 24 gecikme $\%(80,10,10)$ modeli için Şekil 5.14'te hataların otokorelasyonunu veren grafik verilmiştir. Grafiğe göre korelasyonların güven sınırları içinde olduğu ve sıfırın üzerinde tek sütun korelasyonunun olduğunu yani elde edilen tahmin verilerinin güvenilir olduğunu göstermektedir.

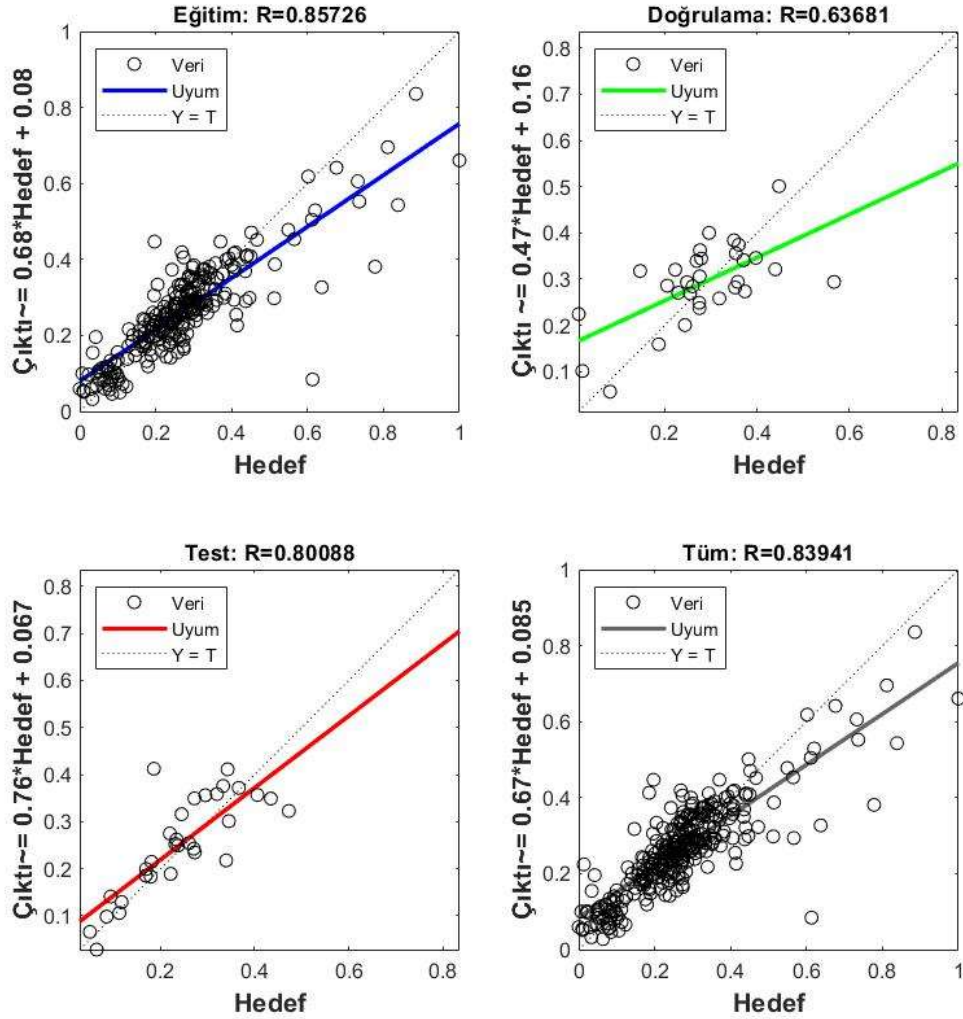


Şekil 5.14. NAR ağı hataların otokorelasyonunu gösteren grafik



Şekil 5.15. NAR ağına ait zaman serisi için çıktı elemanlarının yanıtlarını gösteren grafik

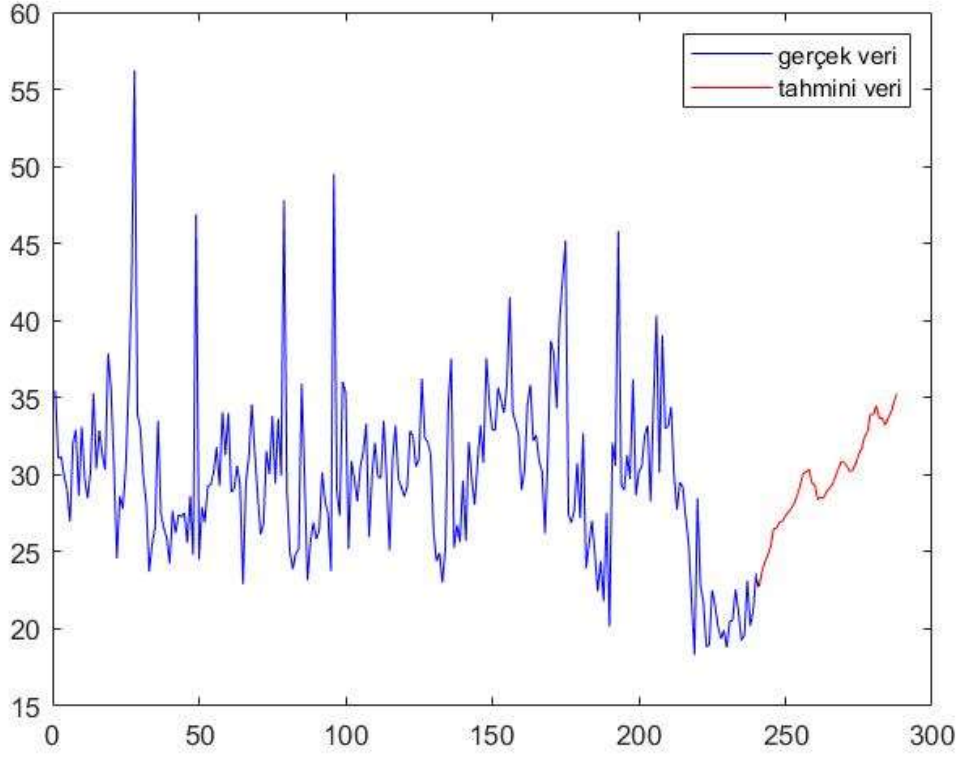
Şekil 5.15'te hedef ve çıktılarının uyumunu gösteren grafikte NARX ağına göre hataların daha belirgin olduğu görülmektedir. Hataların zamana göre durumunu gösteren grafikte ise bazı noktalarda yüksek olduğu görülebilir.



Şekil 5.16. NAR regresyon katsayısını gösteren grafikler

Matlab'ın verdiği regresyon grafikleri çıktılar ve hedefler arasındaki ilişkiyi temsil eder. Şekil 5.16'te ki grafiklere bakıldığında NAR ağına doğrulama hariç R değerlerinin düşük olmadığı söylenebilir. Doğrulama değeri düşük olsa da tüm R değerine bakıldığında çok başarısız bir sonuç olmadığı da söylenebilir. Tüm R değerleri arasında en düşük değer 24 gecikme ile % (60,20,20) oran ve rmsprop eğitim seçeneğine aittir.

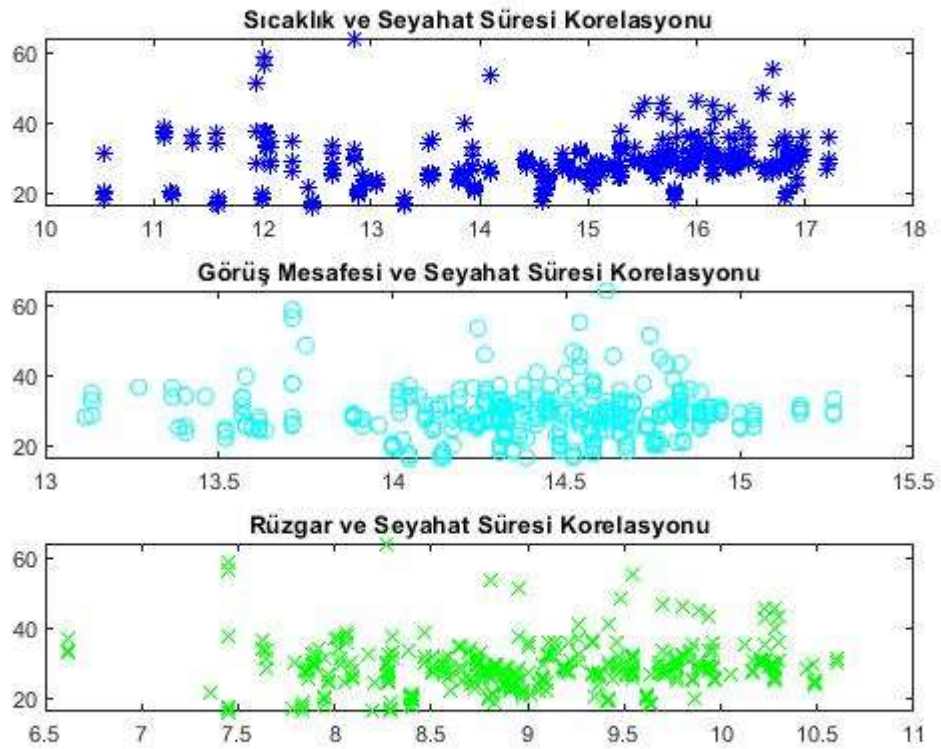
Şekil 5.17’de verilen grafikte tahmini değerler normalleştirilmenin kaldırılmış hali olarak verilmiştir. Grafikte mavi olanlar gerçek veriyi kırmızı olanlar ise tahmini veriyi göstermektedir.



Şekil 5.17. NAR ağı tahmini değerleri gösteren grafik

5.5.2. NARX-LM

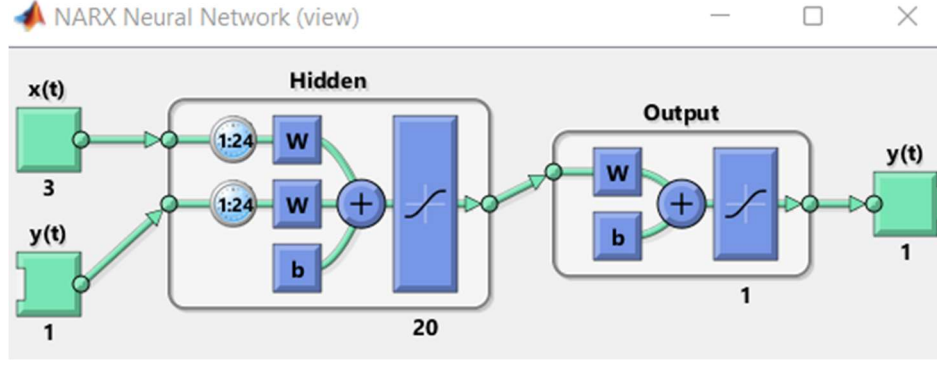
NARX sinir ağı dıřsal girdilerle birlikte tahmin yapan bir ađdır. Seyahat s¼resini etkileyen sıcaklık, g¼r¼ř mesafesi ve r¼zg¼r deęiřkenleri ile seyahat s¼resi tahmin alıřması yapılmıřtır. Analize bařlamadan ¼nce dıřsal verilerde min-max y¼ntemi ile normalizasyon yapılmıřtır. NARX iin de LM algoritması kullanılmıř ve daha ¼nce ara katman n¼ron sayısı bulunmuřtu. NARX iinde ara katman n¼ron sayısı 20 olarak alınmıřtır. Őekil 5.18 seyahat s¼resi ile sıcaklık, g¼r¼ř mesafesi ve r¼zg¼r verilerinin korelasyonunu g¼stermektedir.



Őekil 5.18. Sıcaklık, g¼r¼ř mesafesi ve r¼zg¼r deęiřkenlerinin seyahat s¼resi ile korelasyonu

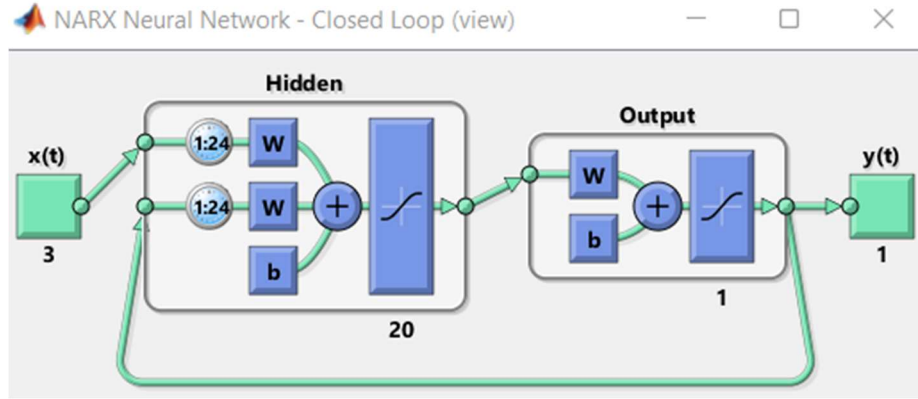
En iyi modeli bulabilmek iin yine

- Gecikme sayıları 3,6,12,24 olarak,
- Eęitim oranları % (Eęitim oranı 60, Doęrulama oranı 20, Test oranı 20), % (Eęitim oranı 70, Doęrulama oranı 15, Test oranı 15) ve % (Eęitim oranı 80, Doęrulama oranı 10, Test oranı 10) olarak,
- Eęitim seenekleri ‘adam’, ‘rmsprop’ ve ‘sgdm’ olarak her biri tek tek denendi.



Şekil 5.19. NARX YSA açık döngü mimarisi

Şekil 5.19 da Matlab'ın oluşturduğu açık döngü 'nün şemasını göstermektedir. Şemaya göre 3 dışsal (eksojen) girdi, ara katmanda 20 nöron, 24 gecikme olduğunu ve her iki katmanda da tansig fonksiyonunun kullanıldığını göstermektedir. Açık döngüde çıkış sistemin gerçek değerlerinden oluşmaktadır.



Şekil 5.20. NARX YSA kapalı döngü mimarisi

Şekil 5.20 de verilen kapalı döngüde ise açık döngüdeki gibi gecikmeleri, nöron sayısını ve kullanılan aktivasyon fonksiyonu görülmektedir. Kapalı döngü tahmini değerleri çıktı olarak vermektedir.

Çizelge 5.4'te NARX 'adam' eğitim seçeneğine ait çizelge verilmiştir. Çizelgeye göre en iyi model R^2 si 92.97, 24 gecikme ile eğitim oranı %60, doğrulama oranı %20 ve test oranı %20 olan modeldir. Eğitim, doğrulama ve test R^2 değerlerine bakıldığında değerlerin birbirine tutarlı olduğunu yani birinin çok yüksek diğerinin çok düşük çıkmadığı söylenebilir. Bu durumda aşırı öğrenme veya az öğrenme durumunun da olmadığı söylenebilir. MAE, MAPE ve MSE değerlerine bakıldığında da en küçük değer 24 gecikme, %(60,20,20) modelinde olmadığını ama bu modeldeki değerlerin de çok yüksek olmadığı söylenebilir.

Oranlara göre her bir eğitim seçeneğinin ne kadar başarılı olduğuna bakarsak da ‘adam’ eğitim seçeneği için % (60,20,20) ile % (70,15,15) ‘in ortalama değerlerinin birbirine çok yakın olduğunu % (80,10,10)’un ise ortalama olarak diğerlerinden daha yüksek olduğu söylenebilir.

Gecikme sayılarına göre bakıldığında ‘adam’ eğitim seçeneği için 24 gecikme sayısı ile ortalama olarak daha yüksek başarı elde edildiği sonra sırasıyla 12 gecikme, 6 gecikme ve 3 gecikmenin geldiği söylenebilir.

Çizelge 5.4. NARX ağına ait adam eğitim seçeneğinin sonuçları

Eğitim seçenekleri	NARX - ADAM											
	3			6			12			24		
Gecikme	%60	%70	%80	%60	%70	%80	%60	%70	%80	%60	%70	%80
	%20	%15	%10	%20	%15	%10	%20	%15	%10	%20	%15	%10
MAE/Test	0,2081	0,2081	0,3537	0,2177	0,2107	0,4901	0,393	0,3685	0,3576	0,4313	0,3543	0,3526
MAE/Eğitim	0,115	0,1149	0,1146	0,1268	0,115	0,1149	0,1268	0,1268	0,1205	0,1254	0,1356	0,1268
MAE/Doğrulama	0,1998	0,2028	0,2171	0,2088	0,2059	0,2188	0,2193	0,2098	0,2169	0,2137	0,2162	0,2165
MAPE/Test	1,4422	1,2519	1,4514	1,2567	1,2587	1,2397	1,2578	1,4827	1,4172	1,4912	1,396	1,4114
MAPE/Eğitim	0,6071	0,7373	0,6832	0,8484	0,7489	0,8188	0,8842	0,9386	0,8348	0,0986	0,9731	0,715
MAPE/Doğrulama	1,2116	1,2115	1,3787	1,2271	1,2272	1,4956	1,2117	1,3795	1,3733	1,3704	1,372	1,3717
MSE/Test	0,0842	0,0881	0,0799	0,0742	0,0883	0,0797	0,0869	0,0782	0,0719	0,0699	0,0698	0,0704
MSE/Eğitim	0,0842	0,0831	0,0707	0,0882	0,0843	0,0733	0,0889	0,08	0,0754	0,0413	0,0741	0,0698
MSE/Doğrulama	0,0865	0,0888	0,0779	0,0848	0,0873	0,0785	0,0868	0,0783	0,0716	0,0666	0,0784	0,0719
Test R ²	0,7315	0,6126	0,7951	0,8336	0,7263	0,8999	0,8365	0,9045	0,6871	0,9586	0,8205	0,8996
Eğitim R ²	0,7922	0,7505	0,8414	0,8416	0,7928	0,893	0,7495	0,8509	0,8408	0,9174	0,8687	0,8742
Doğrulama R ²	0,622	0,8257	0,6883	0,6481	0,8094	0,9052	0,7716	0,8807	0,827	0,9403	0,9569	0,8728
Tüm R ²	0,7451	0,7436	0,8241	0,807	0,7892	0,8953	0,7667	0,857	0,8252	0,9297	0,8793	0,8728

NARX ağına ait ‘rmsprop’ eğitim seçeneğine ait çizelgesi aşağıda verilmiştir. Çizelge 5.5’e göre en iyi model 24 gecikme ile eğitim oranı %80, doğrulama oranı %10 ve test oranı %10 olmaktadır. Eğitim, doğrulama ve test R² değerlerine bakıldığında değerlerin birbirine tutarlı olduğunu yani birinin çok yüksek değerinin çok düşük çıkmadığı söylenebilir. Bu durumda aşırı öğrenme veya az öğrenme durumunun da olmadığı söylenebilir. MAE, MAPE ve MSE değerlerine bakıldığında da en küçük değer 24 gecikme, % (80,10,10) modelinde olmadığını ama bu modeldeki değerlerin de çok yüksek olmadığı söylenebilir.

Oranlara göre ‘rmsprop’ için sırasıyla % (60,20,20), % (70,15,15) ve % (80,10,10)’ un daha başarılı sonuçlar gösterdiği görülüyor. Bu durumda oranlar arttıkça ağın başarı durumunun arttığı söylenebilir.

Gecikme sayılarına göre ‘rmsprop’ da ortalama en yüksek başarıyı 24 gecikme ile sağlamış ve sırasıyla 12 gecikme, 6 gecikme ve 3 gecikme ile devam etmektedir. Bu sonuçlardan gecikme sayısı arttıkça başarı oranının arttığı tam söylenemez.

Çizelge 5.5. NARX ağına ait rmsprop eğitim seçeneğinin sonuçları

Eğitim seçenekleri	NARX - RMSPROP											
	3			6			12			24		
Gecikme	%60	%70	%80	%60	%70	%80	%60	%70	%80	%60	%70	%80
	%20	%15	%10	%20	%15	%10	%20	%15	%10	%20	%15	%10
Oran	%20	%15	%10	%20	%15	%10	%20	%15	%10	%20	%15	%10
MAE/Test	0,2081	0,2081	0,2177	0,2176	0,2086	0,4938	0,2095	0,4943	0,2057	0,3695	0,2084	0,2044
MAE/Eğitim	0,115	0,1149	0,1149	0,1284	0,1269	0,1332	0,1366	0,1269	0,1271	0,1172	0,1174	0,1271
MAE/Doğrulama	0,1998	0,2021	0,2102	0,2086	0,2052	0,1995	0,2142	0,2163	0,1998	0,2099	0,2067	0,242
MAPE/Test	1,2521	1,3384	1,4575	1,3407	1,5194	1,5231	1,4251	1,5364	1,2483	1,3492	1,5358	1,6029
MAPE/Eğitim	0,8654	0,6974	0,6041	0,9553	0,0797	0,0797	0,0604	0,9598	0,8496	0,0772	0,0753	0,0818
MAPE/Doğrulama	1,2129	1,3166	1,3729	1,3387	1,4934	1,3796	1,3715	1,4018	1,5628	1,3481	1,4314	1,5724
MSE/Test	0,0949	0,0961	0,0855	0,0714	0,078	0,0696	0,0713	0,0686	0,0692	0,0776	0,07	0,0954
MSE/Eğitim	0,0346	0,0324	0,0515	0,0577	0,0259	0,0353	0,0234	0,0203	0,047	0,0479	0,0277	0,038
MSE/Doğrulama	0,0741	0,0792	0,0761	0,0815	0,0764	0,0724	0,0726	0,0793	0,0725	0,0781	0,0688	0,0952
Test R ²	0,7617	0,7296	0,7253	0,6629	0,8578	0,7111	0,7358	0,7325	0,9595	0,9459	0,9765	0,9734
Eğitim R ²	0,748	0,8001	0,7849	0,8083	0,8243	0,9114	0,8464	0,8656	0,8497	0,8097	0,9418	0,9467
Doğrulama R ²	0,6358	0,6646	0,6415	0,7194	0,88	0,9127	0,7944	0,8499	0,8523	0,8769	0,9492	0,9802
Tüm R ²	0,7253	0,7691	0,7708	0,7253	0,8352	0,8965	0,8148	0,846	0,8672	0,857	0,9473	0,9568

NARX ağına ait ‘sgdm’ eğitim seçeneğine ait çizelgesi aşağıda verilmiştir. Çizelge 5.6 ya göre en iyi model 24 gecikme ile eğitim oranı %80, doğrulama oranı %10 ve test oranı %10 olmaktadır. Eğitim R², doğrulama R² ve test R² değerlerine bakıldığında en tutarlı ve en yüksek değer bu modelde olduğu görülüyor. MAE, MAPE ve MSE değerleri de bu modelde en küçük değerleri almıştır. Bu modele ait tahmin grafiği Şekil 5.24 de verilmiştir.

Oranlara göre 'sgdm' için ise ortalama olarak sırasıyla % (60,20,20), % (70,15,15) ve % (80,10,10)' un daha başarılı sonuçlar gösterdiği görülüyor.

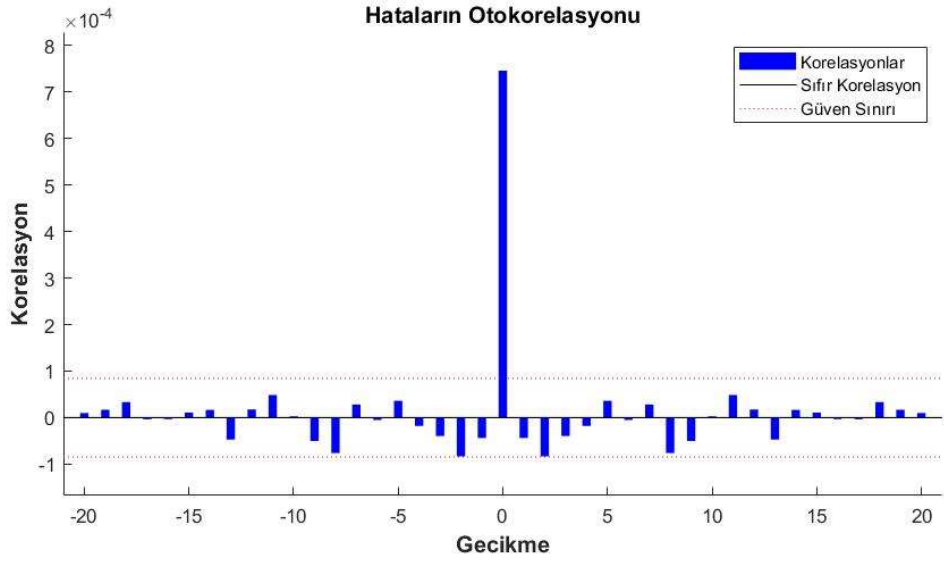
Gecikme sayılarına göre 'sgdm' için de 24 gecikmenin olduğu durum ortalama olarak daha başarılı olup sırasıyla ise 12 gecikme, 3 gecikme ve 6 gecikme gelmektedir.

Çizelge 5.6. NARX ağına ait sgdm eğitim seçeneğinin sonuçları

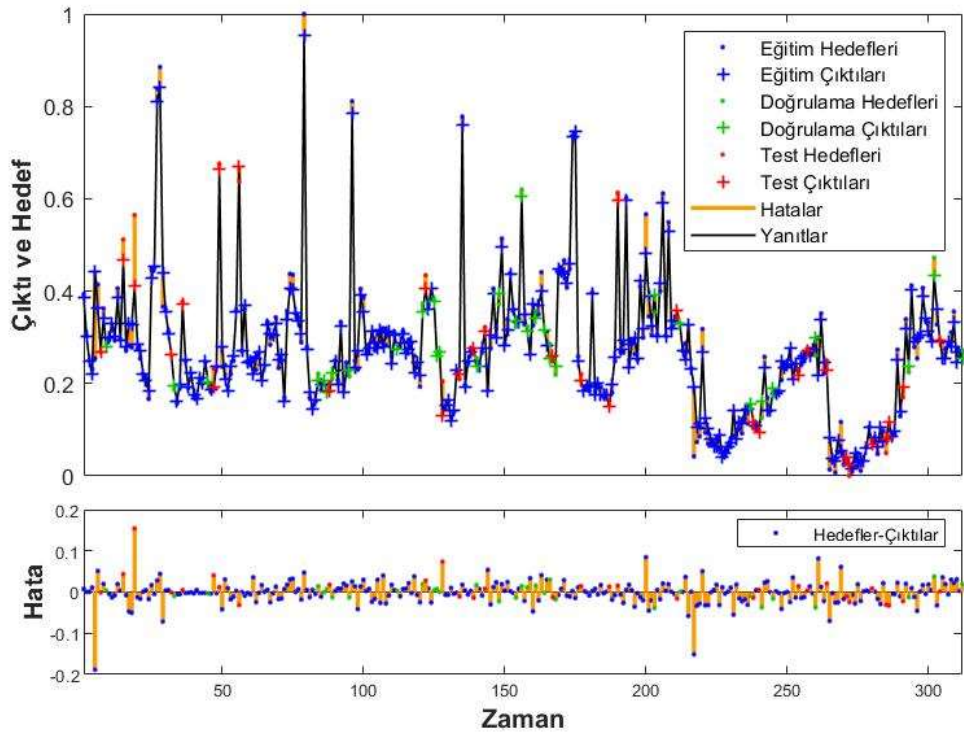
Eğitim seçenekleri	NARX - SGDM											
	3			6			12			24		
Oran	%60	%70	%80	%60	%70	%80	%60	%70	%80	%60	%70	%80
	%20	%15	%10	%20	%15	%10	%20	%15	%10	%20	%15	%10
MAE/Test	0,2081	0,2081	0,3539	0,2103	0,3292	0,3588	0,3679	0,2082	0,3541	0,3585	0,454	0,203
MAE/Eğitim	0,115	0,1149	0,1146	0,1231	0,115	0,1146	0,1464	0,1467	0,121	0,1468	0,1202	0,1145
MAE/Doğrulama	0,1999	0,2012	0,2164	0,2022	0,2089	0,2145	0,2095	0,2088	0,2163	0,217	0,2169	0,1994
MAPE/Test	1,2572	1,2499	1,5432	1,254	1,3769	1,4871	1,3093	1,5779	1,4258	1,3137	1,4446	1,2384
MAPE/Eğitim	0,7255	0,7326	0,6812	0,8466	0,8538	0,6059	0,915	0,0823	0,8096	0,0719	0,0869	0,0562
MAPE/Doğrulama	1,2106	1,2116	1,4576	1,2255	1,3843	1,3754	1,3	1,4624	1,3803	1,3886	1,3751	1,2046
MSE/Test	0,0866	0,0879	0,0847	0,0977	0,0841	0,0864	0,0694	0,0721	0,0759	0,071	0,0852	0,0682
MSE/Eğitim	0,0849	0,073	0,0707	0,0857	0,0848	0,0715	0,0651	0,0368	0,0547	0,0361	0,0293	0,0273
MSE/Doğrulama	0,0875	0,0886	0,0856	0,0966	0,0815	0,0853	0,0749	0,0721	0,0696	0,0824	0,076	0,0648
Test R ²	0,691	0,899	0,8505	0,7667	0,7513	0,6641	0,7664	0,8855	0,9405	0,7725	0,8925	0,9803
Eğitim R ²	0,7539	0,7933	0,7779	0,7316	0,752	0,7924	0,9112	0,9201	0,9061	0,9086	0,9334	0,9831
Doğrulama R ²	0,7087	0,8032	0,9434	0,6781	0,7731	0,8765	0,8889	0,9659	0,8919	0,8608	0,7993	0,9832
Tüm R ²	0,7359	0,8095	0,8064	0,7209	0,7525	0,7931	0,8912	0,9213	0,9066	0,8668	0,9104	0,9825

Şekil 5.21 de gösterilen hata otokorelasyon grafiği tahmin hataların zamanla ilişkisini göstermektedir. Korelasyonların güven sınırları içinde olması ve sıfırın üzerinde tek sütun korelasyonunun olması elde edilen tahmin verilerinin güvenilir olduğunu göstermektedir.

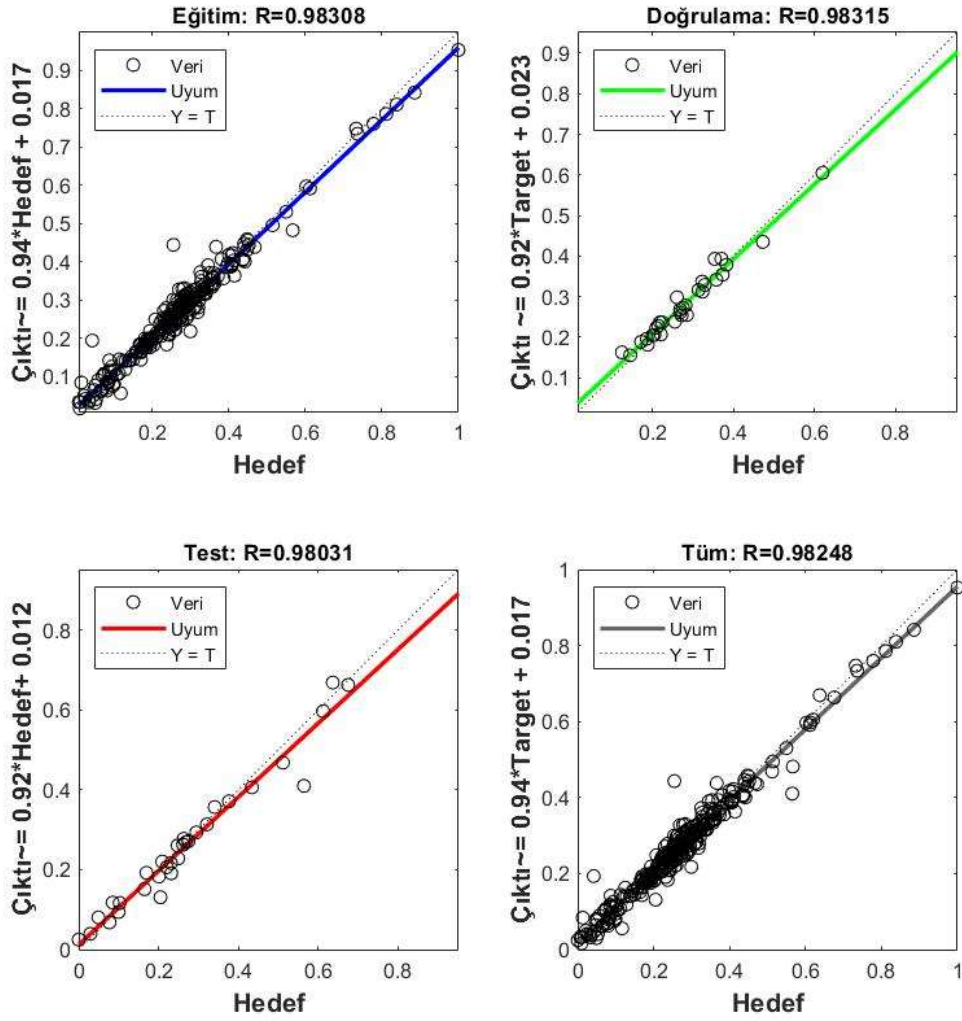
Şekil 5.22 de çıktı ve hedeflerin birbirleriyle uyumunu gösteren grafikte hataların çok olmadığını genelde uyumlu hareket ettikleri söylenebilir. Şeklin altındaki hata grafiğinde ise zamana bağlı hataların durumu görülmektedir. Burada da başlarda hata oranının yüksek olduğunu ama genel olarak çok hata içermediği söylenebilir.



Şekil 5.21. NARX ağı hataların otokorelasyonunu gösteren grafik



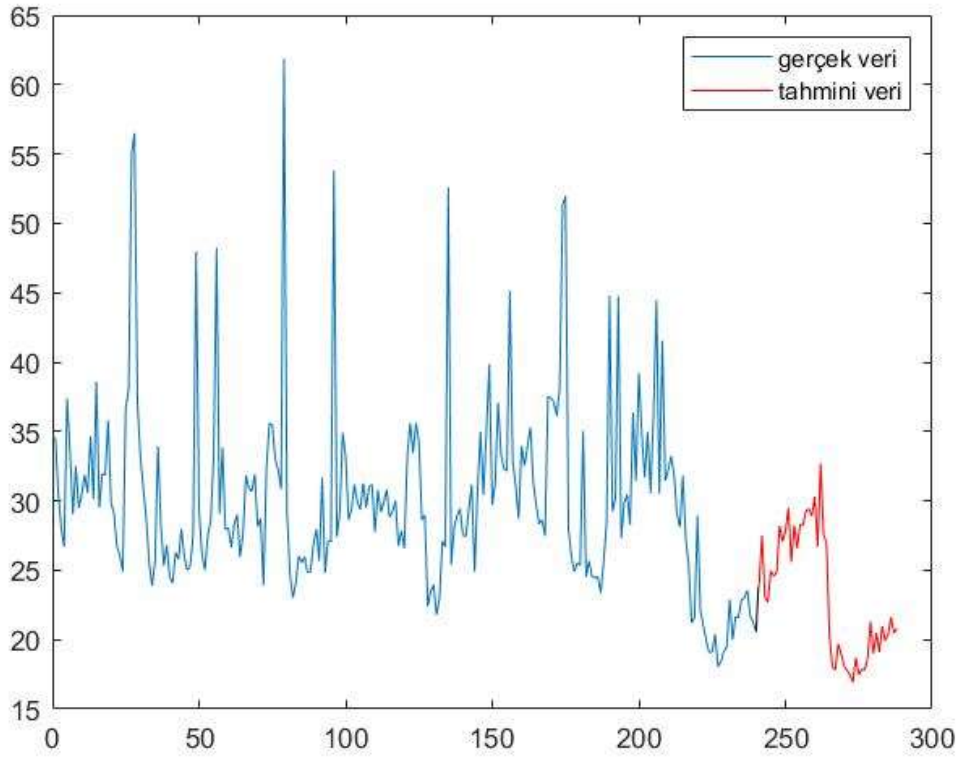
Şekil 5.22. NARX ağına ait zaman serisi için çıktı elemanlarının yanıtlarını gösteren grafik



Şekil 5.23. NARX regresyon katsayısını gösteren grafikler

Çalışan ağın performansının doğrulaması Şekil 5.23'te ki grafiklerden yararlanarak da yapılabilir. Bu grafiklere bakıldığında eğitim, doğrulama, test ve tüm R^2 değerlerinin 98'in üzerinde olduğu regresyon doğrusuna verilerin uyum gösterdiği aynı zamanda ağın başarılı bir performans gösterdiği söylenebilir.

Matlab da normalleştirme işlemi kaldırıldıktan sonra orijinal halleri ile gelen tahmini değerleri gösteren grafik Şekil 5.24'te gösterilmiştir.



Şekil 5.24. NARX ağı tahmini değerleri gösteren grafik

6. SONUÇLAR VE TARTIŞMA

New York şehrinin iki havalimanı arasında 15'er dakikalık toplanan seyahat süresi verilerini daha iyi anlayabilmek için önce hangi dağılıma uydukları bulunmuş daha sonra da güvenilirliğinin araştırılması için 95. yüzdellikleri, tampon süre indeksi, planlanan süre indeksi, başarısızlık ve zamanında oranı ile seyahat süresinin güvenilirliği hakkında bilgiler elde edilmiştir. Seyahat sürelerinin Log-Lojistik (3P) dağılımına uyduğu yani sağa çarpık dağılımlar gösterdiği belirlenmiştir.

Seyahat sürelerinin 95. yüzdelliklerine göre Salı, Çarşamba ve Cuma günlerinin genellikle sabah saatlerinin daha kötü seyahat süresi güvenilirliğine sahip olduğu söylenebilir.

Tampon süre indeksine göre ise Salı günü 9:45-10:00 saatleri arası ortalamaya eklenmesi gereken süre en çoktur. İkinci olarak Cuma günü 9:30-10:00 saatleri arası gelmektedir. Hafta sonu içinse Cumartesi günü 13:30-13:45 saatleri arası en fazla ortalamaya eklenmesi gereken süredir.

Planlanan süre indeksinin hafta içi en fazla olduğu gün ise Çarşamba günü 8:30-8:45 saatleri arası, hafta sonu Cumartesi günü 18:00-18:15 saatleri arasındadır.

Seyahat sürelerinin dağılımı ve güvenilirliği incelendikten sonra tahmini için bu çalışmada yapay sinir ağları yöntemi kullanılmıştır. Veriler zaman serisi verileri olduğu için yapay sinir ağlarının zaman serisini de içeren NAR ve NARX sinir ağları ile çalışma yapılmıştır.

NAR yöntemi ise dışsal girdi değişkenleri olmadan sadece seyahat süresi değişkeni ile 15'er dakikalık 48 adım tahmin etmiş ve sonuçlar tüm R^2 için %83 üzerinde bulunmuştur. Bu değer 24 gecikme, % (80,10,10) ve sgdm eğitim seçeneği modeli ile bulunmuştur. Ayrıca mse, mape ve mae değerlerine de bakıldığında en düşük değerlere sahip değildir ama en düşük değerler arasındadır. NAR için kurulan diğer modellerin çok tutarlı olmadığı bazı modeller için aşırı öğrenme durumunun olabileceği de görülmüştür. Eğitim seçenekleri açısından incelersek tezde çalışılan veri için sgdm ile rmsprop'un başarı durumu arasında çok az bir fark olduğu ve bu eğitim seçeneklerinin adam eğitim seçeneğinden daha iyi olduğu söylenebilir.

NARX yöntemi için seyahat süresini etkileyen diğer faktörlerden sıcaklık, görüş mesafesi ve rüzgar değişkenleri de eklenerek NARX ile seyahat süresi tahmin edilmiş ve eğitim, test, doğrulama ve tüm R^2 değerleri %98 üzerinde bulunmuştur. Ayrıca diğer performans

ölçütlerine de bakıldığında en iyi değerin 24 gecikme, % (80,10,10) ve sgdm eğitim seçeneği kullanılan model olduğu görülmüştür. Daha sonra yine 24 gecikme % (80, 10, 10) ve rmsprop eğitim seçeneği ile %95 R^2 ile 2. en iyi model olduğu görülmüştür. Diğer kurulan modellerin de genel olarak R^2 değerlerinin yüksek ve tutarlı olduğu görülmektedir. En düşük regresyon değerinin (tüm R^2 için) 6 gecikme, %60, %10, %10 oran ve sgdm eğitim seçeneği ile %72,09 olduğu söylenebilir. Bu değerin çok düşük bir değer olmadığı da söylenebilir. Eğitim seçenekleri açısından incelersek tezde çalışılan veri için sgdm 'nin daha başarılı olduğunu sonra sırasıyla ise rmsprop ve adam eğitim seçeneklerinin geldiği söylenebilir.

Tahmin için her iki yöntemi karşılaştırıldığında NARX yönteminin NAR yönteminden daha iyi sonuç verdiği söylenebilir. Bu duruma seyahat süresinin etkileyen dışsal faktörlerin seyahat süresini açıklamaya yardımcı olması da bir etkidir. Ayrıca karşılaştırılan eğitim seçeneklerinden ise çalışmada kullanılan veri için 'sgdm' nin daha başarılı sonuç verdiği söylenebilir. Oranlar içinde %80 eğitim, %10 test ve %10 doğrulama olarak alındığında daha başarılı sonuçlar elde edildiği görülmüştür.

Tezde seyahat sürelerinin normalliği sağlamamasından ve zaman serisi içerdiğinden dolayı yapay sinir ağları kullanılmıştır. Yapay sinir ağları öğrenme ve tahmin yapma konusunda başarılı olsa da her denemede farklı sonuçlar vermesi, ağı oluştururken belli kurallara sahip olmaması, deneme yanılma yollarını içermesi birer dezavantajdır.

KAYNAKLAR

- [1] FHWA Office of Operations, Does Travel Time Reliability Matter ?, <https://ops.fhwa.dot.gov/publications/fhwahop19062/whatis.htm> (Erişim tarihi: 30 Nisan 2022).
- [2] H.H. Hochmair (2016), Spatiotemporal Pattern Analysis of Taxi Trips in New York City, *Transportation Research Record*, (2016), 45-56.
- [3] https://www.fhwa.dot.gov/planning/glossary/glossary_listing.cfm?TitleStart=R (Erişim tarihi : 30 Nisan 2022).
- [4] A. Haghani, Y. Zhang, M. Hamed, Impact of Data Source on Travel Time Reliability Assessment, Final Project Report, Mid-Atlantic Universities Transportation Center, 1-44, (2014).
- [5] M.A.P. Taylor, Travel Time Variability-The Case of Two Public Modes, *Transportation Science*, 16 (1982) 507-521.
- [6] E. Mazloumi, G. Currie, G. Rose, Using GPS Data to Gain Insight into Public Transport Travel Time Variability, *J. Transp. Eng.* 136 (2010) 623–631.
- [7] L.M. Kieu, A. Bhaskar, E. Chung, Public Transport Travel-Time Variability Definitions and Monitoring, *J. Transp. Eng.* 141 (2015).
- [8] P. Chen, R. Tong, G. Lu, Y. Wang, Exploring Travel Time Distribution and Variability Patterns Using Probe Vehicle Data: Case Study in Beijing, *Journal of Advanced Transportation*. (2018) 1–13.
- [9] V.P. Singh, H. Guo, F.X. Yu, Parameter estimation for 3-parameter log-logistic distribution (LLD3) by Pome, *Stochastic Hydrol Hydraul.* 7 (1993) 163–177.
- [10] E. Durán-Hormazábal, A. Tirachini, Estimation of travel time variability for cars, buses, metro and door-to-door public transport trips in Santiago, Chile, *Research in Transportation Economics*. 59 (2016) 26–39.
- [11] Michael.A.P. Taylor, Susilawati, Modelling Travel Time Reliability with the Burr Distribution, *Procedia - Social and Behavioral Sciences*. 54 (2012) 75–83
- [12] M.A. Yazici, C. Kamga, K.C. Mouskos, Analysis of Travel Time Reliability in New York City Based on Day-of-Week and Time-of-Day Periods, *Transportation Research Record*. 2308 (2012) 83–95.
- [13] A. Higatani, T. Kitazawa, J. Tanabe, Y. Suga, R. Sekhar, Y. Asakura, Empirical Analysis of Travel Time Reliability Measures in Hanshin Expressway Network, *Journal of Intelligent Transportation Systems*. 13 (2009) 28–38.
- [14] W. Pu, Analytic Relationships between Travel Time Reliability Measures, *Transportation Research Record: Journal of the Transportation Research Board*. 2254 (2011) 122–130.

- [15] J.W.C. Van Lint, H.J. Van Zuylen, H. Tu, Travel time unreliability on freeways: Why measures based on variance tell only half the story, *Transportation Research Part A: Policy and Practice*. 42 (2008) 258–277.
- [16] M. Qiu, Smart Computing and Communication, Third International Conference, SmartCom 2018, Tokyo, Japan, December, (2018) 10-12.
- [17] Yangxin Lin, Meng Ma, Mengting Bai, Ping Wang, Travel-Time Prediction Methods: A Review, *LNISA*, Vo. 11344, (2018), 67-77.
- [18] E. Özgan, M. Orakoğlu, S. Serin, Ş. Sargin, Kavşaklarda Trafik Akımının Kuyruk Teorisi ile Tahmini, Örnek Bir Uygulama, (2018) 14.
- [19] B. Bustillos, Short Term Travel Time Prediction On Freeways In Conjunction With Detector Coverage Analysis, (2007) 139.
- [20] M. Taşyürek, M. Çelik, Akıllı Durak Sistemindeki Araç Seyahat Sürelerinin Birleşik Yapay Sinir Ağları Kullanarak Tahmini, *European Journal of Science and Technology*. (2020) 72–79.
- [21] N. Değirmenci, A. Akay, Finansal Verilerin ARIMA ve ARCH Modelleriyle Öngörüsü: Türkiye Örneği, *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi*. 12 (2017) 15–36.
- [22] O. Akyıldız, Mobil Araçlar İçin GPS Bilgisi Olmaksızın Öz Navigasyon Sistemi Geliştirilmesi, (2018) 76.
- [23] Ü. Yıldırım, Z. Çataltepe, Örüntü Tanıma ve Öznitelik Seçme Yöntemleri Kullanarak Kısa Zaman Sonraki Yol Trafik Hız Öngörüsü, (2010) 8.
- [24] İ.B. Coşkun, S. Sertok, B. Anbaroğlu, K-Nearest Neighbour Query Performance Analyses On A Large Scale Taxi Dataset: Postgresql Vs. MongoDB, *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. XLII-2/W13 (2019) 1531–1538.
- [25] Kolektif Öğrenme (Ensemble Learning) ve Bagging Algoritması <https://medium.com/@billmuhh/kolektif-%C3%B6%C4%9Frenme-ve-bagging-algoritmas%C4%B1-e8ea3d932b72> (Erişim tarihi: 30 Nisan 2022)
- [26] E. Yakut, B. Elmas, S. Yavuz, Yapay Sinir Ağları ve Destek Vektör Makineleri Yöntemleriyle Borsa Endeks Tahmini, 19 (2014) 139–157.
- [27] M.F. Keskenler, E.F. Keskenler, Geçmişten Günümüze Yapay Sinir Ağları ve Tarihçesi, 5 (2017).
- [28] Sinir Hücreleri (Nöron) Nedir? Yapısı, Özellikleri ve Çeşitleri Nelerdir?, <https://www.tech-worm.com/sinir-hucreleri-noron-nedir-yapisi-ozellikleri-ve-cesitleri-nelerdir/> (Erişim tarihi 30 Nisan 2022).
- [29] Yapay Sinir Ağları, <https://globalaihub.com/yapay-sinir-aglari/> (Erişim tarihi: 30 Nisan 2022).

- [30] B. Ataseven, Yapay Sinir Ağları İle Öngörü Modellemesi, Öneri Dergisi, 10(39) (2013)101-115.
- [31] F. Adıyaman, Talep Tahmininde Yapay Sinir Ağlarının Kullanılması, Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri, İstanbul, 2007
- [32] İ. Çayıroğlu, İleri Algoritma Analizi, <http://www.ibrahimcayiroglu.com/Dokumanlar/IleriAlgoritmaAnalizi/IleriAlgoritmaAnalizi-5.Hafta-YapaySinirAglari.pdf> (Erişim tarihi 30 Nisan 2022)
- [33] Activation Functions in Neural Networks, <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6> (Erişim tarihi 30 Nisan 2022)
- [34] Y.E. Kuyucu, Lojistik Regresyon Analizi (LRA), Yapay Sinir Ağları (YSA) Ve Sınıflandırma Ve Regresyon Ağaçları (C&RT) Yöntemlerinin Karşılaştırılması ve Tıp Alanında Bir Uygulama, (2012).
- [35] D. Tezer, Yapay Sinir Ağları, Lojistik Regresyon ve Destek Vektör Makinesi İstatistik Yöntemlerin Sınıflandırmada Karşılaştırılması, (2018).
- [36] D. Anderson, G. McNeill, Artificial Neural Networks Technology, DACS report, New York, (1992).
- [37] E. Öztemel, Yapay sinir ağları, 3. Baskı, Papatya Yayıncılık, (2012).
- [38] H.H. Nguyen, A Neural Fuzzy Approach to Modeling the Thermal Behavior of Power Transformers,(2007).
- [39] Radial Basis Functions Neural Networks, <https://towardsdatascience.com/radial-basis-functions-neural-networks-all-we-need-to-know-9a88cc053448>, (Erişim tarihi 30 Nisan 2022)
- [40] A. Özcan, Homojen Ve Heterojen Evrimsel Sosyal Ağlarda Bağlantı Tahmini, Doktora Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri, İstanbul, (2017).
- [41] G. İntepe, Teknoloji Tahmini İçin Yeni Bir Model Önerisi, Doktora Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri, İstanbul, (2016).
- [42] G. Zhang, B. E. Patuwo, M. Hu, Forecasting with artificial neural networks: The State of the Art, International Journal of Forecasting,(1998)
- [43] S. Yavuz, M. Deveci, İstatistiksel Normalizasyon Tekniklerinin Yapay Sinir Ağın Performansına Etkisi, Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi. (2012) 22.
- [44] T. Jayalakshmi, A. Santhakumaran, Statistical Normalization and Back Propagation for Classification, IJCTE. (2011) 89–93.
- [45] D. Aşkın, İ. İskender, A. Mamızadeh, Farklı Yapay Sinir Ağları Yöntemlerini Kullanarak Kuru Tip Transformator Sargısının Termal Analizi, (2011) 10.

- [46] K. Levenberg, A method for the solution of certain non-linear problems in least squares, *Quart. Appl. Math.* 2 (1944) 164–168
- [47] Yu H, Wilamowski BM ,Levenberg–Marquardt training. In: *Intelligent systems*, vol 12, (2011) 1–16.
- [48] R. Mokhtari, S. Fakouriyan, R. Ghasempour, Investigating the Effect of Cloud Cover on Radiative Cooling Potential With Artificial Neural Network Modeling, *Front. Energy Res.* 9 (2021).
- [49] R. Sarkar, S. Julai, S. Hossain, W.T. Chong, M. Rahman, A Comparative Study of Activation Functions of NAR and NARX Neural Network for Long-Term Wind Speed Forecasting in Malaysia, *Mathematical Problems in Engineering* (2019) 1–14.
- [50] S. Karasu, A. Altan, Z. Saraç, R. Hacıoğlu, B.E. Üniversitesi, Rüzgâr Hızının Doğrusal Olmayan Otoregresif Sinir Ağları ile Tahmini Prediction of Wind Speed with Non-Linear Autoregressive (NAR) Neural Networks, (2017) 4.
- [51] A. Sarı, M. Aktaş, A. Yönetken, R. Doğan, Güneş Işınım Tahmininde NARX Modelinin Uygulanması, 4 (2021) 1-6.
- [52] F. Di Nunno, F. Granata, Groundwater level prediction in Apulia region (Southern Italy) using NARX neural network, *Environmental Research.* 190 (2020).
- [53] Y. Firat, Application of the Air Quality Forecasting Analysis Using NARX Models, *Gümüşhane Üniversitesi Fen Bilimleri Enstitüsü Dergisi*,10(2) (2020) 442-455.
- [54] E. Demirci, A. Baykaldı, M. Karaatlı, Ticari Kredi Faiz Oranlarının YSA NARX ve VAR Modelleri ile Öngörülmesi, *İşletme Araştırmaları Dergisi*, 12 (2020) 2327–2343.
- [55] M.E. Çimen, S. Kaçar, E. Güteryüz, B. Gürevin, A. Akgül, Kaotik bir hareket videosunun yapay sinir ağları ile modellenmesi, *Balıkesir Üniversitesi Fen Bilimleri Enstitüsü Dergisi.* (2018) 23–35.
- [56] A.S. Mane, S.S. Pulugurtha, Link-level Travel Time Prediction Using Artificial Neural Network Models, in: *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, IEEE, Maui, HI, (2018) 1487–1492.
- [57] S. Sertok, Ö. Yeniay, B. Anbaroğlu, Seyahat Süresi Değişkenliğinin ve Güvenilirliğinin İstatistiksel Analizi, *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi.* (2020) 445–454.
- [58] TLC Trip Record Data, <https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page%20> (Erişim tarihi: 1 Mayıs 2022)
- [59] Weather in Ocak 2015 in New York, <https://www.timeanddate.com/weather/usa/newyork/historic?month=1&year=2015> (Erişim tarihi: 1 Mayıs 2022)

[60] E. Demir, Ö. Saatçiođlu, F. İmrol, Uluslararası Dergilerde Yayımlanan Eğitim Arařtırmalarının Normallik Varsayımları Açısından İncelenmesi, (2016) 20.

EKLER

EK1: Uyum İyiliği Testleri

Koyu mavi zaman dilimleri hariç diğer zaman dilimleri için uyum iyiliği testlerinden Kolmogorov-Smirnov testi uygulanmıştır. Koyu mavi zaman dilimleri için ise diğer bir uyum iyiliği testi olan Anderson-Darling testi uygulanmıştır.

	Pazartesi	Salı	Çarsamba	Persembe	Cuma	Cumartesi	Pazar
7:00-	63	49	63	62	79	63	77
7:15-	88	51	63	75	75	62	84
7:30-	85	68	71	83	76	83	78
7:45-	80	42	51	70	72	85	65
8:00-	80	75	58	66	80	73	68
8:15-	71	58	73	60	89	74	72
8:30-	65	41	65	84	88	57	59
8:45-	71	45	50	46	67	47	68
9:00-	47	49	39	61	69	43	57
9:15-	67	45	63	50	63	46	58
9:30-	82	39	44	65	77	56	60
9:45-	74	46	46	51	60	45	57
10:00-	69	68	57	45	68	50	50
10:15-	64	52	51	41	57	45	40
10:30-	63	52	58	49	48	56	50
10:45-	64	54	56	45	66	48	54
11:00-	62	52	53	65	64	59	62
11:15-	70	60	73	62	86	64	55
11:30-	75	84	78	74	111	64	64
11:45-	88	56	80	74	72	50	68
12:00-	79	58	61	58	97	66	54
12:15-	69	53	55	64	75	47	68
12:30-	69	65	74	59	77	53	92
12:45-	82	83	74	55	78	52	80
13:00-	93	72	76	94	83	49	90
13:15-	116	93	88	97	109	89	117
13:30-	95	93	96	95	113	75	103
13:45-	105	91	92	88	87	74	105
14:00-	113	88	104	90	105	82	117
14:15-	117	90	98	104	114	78	101
14:30-	116	112	102	115	115	86	131
14:45-	114	99	91	107	117	85	120
15:00-	103	111	103	98	107	77	121
15:15-	126	89	87	121	106	68	138
15:30-	105	101	92	108	106	69	105
15:45-	100	116	97	99	117	68	112
16:00-	109	97	67	97	89	53	120
16:15-	108	99	85	108	112	51	99
16:30-	108	100	89	102	115	53	113
16:45-	107	108	105	87	97	51	123
17:00-	114	130	107	90	115	56	127
17:15-	95	118	106	107	134	50	135
17:30-	90	89	110	83	112	52	124
17:45-	96	89	82	63	119	28	127
18:00-	77	86	93	92	93	22	86
18:15-	83	72	87	78	92	22	94
18:30-	68	78	73	56	63	27	95
18:45-	70	95	66	57	71	26	67

