

**T.C.**  
**ULUDAĞ ÜNİVERSİTESİ**  
**SOSYAL BİLİMLER ENSTİTÜSÜ**  
**EKONOMETRİ ANABİLİM DALI**  
**İSTATİSTİK BİLİM DALI**

**MAKROEKONOMİK ZAMAN SERİSİ ANALİZİ**  
**VE YAPAY SİNİR AĞI UYGULAMALARI**

**(DOKTORA TEZİ)**

**Özer ARABACI**

**BURSA, 2007**

**T.C.**  
**ULUDAĞ ÜNİVERSİTESİ**  
**SOSYAL BİLİMLER ENSTİTÜSÜ**  
**EKONOMETRİ ANABİLİM DALI**  
**İSTATİSTİK BİLİM DALI**

**MAKROEKONOMİK ZAMAN SERİSİ ANALİZİ**  
**VE YAPAY SİNİR AĞI UYGULAMALARI**

**(DOKTORA TEZİ)**

**Özer ARABACI**

**Danışman:**

**Prof. Dr. Mustafa AYTAÇ**

**BURSA, 2007**

T. C.

**ULUDAĞ ÜNİVERSİTESİ**

**SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜNE**

**Ekonometri** Anabilim/~~Anasana~~ Dalı, **İstatistik** Bilim Dalı'nda **U2004085** numaralı **Özer ARABACI**'nin hazırladığı “**MAKROEKONOMİK ZAMAN SERİSİ ANALİZİ VE YAPAY SİNİR AĞI UYGULAMALARI**” konulu **Doktora Tezi Çalışması** ile ilgili tez savunma sınavı, **27/11/ 2007** günü **12:00- 14:00** saatleri arasında yapılmış, sorulan sorulara alınan cevaplar sonunda adayın tezinin/~~çalışmasının~~ **BAŞARILI** (başarılı/~~başarısız~~) olduğuna **OYBİRLİĞİ** (oybirliği/~~oy çokluğu~~) ile karar verilmiştir.

Sınav Komisyonu Başkanı  
Prof. Dr. Ercan DÜLGEROĞLU  
Uludağ Üniversitesi

Üye (Tez Danışmanı)  
Prof. Dr. Mustafa AYTAÇ  
Uludağ Üniversitesi

Üye  
Prof. Dr. Nemci GÜRSAKAL  
Uludağ Üniversitesi

Üye  
Prof. Dr. Ercan DÜLGEROĞLU  
Uludağ Üniversitesi

Üye  
Prof. Dr. Erkan IŞIĞIÇOK  
Uludağ Üniversitesi

Üye  
Prof. Dr. Şehamet BÜLBÜL  
Marmara Üniversitesi

Ana Bilim Dalı Başkanı  
Prof. Dr. Nemci GÜRSAKAL

27/11/ 2007

Enstitü Müdürü  
Prof. Dr. Mustafa AYTAÇ

## ÖZET

Yazar : Özer Arabacı  
Üniversite : Uludağ Üniversitesi  
Anabilim Dalı : Ekonometri  
Bilim Dalı : İstatistik  
Tezin Niteliği : Doktora Tezi  
Sayfa Sayısı : XII + 197  
Mezuniyet Tarihi : .... /.... / ...  
Tez : Prof. Dr. Mustafa AYTAÇ  
Danışman(lar)ı

### **Makroekonomik Zaman Serisi Analizi ve Yapay Sinir Ağı Uygulamaları**

Son yıllarda zaman serisi analizinde yapay sinir ağı modellerinin kullanımına ilgi artmaktadır. Doğrusal dışı modelleme başarısı, veri setinden öğrenebilme yeteneği ve veri yaratma sürecine kısıt getirmemesi gibi özellikler yapay sinir ağı modellerini çekici kılmaktadır. Diğer taraftan farklı bir terminolojiye sahip olması ve parametrik olmayan doğası ise bir dezavantaj olarak görülmektedir. Bu çalışmada iktisadi zaman serilerinin sahip olduğu farklı anahtar özellikler durumunda yapay sinir ağı modellerinin kullanımının uygunluğu araştırılmaktadır. Bu amaçla mevsimsellik, yapısal kırılma, volatilité ve doğrusal dışılık gibi özelliklere sahip farklı zaman serileri kullanılmıştır. İlk olarak bu seriler sahip oldukları bu anahtar özellikler durumunda kullanılan geleneksel modellerle modellenmiştir. İkinci olarak ilgili seriler yapay sinir ağı modelleriyle modellenmiş ve önraporlama performansları karşılaştırılmıştır. Elde edilen sonuçlar volatilité dışında diğer durumlarda yapay sinir ağı modellerinin kullanılabilceğini destekler yöndedir.

**Anahtar Sözcükler:** Zaman Serileri, Yapay Sinir Ağı, İleri Beslemeli Ağ, ARIMA Modelleri, Önraporlama

## ABSTRACT

Yazar : Özer Arabacı  
Üniversite : Uludağ Üniversitesi  
Anabilim Dalı : Ekonometri  
Bilim Dalı : İstatistik  
Tezin Niteliği : Doktora Tezi  
Sayfa Sayısı : XII + 197  
Mezuniyet Tarihi : .... /.... / ...  
Tez : Prof. Dr. Mustafa AYTAÇ  
Danışman(lar)ı

### **Macroeconomic Time Series Analysis and Artificial Neural Networks Applications**

There is an increasing interest in time series analysis using artificial neural networks models (ANN) in recent years. ANN models are attractive because of their nonlinear modelling success, learning from data capability. On the other hand it is a disadvantage that it has a different nomenclature and nonparametric nature. In this study it was investigated that in which cases it is useful to use ANN models in economic time series analysis. For this purpose, some economic time series which have different key features such as seasonality, structural break, volatility and nonlinearity are used. Firstly these series are modelled by using traditional methods. Secondly same series are modelled by using ANN models and forecasting performance of those models are compared. Results are supported that except for volatility, ANN models can be used as a time series analysis method.

**Key Words:** Time Series, Artificial Neural Networks, Feedforward Networks, ARIMA Models, Forecasting.

## *İÇİNDEKİLER*

<i>TEZ ONAY SAYFASI</i>	i
<i>ÖZET</i>	ii
<i>ABSTRACT</i>	iii
<i>İÇİNDEKİLER</i>	iv
<i>TABLolar</i>	v
<i>ŞEKİLLER</i>	vii
<i>GİRİŞ</i>	1

### **BİRİNCİ BÖLÜM**

#### **ZAMAN SERİSİ ANALİZİNDE ANA BAŞLIKLAR**

<b>I. ANAHTAR ÖZELLİKLER</b>	<b>4</b>
A. DURAĞANLIK KAVRAMI	4
B. TREND	6
1. Gecikme Yapısının Belirlenmesi	10
2. Dickey-Fuller ve Genişletilmiş Dickey-Fuller Testleri	11
C. MEVSİMSELLİK	13
1. Periyodik Otoregresif Model	14
2. Periyodik Otoregresif Modelde Birim Kök Testi	16
D. YAPISAL KIRILMA	17
E. VOLATİLİTE MODELLEMESİ	21
1. ARCH	22
2. GARCH	23
F. DOĞRUSAL DIŞILIK	25
1. Eşik Değer Otoregresif Model (TAR)	28
2. Düzgün Geçiş Modelleri (STAR)	32
<b>II. ÖNRAPORLAMA</b>	<b>36</b>

**İKİNCİ BÖLÜM**  
**YAPAY SİNİR AĞLARINA GİRİŞ VE TEMEL MODELLER**

<b>I. YAPAY SİNİR AĞLARINDA TARİHSEL GELİŞİM</b>	<b>44</b>
<b>II. YAPAY SİNİR AĞLARINDA TERMİNOLOJİ VE TEMEL MODELLER</b>	<b>48</b>
A. YAPAY SİNİR AĞLARINDA TEMEL YAPI VE ELEMANLAR	48
B. McCULLOCH – PITTS MODELİ	52
C. BASİT ALGILAYICI MODELİ	54
D. ADAPTİF DOĞRUSAL ELEMAN AĞI	64

**ÜÇÜNCÜ BÖLÜM**  
**YAPAY SİNİR AĞI MODELLERİNİN**  
**ZAMAN SERİSİ ANALİZİNDE KULLANIMI**

<b>I. ÖĞRETMELİ ÖĞRENME STRATEJİSİ</b>	<b>74</b>
<b>II. ÇOK KATMANLI ALGILAYICILAR</b>	<b>80</b>
<b>III.TAHMİN</b>	<b>88</b>
A. GERİ YAYILIM ALGORİTMASI	89
B. DOĞRUSAL OLMAYAN EN KÜÇÜK KARELER	96
<b>IV.AĞ SEÇİMİ</b>	<b>100</b>
A. TEORİK BİLGİ KRİTERLERİ	105
1. Schwarz Bilgi Kriteri	105
2. Akaike Bilgi Kriteri	106
B. DİĞER YAKLAŞIMLAR	107
1. Çapraz Geçerlilik	107
2. Erken Durdurma Metodu	108
3. Hata Kriterlerinin Kullanımı	109
<b>V. GERİ DÖNÜŞÜMLÜ AĞLAR</b>	<b>110</b>
<b>VI. YAPAY SİNİR AĞLARI İLE ZAMAN SERİSİ ANALİZİNE YENİDEN BAKIŞ</b>	<b>112</b>

## DÖRDÜNCÜ BÖLÜM

# TÜRKİYE EKONOMİSİNE AİT DEĞİŞİK ZAMAN SERİSİ ÖZELİKLERİ GÖSTEREN MAKRO EKONOMİK SERİLER ÜZERİNDE YAPAY SİNİR AĞI MODELLERİNİN PERFORMANSLARININ ÖLÇÜLMESİ

I. YÖNTEM	121
II. MEVSİMSELLİK İÇEREN BİR SERİ OLARAK ÖZEL NİHAİ TÜKETİM SERİSİNİN MODELLENMESİ	123
III. İMKB-100 HAFTALIK GETİRİLERİ İÇİN VOLATİLİTE MODELLEMESİ	135
IV. YAPISAL KIRILMA GÖSTEREN BİR SERİ OLARAK SANAYİ ÜRETİM ENDEKSİNİN MODELLENMESİ	148
V. KAPASİTE KULLANIM ORANLARI SERİSİNDE DOĞRUSAL DIŞILIĞIN MODELLENMESİ	158
VI. AĞIRLIKLANDIRILMIŞ PENCERE YÖNTEMİYLE HAM PETROL VARİL FİYATLARININ ÖN RAPORLANMASI	168
SONUÇ	175
TERİMLER SÖZLÜĞÜ	183
ÖZGEÇMİŞ	198



## TABLolar

<i>Tablo 1: Önraporlama Performansı Ölçütleri</i>	39
<i>Tablo 2: “VE”, “VEYA”, “VE DEĞİL” Problemleri</i>	55
<i>Tablo 2: XOR Problemi</i>	56
<i>Tablo 3: “VE” Problemi</i>	59
<i>Tablo 4: “VEYA” Problemi</i>	67
<i>Tablo 6: Uygun PEAR (p) Modelinin Seçimi</i>	126
<i>Tablo 7: PEAR (1) Modeli Örneklem içi Önraporlama Ölçütleri</i>	128
<i>Tablo 8: NONT Serisi İçin Ağ Seçimi</i>	130
<i>Tablo 9: 5-4-1 FNN Modeli Örneklem içi Önraporlama Ölçütleri</i>	132
<i>Tablo 10: Özel Nihai Tüketim İçin Gerçekleşen Değerler ve Expost Önraporlama Performansları</i>	133
<i>Tablo 11: İMKB 100 Serisi Getirileri (R) İçin ADF Testi</i>	136
<i>Tablo 12: İMKB 100 Serisi Getirileri İçin En İyi Ortalama Modeli</i>	137
<i>Tablo 13: İMKB 100 Serisi Getirileri İçin ARCH LM Testi</i>	138
<i>Tablo 14: İMKB 100 Haftalık Getiri Serisi İçin Model Tahminleri</i>	139
<i>Tablo 15: İMKB 100 Haftalık Getiri Serisi İçin 3-4-1 FNN Önraporlama Performans Ölçütleri</i>	146
<i>Tablo 16: SÜE Serisi İçin ADF Testi</i>	149
<i>Tablo 17: SÜE Serisi İçin Peron 1997 testi</i>	150
<i>Tablo 18: <math>\Delta SÜE</math> Serisi İçin Model Tahminleri</i>	152
<i>Tablo 19: <math>\Delta SUE</math> Serisi İçin Ağ Seçimi</i>	154
<i>Tablo 20: <math>\Delta SUE</math> İçin 4-6-1 FNN Önraporlama Performans Ölçütleri</i>	155
<i>Tablo 21: <math>\Delta SUE</math> İçin Gerçekleşen Değerler ve Expost Önraporlama Performansları</i>	156
<i>Tablo 22: KKO Serisi İçin ADF testi</i>	159

<i>Tablo 23: Uygun ARMA(p,q) Modelinin Seçimi</i>	159
<i>Tablo 24: ARMA(3,2) Modeli</i>	160
<i>Tablo 25: KKO İçin 2-3-1 FNN Önraporlama Performans Ölçütleri</i>	162
<i>Tablo 26: Eşik Değer Modeli İçin En Uygun Modelin Belirlenmesi</i>	163
<i>Tablo 27: KKO İçin TAR Modeli Önraporlama Performans Ölçütleri</i>	164
<i>Tablo 28: LSTAR Model Spesifikasyonunun Belirlenmesi</i>	165
<i>Tablo 29: KKO İçin LSTAR Önraporlama Performans Ölçütleri</i>	166
<i>Tablo 30: KKO İçin Ex-post Önraporlamalar</i>	167
<i>Tablo 31: COP için ADF Testi</i>	169
<i>Tablo 32: 2-3-1 İleri Beslemeli Ağı için Bazı Değerlendirme Ölçütleri</i>	172
<i>Tablo 33: Çekirdek COP için ADF testi</i>	172
<i>Tablo 34: AR(2) Modeli için Bazı Değerlendirme Ölçütleri</i>	173
<i>Tablo 35: COP Serisi için Gerçekleşen Değerler ve Expost Önraporlamalar</i>	174

## ŞEKİLLER

<i>Şekil 1: Deterministik ve Stokastik Trende Sahip İki Seri</i>	9
<i>Şekil 2: Yapısal Kırılmalı Bir Seri</i>	18
<i>Şekil 3: Önraporlama ile İlgili Ana Kavramlar</i>	37
<i>Şekil 4: Biyolojik Bir Sinir Hücresinin Yapısı</i>	49
<i>Şekil 5: Yapay Bir Sinir Hücresinin Yapısı</i>	50
<i>Şekil 6: McCulloch – Pitts Yapay Sinir Hücresi</i>	52
<i>Şekil 7: (a) Ve – (b) Veya – (c) Ve değil Problemleri</i>	55
<i>Şekil 8: XOR Problemi</i>	56
<i>Şekil 9: Basit Algılayıcı Modeli</i>	57
<i>Şekil 10: Teorik Bir Sınıflandırma Problemi Çözümü</i>	58
<i>Şekil 11: “VE” Problemi</i>	60
<i>Şekil 12: “VE” Probleminin Basit Algılayıcı İle Çözümü</i>	63
<i>Şekil 13: ADALİNE Ünitesi</i>	64
<i>Şekil 14: “VEYA” Problemi</i>	67
<i>Şekil 15: “VEYA” Probleminin ADALİNE İle Çözümü</i>	71
<i>Şekil 16: İki Adaline Ünitesinden Oluşan Madaline Ağ</i>	72
<i>Şekil 17: Üç Katmanlı Algılayıcı</i>	81
<i>Şekil 18: İki Katmanlı İki Gizli Üniteli İleri Beslemeli Ağ</i>	82
<i>Şekil 19: Sıklıkla Kullanılan Transfer Fonksiyonları</i>	83
<i>Şekil 20: Problem Karmaşıklığı</i>	85
<i>Şekil 21: 3-2-1 İleri Beslemeli Ağ</i>	86
<i>Şekil 22: 3-2-1 İleri Beslemeli Ağ (Fonksiyon Gösterimi)</i>	87
<i>Şekil 23: Belirli Bir Aralıkta Fonksiyon Kompleksliği</i>	102

<i>Şekil 24: 1-3-1 Ağ ile Çeşitli Fonksiyonlara Yakınsama</i>	103
<i>Şekil 25: 1-5 -1 Ağı ile <math>h = \cos(9x+2)</math> Fonksiyonuna Yakınsama</i>	103
<i>Şekil 26: Genelleştirme Problemi</i>	104
<i>Şekil 27: Erken Durdurma Metodu</i>	108
<i>Şekil 28: Kısmi Geri Dönüşümlü Ağ</i>	110
<i>Şekil 29: Kayan Pencere Yöntemi</i>	116
<i>Şekil 30: Özel Nihai Tüketim (1987Q1 - 2005:Q4)</i>	124
<i>Şekil 31: Özel Nihai Tüketim: Çeyrekler Bazında Ortalamalar</i>	124
<i>Şekil 32: Özel Nihai Tüketim: Kısıtlı Model Kalıntı Kareler</i>	125
<i>Şekil 33: Özel Nihai Tüketim: Kısıtsız Model Kalıntı Kareler</i>	126
<i>Şekil 34: Özel Nihai Tüketim – PEAR(1) Modeli Önraporlama Performansı</i>	127
<i>Şekil 35: NONT Serisi</i>	129
<i>Şekil 36: Özel Nihai Tüketim 5-4-1 FNN Modeli Önraporlama Performansı</i>	131
<i>Şekil 37: Özel Nihai Tüketim İçin Expost Önraporlama Performansları</i>	133
<i>Şekil 38 : İMKB 100 Haftalık Getiri (20.07.1987-09.04.2007)</i>	136
<i>Şekil 39: İMKB 100 Haftalık Getirileri İçin Ortalama Modeli Önraporlaması</i>	140
<i>Şekil 40: İMKB 100 Haftalık Getirileri İçin Varyans Modeli Önraporlaması</i>	140
<i>Şekil 41: 2-3-1 FNN Cross Validation Eğitim Performansı</i>	142
<i>Şekil 42: 3-4-1 FNN Cross Validation Eğitim Performansı</i>	143
<i>Şekil 43: 2-5-1 FNN Cross Validation Eğitim Performansı</i>	144
<i>Şekil 44: 3-5-1 FNN Cross Validation Eğitim Performansı</i>	145
<i>Şekil 45: 3-4-1 FNN Ağı ile İMKB-100 Haftalık Getirilerinin Önraporlaması</i>	146
<i>Şekil 46 : Sanayi Üretim Endeksi (1980Q1-2006Q3)</i>	149
<i>Şekil 47 : Sanayi Üretim Endeksi Birinci Farkı (1980Q2-2006Q3)</i>	151
<i>Şekil 48: ARMA (3,4) Örneklem İçi Önraporlama Değerleri</i>	153
<i>Şekil 49 : Ağın Eğitiminde İterasyonlar Boyunca MSE değişimi</i>	154

<i>Şekil 50 : 4-6-1 FNN Örneklem İçi Önraporlama Değerleri</i>	<i>155</i>
<i>Şekil 51: <math>\Delta</math>SUE Serisi İçin Gerçekleşen Değerler ve Expost Önraporlamalar</i>	<i>156</i>
<i>Şekil 52: Kapasite Kullanım Oranları (02.1991-12.2006)</i>	<i>158</i>
<i>Şekil 53: ARMA(3,2) Modeli Örneklem İçi Önraporlaması</i>	<i>160</i>
<i>Şekil 54: 2-3-1 FNN İçin Cross Validation Yönteminde Erken Durdurma</i>	<i>161</i>
<i>Şekil 55: 2-3-1 FNN ile KKO serisi Örneklem İçi Önraporlama</i>	<i>162</i>
<i>Şekil 56: TAR ile KKO Serisi Örneklem İçi Önraporlama</i>	<i>164</i>
<i>Şekil 57: LSTAR ile KKO Serisi Örneklem İçi Önraporlama</i>	<i>166</i>
<i>Şekil 58: Ham Petrol Varil Fiyatları Ocak 1991-Mayıs 2007</i>	<i>168</i>
<i>Şekil 59: Ham Petrol Varil Fiyatları Kasım 2001-Mayıs 2007</i>	<i>170</i>
<i>Şekil 60: 2-3-1 FNN ile COP Serisi Örneklem İçi Önraporlama</i>	<i>171</i>
<i>Şekil 61: AR(2) Modeli ile COP Serisi Örneklem İçi Önraporlama</i>	<i>173</i>
<i>Şekil 62: COP Serisi Gerçekleşen Değerler ve EXPOST Önraporlamalar</i>	<i>174</i>

## ***GİRİŞ***

Modern hayatın gösterdiği gelişim ve bu gelişimin hızı karşısında her alanı daha karmaşık hale gelen günlük yaşam, geleceğe bakışın daha güvenilir olmasını gerektirmektedir. Bu nedenle geleceğe ait planların ve bu planlar için gerekli tahminlerin daha tutarlı bir şekilde yapılması son derece önemlidir. Bu önemin bir parçası olarak, iktisadi serilerin önraporlamalarındaki tutarlılık iktisadi birimlerin politika hedeflerini oluşturmalarında hayati bir bileşendir. Bu anlamda son 10 yılda popüleritesi her geçen gün artmakta olan Yapay Sinir Ağı Modellerinden yeni ve kullanışlı bir modelleme aracı olarak sıklıkla söz edilmektedir. Yüksek örüntü tanıma kabiliyeti, evrensel yakınsayıcı özelliği ve bu nedenle de her tür karmaşık ilişkiyi yakalayabilme yeteneği, regresyon yaklaşımının tersine veri setinin dağılımına ve tahmin edilecek parametrelerin doğrusallığına ilişkin herhangi bir varsayım gerektirmemesi gibi çekici özelliklerinin yanında, kendine özgü farklı bir terminolojisinin olması, tüm sürecin bilgisayarlarda işlediğine dair kara kutu yakıştırması, parametrelerinin yorumlanamaması gibi dezavantajları da olan bu modelleme ailesinin daha yakından incelenmesi gerekmektedir.

Bu tez çalışmasında “zaman serisi analizinde yapay sinir ağları yararlı bir analiz aracı olarak kullanılabilir mi?” sorusunun cevaplanması temel problematik olarak ele alınmıştır. Bu nedenle çalışmanın bakış açısı zaman serilerinin sahip olduğu bazı anahtar özellikler üzerinden şekillenmiştir. Bu anahtar özellikler çalışmada trend, mevsimsellik, volatilité, yapısal kırılma ve doğrusal dışılık olarak ele alınmıştır. Böylece ilk bölümde bu özellikler, bu özelliklere sahip zaman serilerinin tanımlanması için gerekli testler, kullanılacak zaman serisi modelleri ve önraporlama üzerinde durulmuştur.

İkinci bölümde yapay sinir ağı modellerinin kendine özgü terminolojisi, tarihsel gelişimi ve temel modellerin tanıtılması hedeflenmiştir. Böylece zaman serisi analizinde kullanılan yapay sinir ağı modellerinin incelenmesi için bir

alt yapı hazırlanmıştır. Özellikle tek bir nöronun ve buna bağlı olarak da temel modellerin nasıl çalıştığı üçüncü bölümde ele alınan daha karmaşık ağ yapılarının anlaşılması için önem arz etmektedir.

Üçüncü bölümde, zaman serisi analizinde kullanılan yapay sinir ağı modelleri öğretmenli öğrenme stratejisini kullandıkları için, istatistiksel öğrenme teorisi kısaca tanıtılmış, daha sonra da Çok Katmanlı Algılayıcılar ve Geri Dönüşümlü Ağlar tartışılmıştır. Bu tartışmada tahmin, katman sayısı ve katmanlarda yer alacak nöron sayısı problemleri ön planda tutulmuştur. Bu bölümde son olarak, yapay sinir ağı modellerinin zaman serisi analizinde kullanımını için adimsal bir süreç sunulmuştur.

Son olarak dördüncü bölüm olan uygulama bölümünde, Türkiye ekonomisine ait ve yukarıda sözü edilen anahtar özelliklere sahip seriler üzerinde yapay sinir ağı modellerinin performansları birinci bölümde tartışılan zaman serisi modellerinin performanslarıyla karşılaştırılmıştır. Böylece, zaman serisi analizinde yapay sinir ağı modellerini kullanmanın avantajlı olduğu durumların belirlenmesi hedeflenmiştir.

## BİRİNCİ BÖLÜM

### ZAMAN SERİSİ ANALİZİNDE ANA BAŞLIKLAR

Makroekonomik veriler eşit aralıklı zaman noktalarında ölçülmüş istatistiklerden oluşmaktadır. Bu istatistikler genellikle haftalık, aylık, üç aylık, yıllık şekilde derlenirler. Bu nedenle makroekonomik veriler üzerinden yapılacak analizler çoğunlukla zaman serisi analizi yaklaşımı çerçevesinde şekillenir. Ele alınan bir zaman serisinin diğer serilerden en önemli farkı, içerdiği gözlem değerlerinin bir sıraya sahip olmasıdır. Verilerin bir zaman sırasına göre ele alınması, arka planda bir veri yaratma süreci mantığına işaret eder. Bu nedenle ele aldığımız serinin stokastik bir süreç olduğu düşünülür. İşte bu durumda tek değişkenli zaman serisi yaklaşımı açısından problem, serinin durağan bir stokastik süreçle ifade edilip edilemeyeceğidir. Eğer seri durağan bir stokastik süreçle ifade edilemez ise değil karmaşık istatistiksel modelleme tekniklerin kullanımı, basit moment tahminlerinin yapılması bile sakıncalıdır. Böylece durağanlık testlerine olan ilgi başlamıştır. Aslında son 15 - 20 yılda tek değişkenli zaman serisi analizi başlığı altında çığ gibi büyüyen bu literatürün test etmeye çalıştığı, zaman serilerinin hemen hepsinin sahip olduğu anahtar bazı özelliklerin, ilgili serinin durağan bir stokastik süreçle ifade edilebilmesini etkileyip etkilemediğidir. Bu anahtar özellikler Trend, Mevsimsellik, Kırılma, Volatilité ve Doğrusal dışılıktır. Sonuçta bu bölümde ele alınacak başlıklar bu anahtar özelliklerden oluşacaktır. Öte yandan yukarıda da belirtildiği gibi, durağanlığın test edilmesi ile ilgili literatür geniş bir başlık olduğundan, konu burada sadece uygulama bölümünde Türkiye ekonomisine ilişkin kullanılacak serilerin gösterdikleri özelliklerle sınırlandırılmıştır.



## I. ANAHTAR ÖZELLİKLER

Yukarıda belirtildiği gibi zaman serileri analizinin tarihsel hikayesi zaman serilerinin gösterdiği ana özelliklerin modelleme konusunda ortaya çıkardıkları problemlerin çözümlenmesi paralelinde şekillenmiştir. Bu nedenle de tek bir model çerçevesinde konuyu ele alabilme yaklaşımı yetersiz kalmaktadır. Bu başlık altında ele alınacak konular, zaman serilerinin gösterdikleri farklı özelliklerin uzun uzun tartışılmasından ziyade, söz konusu özelliklerin farklı modellerle nasıl ele alınabileceğinin üzerinde durmaktadır. Çünkü doğru modelin seçimi ve buna bağlı olarak doğru öngörülerin elde edilmesi, bu tezin asıl problemiği olan “Yapay sinir ağlarının ekonomik zaman serilerinde bir analiz aracı olarak yararlı bir şekilde kullanılabilir mi?” sorusunun cevabının alınmasında son derece önemli bir yer teşkil etmektedir.

### A. DURAĞANLIK KAVRAMI

Durağanlık varsayımı, ilgili serinin olasılık yoğunluk fonksiyonunun zaman içerisinde değişmediğinin varsayımıdır. Serinin olasılık yapısı zaman içerisinde sabit kalıyorsa seri durağandır (*Pindyck ve Rubinfeld, 1981, s.497*). Durağanlık varsayımının kullanımı bazen değişik şekillerde karşımıza çıkabilmektedir. Çünkü pratikte yukarıdaki varsayım çok kısıtlayıcı olabilmekte ve test edilememektedir. Bu nedenle bir serinin ilk iki momentinin yani ortalamasının [ $E(y_t) = \mu_y$  tüm t için], ve varyansının [ $Var(y_t) = \sigma_y^2$  tüm t için], zaman boyunca değişmemesi ve  $y_t$ 'nin kendi tüm gecikmeli değerleri ile olan kovaryansının zaman boyunca sabit olması [ $Cov(y_t, y_{t+k}) = \gamma_k$ ] koşullarının yerine gelmesi *zayıf durağanlık* olarak adlandırılır. Ayrıca bu duruma kovaryans durağanlık yada ikinci dereceden durağanlık da denilmektedir. Durağanlıkla ilgili bir başka kavram *trend durağanlıktır*. Durağan olan bir zaman serisi  $y_t$ 'nin, sahip olduğu deterministik trendin etkisiyle ortalamasının değişmesi ile seri durağan

dışı bir yapı gösterebilir. Bu durumda şöyle bir model,

$$y_t = \mu + \phi t + \varepsilon_t$$

ileri sürülebilir. Eğer deterministik trend serinin durağan olmasını engelliyorsa, bu modelle trend etrafında  $\mu$  ortalamalı bir durağan stokastik süreç  $\varepsilon_t$  tarafından açıklanıyor olacaktır. Sıklıkla kullanılan bir diğer kavram *fark durağanlıktır*. Eğer seri stokastik bir trende sahipse yada bir başka ifade ile rassal yürüyüş olarak tanımlanıyorsa bunun anlamı serinin belirli bir ortalama etrafında saçılmadığı ve sabit bir varyansa sahip olmadığıdır. Bu durumda stokastik trend yapısının dışlanmasıyla seri durağanlaştırılabilir. Bu işlem fark alma işlemidir. Rassal yürüyüş serisi,

$$y_t = y_{t-1} + \varepsilon_t \text{ ise,}$$

her iki tarafın birinci farkı alındığında,

$$y_t - y_{t-1} = \varepsilon_t$$

$$\Delta y_t = \varepsilon_t \text{ olur.}$$

Böylece,  $\varepsilon_t \sim iid (0, \sigma_\varepsilon^2)$  olduğundan fark serisi durağan süreçle aynı yapıda olur.

Kısaca, fark alınarak durağanlaşan seriler fark durağan seri olarak adlandırılırken, deterministik trendden sapmaların durağan olduğu seriler ise trend durağan seriler olarak adlandırılır (*Chatfield, 1997, 462*). Son olarak durağan süreçler I(0) yani sıfırıncı dereceden entegre, bir fark alındığında durağan olan süreçler I(1) yani birinci dereceden entegre ve iki fark alındığında durağanlaşan süreçler ise I(2) yani ikinci dereceden entegre süreçler olarak adlandırılırlar.

## B. TREND

Ekonomik zaman serileri için en sık rastlanan ortak özellik pozitif bir trende sahip olmaktır. Elbette bunun en büyük nedeni, iktisadi büyüklüklerin nüfus artışı, teknolojik gelişmeler vb. nedenlerden dolayı artma eğiliminde olmasıdır.  $y_t$  serisi birinci dereceden otoregresif bir yapıya sahip olsun, yada diğer bir ifade ile AR(1) süreci ile temsil edilsin, bu durumda,

$$y_t = \phi_1 y_{t-1} + \varepsilon_t \quad \text{burada, } \varepsilon_t \sim (0, \sigma_\varepsilon^2) \text{ iid dir.}$$

Ekonomik zaman serileri için genel olarak  $\phi_1$  katsayısının alabileceği değerler  $0 \leq \phi_1 \leq 1$  arasındadır. Çünkü  $\phi_1 < 0$  durumunda seri sönen bir seriye dönecektir. Öte yandan  $\phi_1 > 1$  durumunda ise seri limitsiz olarak büyür (*Patterson, 2000, s.209*).  $0 \leq \phi_1 \leq 1$  şeklindeki genel duruma dönülecek olursa burada gözlemlenebilecek iki özel değer söz konusudur. Bunlar  $\phi_1 = 0$  ve  $\phi_1 = 1$  durumlarıdır.

$\phi_1 = 0$  durumunda,  $y_t = \varepsilon_t$  olacaktır.  $\varepsilon_t \sim (0, \sigma^2)$  iid olduğundan,  $y_t \sim (0, \sigma^2)$  iid olmasını gerektireceğinden  $y_t$  serisi beyaz gürültü olarak adlandırılır (*Dickey, 2005, s.2*).

$\phi_1 = 1$  olduğu durumda,  $y_t = y_{t-1} + \varepsilon_t$  olacaktır. Bu model ise pür rassal yürüyüş olarak adlandırılır. Bu modelde herhangi bir dönemdeki rassal şok, zaman içerisinde kaybolmadan diğer dönemleri de etkiler. Bu durum şu şekilde gösterilebilir.

$$y_t = y_{t-1} + \varepsilon_t \text{ ise,}$$

$$y_1 = y_0 + \varepsilon_1$$

$$y_2 = y_1 + \varepsilon_2 \text{ olduğundan,}$$

$$y_2 = y_0 + \varepsilon_1 + \varepsilon_2$$

$$y_3 = y_2 + \varepsilon_3$$

olacağından,

$$y_3 = y_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3$$

ve genel bir ifade olarak,

$$y_t = y_0 + \sum_{i=1}^t \varepsilon_i$$

yazılabilir.

Özet olarak,  $y_t$  serisinin zaman içerisindeki değerleri, başlangıç değeri olan  $y_0$  ve rassal terimin o ana kadarki tüm değerlerinin toplamından oluşmaktadır. Bu durumda rassal yürüyüş modeli sonsuz bir belleğe sahip olup, rassal şokların etkisi altında ve ne yönde hareket edeceği önceden tahmin edilemeyen bir yapıdır (*Patterson, 2000, s.210*). Bu yapıya bir ekleme  $\phi_1 = 1$  olan AR(1) yapısına sabit bir terim ( $\mu$ ) konularak yapılabilir. Böylece,

$$y_t = \mu + y_{t-1} + \varepsilon_t$$

olacaktır.

Burada,  $\mu$  *yığılma* terimi adını alır. Yığılma teriminin burada önemi  $\varepsilon_t$ 'deki rassal gerçekleşme ne olursa olsun  $\mu$ 'den dolayı  $y_t$  serisinin yukarı hareket edeceğidir. Şöyle ki,

$$y_1 = \mu + y_0 + \varepsilon_1$$

$$y_2 = 2\mu + y_0 + \varepsilon_1 + \varepsilon_2$$

$$y_3 = 3\mu + y_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3$$

ve genel bir ifade kullanarak,

$$y_t = y_0 + t\mu + \sum_{i=1}^t \varepsilon_i$$

bu modele kayan rassal yürüyüş modeli adı verilip, birinci farkı alındığında,

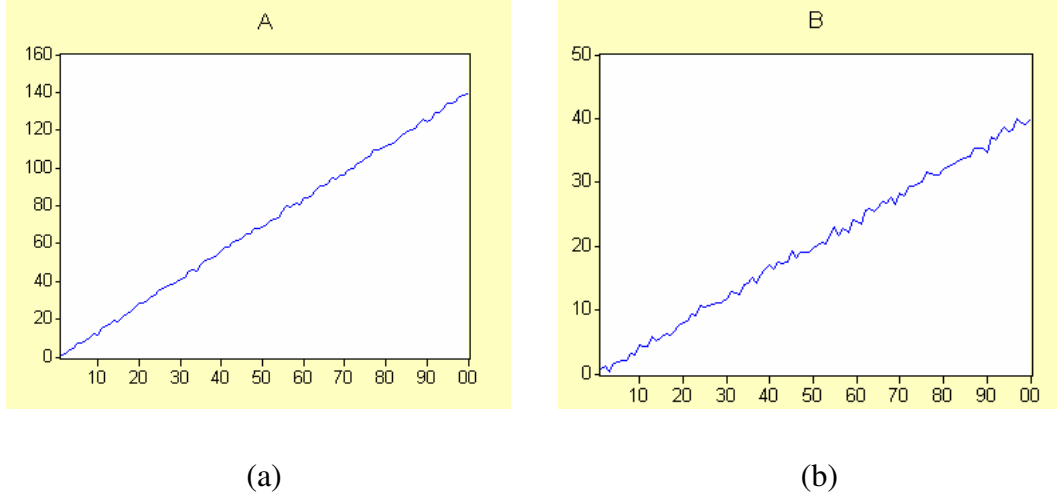
$$y_t - y_{t-1} = \mu + \varepsilon_t$$

yada,

$$\Delta y_t = \mu + \varepsilon_t$$

elde edilir.

Bu modelin özelliği,  $\varepsilon_t \sim (0, \sigma_\varepsilon^2)$  iid olduğundan, birinci fark serisi  $\mu$  ortalamaya sahip olan, pür rassal yürüyüş modeli olmasıdır (*McCabe / Tremayne, 1995, s.1016*). Zaman serisi analizi açısından ise bir örnek yardımıyla şu açıklamanın getirilmesi yerinde olacaktır.  $A_t = 0,4t + \varepsilon_t$  şeklindeki deterministik trende sahip olan seri ile,  $B_t = 0,4 + B_{t-1} + \varepsilon_t$  şeklinde stokastik trende sahip olan serinin aralarındaki fark bir takım testler yapılmadan anlaşılabilir. Ancak burada  $A_t$  serisi deterministik trend dışında belirli bir ortalama etrafında dalgalanırken,  $B_t$  serisi belirli bir ortalamaya dönme kuvvetinden yoksun bir seridir (*Franses, 1998, s.71*). Çünkü,  $A_t$  serisi deterministik trend ve durağan bileşenden oluşmuşken,  $B_t$  serisi stokastik trend ve durağan bileşenden oluşmuştur (*Hatanaka, 1996, s.16*). Aşağıda şekil 1 (a) ve (b)'de sırasıyla  $A_t$  ve  $B_t$  serileri görülmektedir. İki seri birbirlerinden açık bir fark ortaya koymamaktadır.



**Şekil 1: Deterministik ve Stokastik Trende Sahip İki Seri**

Böylece buraya kadar sıralanan özelliklere sahip genel bir zaman serisi modeli,  $y_t = \mu + \phi_1 y_{t-1} + \beta t + \varepsilon_t$  şeklinde ifade edilebilir. Buradan hareketle aşağıdaki 5 ayrı model söz konusudur (*Patterson, 2000, s.226*).

Model	Parametreler	Açıklama	Özellik
1	$\mu \neq 0,  \phi_1  < 1, \beta \neq 0$	Deterministik trend ve durağan AR(1) bileşeni	I(0)
2	$\mu \neq 0,  \phi_1  = 1, \beta \neq 0$	Kayan Rassal yürüyüş ve deterministik trend	I(1)
3	$\mu \neq 0,  \phi_1  = 1, \beta = 0$	Kayan rassal yürüyüş	I(1)
4	$\mu \neq 0,  \phi_1  = 0, \beta \neq 0$	Deterministik trend	I(0)
5	$\mu = 0,  \phi_1  = 1, \beta = 0$	Pür rassal yürüyüş	I(1)

## 1. Gecikme Yapısının Belirlenmesi

Yukarıda ifade edilen modeller sadece birinci dereceden bir AR(1) yapısını açıklamak için kullanılmaktadır. Aslında daha gerçekçi bir yaklaşım ek bazı otoregresif ve hareketli ortalama bağımlılığın seride yer alabileceğidir (*Stock / Watson, 1986, s.148*). Çünkü p. dereceden bir otoregresif yapı düşünülecek olursa,

$$Y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

bu durumda AR(1) süreci ile ele alınan süreçte

$$Y_t = \phi_1 y_{t-1} + \varepsilon_t$$

hata terimi, temiz bir seri değil, aşağıda gösterildiği gibi serisel olarak korelasyonlu olur

$$\varepsilon_t = \phi_2 y_{t-2} + \phi_3 y_{t-3} + \dots + \phi_p y_{t-p} + v_t$$

Bu nedenle kalıntılardaki serisel korelasyonun ortadan kaldırılması gereklidir. Bu amaçla kullanılacak başlıca iki temel yaklaşım bulunmaktadır. Bunlardan birincisi teorik bilgi kriterlerinden yararlanmak, ikicisi ise genelden özele ya da özelden genele modelleme stratejisini kullanmaktır.

- **Teorik Bilgi Kriterlerinin Kullanımı** ile gecikme yapısının derecesi belirlenirken cimrilik prensibi dikkate alınarak, bir ceza fonksiyonu yardımıyla fonksiyonel gecikmelerin sayısı mümkün olduğunca minimize edilir. Pratik olarak, daha sonra sunulacak olan Akaike (AIC) ve Schwarz (SIC) bilgi kriterlerinden elde edilen değerlerin marjinal anlamlılık düzeyleri kullanılarak kalıntılarda serisel korelasyonun olup olmadığı test edilebilir (*Gonzalo / Pitaraki, 2005, s.401*).

- **Genelden Özele ve Özelden Genele Modelleme Stratejilerinin** kullanımında ise sırasıyla genelden özele stratejisinde, birisi m gecikmeli ve diğeri m+n gecikmeli iki modelden son n gecikmenin bileşik olarak sıfıra eşit olduğu boş hipotezi Wald testi yardımıyla sınanır. Özelden genele stratejisinde ise, gecikme uzunluğunun belirlenmesinde n=1 alınarak ve birer birer artırılarak t testi yardımıyla modele dahil edilen son gecikmenin anlamlılığı test edilir (*Sevüktekin ve Nargeleçekenler, 2007, s.325*). Diğer taraftan genelden özele yaklaşımı, özelden genele yaklaşımına göre daha çok tercih edilir.

## 2. Dickey-Fuller ve Genişletilmiş Dickey-Fuller Testleri

Said ve Dickey (1984), daha önce Dickey ve Fuller (1979) tarafından ileri sürülen ve sadece AR(1) yapısında birim kökü test eden testi (DF testi) daha yüksek dereceden AR ve MA yapılarında da çalışabilecek şekilde genişletmişlerdir (ADF testi).

DF testi için,

$$y_t = \phi_1 y_{t-1} + \varepsilon_t$$

şeklindeki AR (1) modelini ele alalım. Eşitliğin her iki tarafının da farkı alındığında model,

$$y_t - y_{t-1} = \phi_1 y_{t-1} - y_{t-1} + \varepsilon_t$$

$$\Delta y_t = (\phi_1 - 1) y_{t-1} + \varepsilon_t$$

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t$$

şeklini alacaktır. Öte yandan burada,



$$\gamma = \phi_1 - 1 \quad \text{ve} \quad \Delta y_t = y_t - y_{t-1} \text{ dir.}$$

Ele alınan modelde stokastik trendin varlığının test edilmesi aslında boş hipotez olarak otoregresif parametrenin 1'e eşit olmasının test edilmesine eşittir ve alternatif hipotezinde 1'den küçük olarak oluşturulması serinin trend durağan olduğunu belirtir (*Phillips ve Xiao, 1998, s.3*). Bu durumda, DF testinde, boş ve alternatif hipotezler şu şekildedir.

$$H_0 : \gamma = 0 \quad (\phi_1 = 1)$$

$$H_1 : \gamma < 0 \quad (\phi_1 < 1)$$

Birim kök testi yapılırken, ele alınan bu model en küçük kareler yöntemiyle tahmin edilir ve  $\hat{\gamma}$  parametresinin 1'e eşit olup olmadığı, t testi mantığında test edilir. Buna  $\hat{\tau}$ -test istatistiği de denilir.  $\phi_1$  sifira ne kadar yakın ise serinin o kadar çabuk durağanlaşacağı açıktır. Ayrıca, burada alternatif hipotezin tek yanlı kurulmasının nedeni ise testin gücünü mümkün olabildiğince arttırmaktır. Tahmin edilen  $\hat{\gamma}$  parametresi için t istatistiği mantığında hesaplanan  $\hat{\tau}$ -test istatistiğinin karşılaştırılacağı kritik değerler ise  $\hat{\tau}$  istatistiği, student t dağılımına göre negatif değerlere doğru çarpıklık gösterdiği için DF testi kritik değer tablosundan alınan değerle karşılaştırılır (*Redebusch, 1992, s.664*). Fakat, seride bir sabit terim yada deterministik trend varsa bu durumda bunların da regresyona eklenerek testin yapılması gerekmektedir. Böylece, bu durumda sırasıyla kesme eklenmiş modelde  $\hat{\tau}_\mu$  ve deterministik trend eklenmiş modelde ise  $\hat{\tau}_\beta$  test istatistikleri kullanılacaktır.

Daha önce gecikme uzunluğunun belirlenmesi konusunda belirtildiği gibi, yüksek mertebeden bir ARMA yapısıyla çalışılıyorsa test istatistiğini hesaplamak için tahmin edilen regresyon denkleminin hata terimlerinde otokorelasyon olacaktır. Bu sorunun giderilmesi amacıyla Said ve Dickey (1984) Genişletilmiş Dickey Fuller testini geliştirmişlerdir. Uygun gecikme yapısının belirlenmesinden sonra tahmin edilmesi gereken denklem,

$$\Delta Y_t = \mu + \beta t + \gamma Y_{t-1} + \sum_{j=1}^p \alpha_j \Delta Y_{t-j} + \varepsilon_t$$

şeklindedir. ADF testi, DF testinde kullanılan modele göre p adet fazla regresör içerir (Serletis, 1992, s.393). Denklem tahmin edildikten sonra  $\gamma$  parametresinin t değeri mantığında hesaplanan test istatistiği kritik değerlerle karşılaştırılır. ADF testi sol taraflı bir test olduğundan dolayı, eğer test istatistiği kritik değerden küçük ise birim kök hipotezi red edilmiş olacaktır.

### C. MEVSİMSELLİK

Ekonomik zaman serileri üzerine mevsimsellik ile ilgili olarak ilk çalışma James W. Gilbert tarafından 1854 yılında yapılmıştır. Bir banker olan Gilbert İngiltere Bankası banknotlarına olan talebin Ocak, Nisan, Temmuz ve Ekim aylarında oldukça yüksek olduğunu bulmuştur. Öte yandan zaman serlerinde periyodik hareketliliği ilk açıklayan kişi ise bir astronom olan Herscheldir. Herschel 1801 yılında güneş tutulmaları ile buğday fiyatları arasındaki ilişkiyi bulmuştur (Hylleberg, 1992, s.15). Devreler bir yıldan uzun olmamak şartıyla bir zaman serisindeki tekrarlanan döngüsel hareketlerin tümüne mevsimsel dalgalanma adı verilir (Serper,2004,s.398). Zaman serisi analizinde mevsimselliği düzeltilmiş serileri kullanmak oldukça sık başvurulan bir yöntemdir. Mevsimselliğin düzeltilmesi konusunda kullanılan filtre temelli yaklaşımlar X-11 tipi metotlar olarak adlandırılır. Bu prosedür başlıca 3 temel adım içerir. Birinci adımda, hareketli ortalama ile trend tahmin edilir. İkinci adımda, tahmin edilen trend seriden alınır. Üçüncü adımda, trendden arındırılmış seriden mevsimsel bileşen hareketli ortalama yöntemi ile alınır. Burada önemli olan nokta, seriden trendin tam olarak arındırılmadığı durumda mevsimsel bileşenin tam olarak bilinemeyeceğidir. Ancak diğer taraftan mevsimsel bileşenden arındırılmamış bir seride trend de tam olarak tanımlanamaz. Bu nedenle, X-11 yöntemi iteratif bir süreç olarak uygulanır (Kenny / Durbin, 1982, s.76).

Ancak bazı durumlarda, serilerden mevsimsel düzeltme ile dışlanan bilgi, politika geliřtirmede ve geliřtirilen politikaların uygulanmasının zamanlamasında oldukça önemli olabilir. Bu nedenle de mevsimsel örüntünün seriden alınmaması gerekebilir. Bu durumda, ARIMA modelleme stratejisinin bir uzantısı olan mevsimsel modelleme yöntemi SARIMA modelleri modelleme aracı olarak önerilmektedir. Öte yandan SARIMA modellerinin, kullanılan örnekleme k gecikme uzunluğunda kısaltmasından dolayı bir alternatif yöntem olan Periyodik Modeller ön plana çıkmaktadır (*Franses, 1998, s.118*). Ayrıca Tiao ve Grupe (1980), Franses (1996), Boswijk, Franses ve Haldrup (1997) ve (Osborn1988) periyodik modellerin öngörü performanslarının periyodik olmayan modellerden daha üstün olduğunu belirtmektedirler. Bu nedenlerle bu başlık altında periyodik modellerin kullanımı tartışılacaktır.

### 1. Periyodik Otoregresif Model

Periyodik otoregresif model (PEAR), AR modelinin otoregresif parametresinin zaman içerisinde mevsimler boyunca dinamikleřtirilmesine dayanır. Böylece, örneğın çeyrek yıllık veri seti ile çalışıldığı durumda, her bir çeyreğın diğerlerine göre göreceli etkisi ayrıştırılabilir (*Kurozimi, 2002, s.243*). Aslında bu yöntem ekonomik birimlerin farklı mevsimlerde farklı davranış kalıpları göstermelerinden dolayı oldukça kullanışlı bir yoldur. PEAR (p) modeli,

$$Y_{s,t} = \mu_s + \phi_{1,s}Y_{t-1} + \dots + \phi_{p,s}Y_{t-p} + \varepsilon_{s,t}$$

şeklinindedir. Çeyrek yıllık veri seti ile çalışıldığı durumda (s=4) ve gecikme uzunluğu 1 alındığında (p=1), yukarıdaki model açık olarak,

$$\begin{aligned}
Y_{4,t} &= \mu_4 + \phi_{1,4}Y_{3,t} + \varepsilon_{4,t} \\
Y_{3,t} &= \mu_3 + \phi_{1,3}Y_{2,t} + \varepsilon_{3,t} \\
Y_{2,t} &= \mu_2 + \phi_{1,2}Y_{1,t} + \varepsilon_{2,t} \\
Y_{1,t} &= \mu_1 + \phi_{1,1}Y_{4,t-1} + \varepsilon_{1,t}
\end{aligned}$$

şeklini alacaktır. Görüldüğü gibi periyodik modelde birinci dereceden bir AR yapısını açıklamak için dört otoregresif parametre kullanılmaktadır.

Bu şekilde bir modelleme için seçilebilecek iki farklı yol vardır. Bunlardan birincisi, periyodik olmayan model tahmininden elde edilen kalıntılar üzerinden periyodik varyasyonun varlığını sınanan bir takım testlere başvurmaktır. İkincisi ise, PEAR (p) modelinde p yapısını model seçim kriterlerine göre belirledikten sonra tahmin edip, otoregresif parametrelerde periyodik varyasyonun varlığının bir takım testlerle sınanmasıdır (*Franses, 1998, s. 120*). Bu çalışmada uygulama aşamasında ikinci yol kullanıldığından dolayı burada da ikinci yolun ayrıntıları sunulmuştur.

İlk olarak karar verilmesi gereken nokta kaçınıcı dereceden bir model kullanılacağıdır. Bu amaçla daha önce değinildiği gibi AIC ve SIC kriterleri kullanılabilir. Otoregresif yapı doğru olarak belirlendikten sonra ikinci olarak, modelin tahmini yapılacaktır. Bu amaçla aşağıdaki modelde doğru otoregresif yapı ele alındığı varsayımıyla,

$$Y_t = \mu_s + \phi_{1,s}Y_{t-1} + \dots + \phi_{p,s}Y_{t-p} + \varepsilon_{s,t}$$

ve  $\varepsilon_{s,t}$ 'nin normalliği varsayımı altında en çok olabirlik tahmini ve olağan en küçük kareler tahmincilerinin birbirlerine eş değer olacağından en küçük kareler yöntemi ile model tahmin edilebilir.

Üçüncü aşama ise otoregresif parametrelerde periyodik varyasyonun testidir. Boswijk ve Franses (1996), aşağıdaki hipotez çiftini test eden ve  $F(3p, n-(s+sp))$  dağılımına sahip olan bir F testi önermektedirler.

$$F = [(n-l)/s]/[(RSS_1 - RSS_0)/RSS_1]$$

şeklindeki bu testte  $RSS_1$  ve  $RSS_0$  sırasıyla kısıtsız ve kısıtlı modellerin kalıntı kareler toplamları olmak üzere ve  $n$  örneklem sayısı,  $l$ 'de bir yıldaki dönem sayısını belirtmek üzere,

$$H_0 : \phi_{p,s} = \phi_i$$

$$H_1 : \phi_{p+1,s} \neq \phi_i$$

hipotezi sınanır. Testin değeri  $3p$  ve  $n-(s+sp)$  serbestlik dereceli tablo değerini aşarsa otoregresif parametrelerde periyodik varyasyonun varlığı kabul edilir. Böylece, otoregresif parametrelerde periyodik varyasyonun varlığının kabulü ile, son aşama olan birim kök testi aşamasına başlanabilir.

## 2. Periyodik Otoregresif Modelde Birim Kök Testi

Mevsimsel etkinin büyük ölçüde deterministik yapı içerdiği düşünülse de aslında takvim ve hava koşullarına bağlı değişimler bir süre sonra ekonomik birimlerin kendi fayda fonksiyonlarının ve alışkanlıklarının değişmesini, sabitlenmemesini beraberinde getirir. Bu gibi nedenlerle, mevsimsellik stokastik bir yapıda gösterebilir (*Ashworth ve Thomas, 1999, s.736*).

Bir PEAR (1) süreci,

$$Y_{s,t} = \alpha_s Y_{s-1,t} + \varepsilon_{s,t} \quad s=1,2,3,4 \text{ ve } t=1,2,3,\dots$$

şeklinde ifade edilsin. Ayrıca burada,  $\varepsilon_{s,t} \sim N(0, \sigma^2)$  dir. Burada durağanlık koşulu  $\alpha_1 \alpha_2 \alpha_3 \alpha_4$  çarpımına bağlıdır. Eğer süreç durağan ise,  $|\alpha_1 \alpha_2 \alpha_3 \alpha_4| < 1$ , eğer periyodik olarak entegre ise,  $\alpha_1 \alpha_2 \alpha_3 \alpha_4 = 1$  olacaktır (*Castro ve Osborn, 2005, s.7*). Böylece boş hipotez ve alternatif hipotezler,

$$H_0 : \alpha_1 \alpha_2 \alpha_3 \alpha_4 = 1$$

$$H_1 : \alpha_1 \alpha_2 \alpha_3 \alpha_4 \neq 1$$

şeklinde olacaktır. Boş hipotezin getirdiği kısıt altındaki regresyon,

$$Y_t = \alpha_1 D_{1,t} Y_{t-1} + \alpha_2 D_{2,t} Y_{t-1} + \alpha_3 D_{3,t} Y_{t-3} + (\alpha_1 \alpha_2 \alpha_3)^{-1} D_{4,t} + \varepsilon_t$$

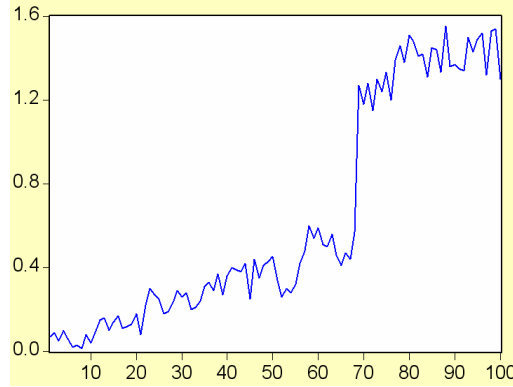
şeklinindedir. Burada D mevsimsel kukla değişkenleri temsil etmektedir. Bu model doğrusal dışı en küçük kareler yöntemi ile tahmin edilebilir. Bu modelin kalıntı kareler toplamı  $RSS_0$  ve kısıtsız genel modelin kalıntı kareler toplamı ise  $RSS_1$  olmak üzere, Boswijk ve Franses (1996)'nın önerdiği en çok olabilirlik oranı testi,

$$LR = n \log (RSS_0 / RSS_1)$$

şeklinde olup, 4 serbestlik dereceli bir  $\chi^2$  dağılımına uyar.

#### **D. YAPISAL KIRILMA**

Perron (1989)'a göre ekonomide bazı kriz ve benzeri dönemler serilerin ortalamaları üzerinde kalıcı etkiler bırakabilmektedir. Bu durumda problem serinin durağanlığının ADF tipi testlerle test edilememesidir. Çünkü böyle bir durumda, şekil 2'de de gösterildiği gibi seride kırılmanın ortaya çıktığı dönem öncesinde küçük değerler(büyük değerler) yine küçük değerler(büyük değerler) tarafından izlenirken kırılma sonrasında göreceli büyük değerler (küçük değerler) yine göreceli büyük değerler (küçük değerler) tarafından izlenir.



**Şekil 2: Yapısal Kırılmalı Bir Seri**

Böylece  $y_{t-1}$ 'in otoregresif katsayısının 1'e eşitliğini test eden birim kök testleri birim kökü red etmeme yönünde yanlış olurlar. Aslında kırılmanın olduğu dönem öncesinin ve sonrasının ayrı ayrı test edilmesi bir çözüm olarak görülse de bu durumun serbestlik derecesinde ortaya çıkaracağı azalma nedeniyle ilgi görmemektedir (Enders, 2004, s.202). Ayrıca bir çok durumda kırılmanın yeri de tam olarak bilinmediğinden seride kırılma dönemini tespit etmek için, Bai (1997), Bai ve Peron (1998), Zivot ve Andrews (1992), Vogelsang (1997), Lumsdaine ve Papell (1997) tarafından geliştirilmiş ayrı ayrı testler bulunmaktadır. Bu çalışmanın uygulama bölümünde, literatürde yaygın olarak kabul görmüş olan Vogelsang ve Perron (1998) yaklaşımı kullanılmıştır.

$T_B$  kırılma olan dönemi ve  $T$  toplam gözlem sayısı olmak üzere seride ani bir kırılma olduğunu düşünelim. Bu tip bir durum “eklenebilir aşırılık” (Additive outlier (AO)) durumu olarak bilinmektedir. (Öte yandan bu kırılma seride zaman içerisinde daha yumuşak bir şekilde ortaya çıkıyorsa bu tip bir duruma verilen ad ise “yenilikçi aşırılık” (innovational outlier)). AO durumunda olası üç ayrı model söz konusudur. Bunlar,

$$Y_t = \mu + \beta t + \theta DU_t + u_t \quad (\text{Model 1})$$

$$Y_t = \mu + \beta t + \theta DU_t + \gamma DT_t + u_t \quad (\text{Model 2})$$

$$Y_t = \mu + \beta t + \gamma DT_t + u_t \quad (\text{Model 3})$$

Burada,

$$DU_t = \begin{cases} 0 & t \leq T_B & \text{ise} \\ 1 & t \geq T_B + 1 & \text{ise} \end{cases}$$

ve

$$DT_t = \begin{cases} 0 & t \leq T_B & \text{ise} \\ 1 & t \geq T_B + 1 & \text{ise} \end{cases}$$

şeklindedir. Birinci model sadece düzeyde bir kırılma içerir, ikinci model hem kesme hem de eğimde bir kayma içerir, üçüncü model ise sadece eğimde kayma içerir. Öte yandan bu üç modelde de hata yapısı durağan bir süreç olarak varsayılmaktadır.

Hata teriminin durağan olduğu alternatif hipotezi altında  $y_t$  kırılmalı trend etrafında durağandır. Birim kök boş hipotezi altında ise,  $y_t$   $I(1)$ 'dir. Bu durumda fark serisi  $\Delta y_t$ , durağan bir süreçtir ve,

$$\Delta Y_t = \beta + \theta D_t + v_t$$

$$\Delta Y_t = \beta + \theta D_t + \gamma DU_t + v_t$$

$$\Delta Y_t = \beta + \gamma DU_t + v_t$$

eklindedir. Burada ayrıca,

$$D_t = \begin{cases} 0 & \text{eğer } t \neq T_B + 1 \\ 1 & \text{eğer } t = T_B + 1 \end{cases}$$

olup, hata terimi  $v_t$ , durağandır.

Bu şekilde birim kökü test etme prosedürü başlıca iki adımda gerçekleştirilir. Birinci adımda yukarıdaki modellerden bir tanesi trendden



ayrıştırılarak kalıntılar ( $\tilde{Y}_t$ ) elde edilir, İkinci adımda ise bu kalıntılara dayalı aşağıdaki regresyonlar tahmin edilir,

$$\tilde{Y}_t = \alpha \tilde{y}_{t-1} + \sum_{i=0}^k w_i D_{t-i} + \sum_{i=0}^k c_i \Delta \tilde{y}_{t-i} + \varepsilon_t \quad (1. \text{ ve } 2. \text{ model için})$$

$$\tilde{Y}_t = \alpha \tilde{y}_{t-1} + \sum_{i=0}^k c_i \Delta \tilde{y}_{t-i} + \varepsilon_t \quad (3. \text{ model için})$$

ve  $\alpha=1$  hipotezi test edilir.

Burada karşılaşılabilecek problem daha önce de belirtildiği gibi kukla değişkenlerin oluşturulmasında kullanılan kırılmanın döneminin tespitidir. Eğer kırılma dönemi biliniyorsa,  $\alpha=1$  hipotezinin testinde kullanılan test istatistiği Peron (1989) ya da Peron ve Vogelsang (1993)'de önerilen kritik değerlerle karşılaştırılabilir. Burada,  $\lambda = \frac{T_B}{T}$  nisbi kırılma yansımasıdır ve kritik tablo değerinin belirlenmesinde de rol oynar (Sevüktekin / Nargeleçekenler, 2007, s.407).

$T_B$ 'nin bilinmediği durumda ise Vogelsang ve Peron (1998) şu prosedürü önermektedirler.

Kırılma dönemi bilinmediği için,  $\lambda = \frac{T_B}{T}$  olmak üzere ve toplam gözlem sayısının ilk ve son %15'lik kısmı dışlanarak elde kalan gözlemler üzerinden,  $T_B=1,2,3,\dots,T$  şeklinde ayrı ayrı regresyon tahminleri yapılır.  $t_\theta$ ,  $\theta = 0$  ve  $t_\gamma$ 'de  $\gamma = 0$  hipotezlerini test eden istatistikler olmak üzere,

- Birinci modelde eğer  $\theta > 0$  ise maksimum  $|t_\theta|$ 'yi veren, ya da  $\theta < 0$  ise minimum  $|t_\theta|$ 'yi veren dönemin kırılma dönemi olduğu kabul edilir.

- İkinci modelde eğer  $\gamma > 0$  ise maksimum  $|t_\gamma|$ 'yi veren, ya da  $\gamma < 0$  ise

minimum  $|t_\gamma|$ 'yi veren dönemin kırılma dönemi olduğu kabul edilir

- Üçüncü modelde eğer  $\gamma > 0$  ise maksimum  $|t_\gamma|$ 'yi veren, ya da,  $\gamma < 0$  ise minimum  $|t_\gamma|$ 'yi veren dönemin kırılma dönemi olduğu kabul edilir.

Vogelsang ve Peron bu prosedürün kullanımıyla doğru kırılma döneminin seçilebileceğini asimptotik olarak göstermişlerdir. Bu nedenle de, yine Peron (1989) yada Peron ve Vogelsang (1992)'de verili kritik değerler kullanılabilir.

## E. VOLATİLİTE MODELLEMESİ

Engle (1982) zaman serisi modellemesinde hata teriminin sabit varyanslı olma varsayımının geçerli olmayabileceğini İngiltere'ye ait enflasyon serisini inceleyerek ortaya koymuştur. Aslında, alışlagelmiş olarak niteleyebileceğimiz yaklaşımda zaman serilerindeki temel problem otokorelasyon olarak düşünülmekte ve değişen varyans problemi yatay kesit verileriyle çalışıldığında karşılaşılabilecek bir problem olarak görülmekteydi (*Greene, 1997, s.569*). Bu nedenle de ele alınan zaman serisi sabit varyansa sahip olmadığı durumda bildik yöntemler modellemede ve öngörü yapmada başarılı olamamaktadır. Otoregresif koşullu değişen varyans (ARCH) modelleri ise varyansa öngörü hatalarının karelerinin bir fonksiyonu olarak değişim serbestliği getirmektedir. Aşağıdaki AR (1) modeli ele alındığında,

$$y_t = a_0 + \phi_1 y_{t-1} + \varepsilon_t$$

$\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$  olmak üzere,  $y_t$ 'nin koşulsuz ortalaması sıfır ve koşullu ortalaması

ise  $\phi_1 y_{t-1}$  iken, koşullu varyansı  $\sigma_\varepsilon^2$  ve koşulsuz varyansı da  $\frac{\sigma_\varepsilon^2}{1-\phi_1^2}$  şeklindedir.

(Engle, 1982, s.987).  $\phi_1$ ' in alabileceği değerler  $0 < \phi_1 < 1$  olduğundan dolayı koşullu varyans koşulsuz varyanstan daha küçük değer almaktadır.

## 1. ARCH MODELİ

Yukarıda AR(1) modelinde sözü edilen öngörü  $y_{t+1}$ 'in koşullu varyansı,

$$\begin{aligned} \text{Var}(y_{t+1} | y_t) &= E[y_{t+1} - \phi_1 y_t]^2 \\ &= E(\varepsilon_{t+1})^2 \end{aligned}$$

şeklindedir. Sabit varyans varsayımının geçersiz olduğu durumda, koşullu varyansın bir AR(q) modeli ile tahmini basit bir şekilde yapılabilir (Enders, 2004, s114). Bu yaklaşım,

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \dots + \alpha_q \hat{\varepsilon}_{t-q}^2 + v_t$$

şeklinde gerçekleştirilebilir. Burada  $v_t$  beyaz gürültü sürecidir. Langrange çarpanları testi yardımıyla yukarıdaki tahmin sürecinin bir AR(q) modeli olarak ele alınması durumunda ARCH etkisinin varlığı test edilebilir.

LM = (T-q)  $R^2$  şeklinde hesaplanan test istatistiği q serbestlik dereceli bir  $\chi^2$  dağılımına sahiptir. Bu durumda,

$$\begin{aligned} H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_q = 0 \\ H_1 : \alpha_1 \neq \alpha_2 \neq \dots \neq \alpha_q \neq 0 \end{aligned}$$

şeklindeki hipotez takımı test edilerek  $LM < \chi^2_{q \text{ tablo}}$  durumunda  $H_0$  red edilerek ARCH etkisinin varlığına ve model spesifikasyonunun uygun olduğuna karar verilebilir. Böylece tüm hata gecikmelerinin katsayıları sıfıra eşittir boş hipotezi red edilemediği durumda tahmin edilen varyans  $\alpha_0$ 'a eşit olur. Bunun anlamı

sabit varyans varsayımının geçerliliğidir. Diğer taraftan en az bir hata gecikmesinin katsayısı anlamlı derecede sıfırdan farklı ise örneğin  $\alpha_1$ , bu durumda  $x_t = \phi_1 y_{t-1}$  olmak üzere, ARCH (1) modeli,

$$\begin{aligned} y_t | Y_{t-1} &= (x_t b, h_t) \\ h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \\ \varepsilon_t &= y_t - x_t b \end{aligned}$$

şeklinde olacaktır. Böylece varyans en çok olabilirlik prosedürü yoluyla elde edilir (Poon / Granger, 2003, s.484). Bu modelde,  $\alpha_0 > 0$  ve  $0 < \alpha_1 < 1$  kısıtları vardır. Bu ikinci kısıtlama sürecin kararlılığının sağlanması için yada diğer bir ifade ile sonsuz bir varyansa sahip olmaması için gereklidir (Işığışık, 1999, s.4).

## 2. GARCH MODELİ

Bollerslev (1986) Engle'in yaklaşımını genişleterek koşullu varyansın ARMA yapısında da olduğu durumda tahminine olanak sağlamıştır. Böylece ARCH modellerine göre daha etkin geçmiş bilgiye sahip ve daha esnek bir gecikme yapısında olan bu model Genelleştirilmiş Otoresif Koşullu Değişen Varyans modeli (GARCH) adını almıştır.

Bu durumda hata yapısı,

$$\varepsilon_t = v_t \sqrt{h_t}$$

şeklinde ele alınarak ve  $\sigma_v^2 = 1$  olmak üzere,  $\varepsilon_t$ 'nin beklenen değeri,

$$E(\varepsilon_t) = E(v_t, (h_t)^{1/2}) = 0$$

şeklindedir. Koşullu ve koşulsuz ortalamaların sıfıra eşit olduğu bu model yapısı,

$$\begin{aligned}
y_t | \psi_{t-1} &\approx (0, h_t) \\
h_t &= \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i} + \sum_{j=1}^q \beta_j h_{t-j} + v_t \\
\varepsilon_t &= y_t - x_t b
\end{aligned}$$

formundadır.

Bu modeldeki kısıtlar,

$$\alpha_0 > 0, \quad \alpha_i \geq 0, \quad \beta_j \geq 0 \quad \text{ve} \quad \sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$$

şeklindedir. Öte yandan son kısıt zayıf durağanlığı işaret eder (*Cao ve Tsay, 1992, s.167*).

GARCH yapısının varlığı yine ARCH yapısının teşhisi gibi aynı mantıktaki LM testi ile test edilebilir. Ancak bu durumda hipotez takımı biraz daha karmaşık hal alır (*Sevüktekin / Nargeleçekenler, 2006, s.253*).

$$\begin{aligned}
H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_q = \beta_1 = \beta_2 = \dots = \beta_p = 0 \\
H_1 : \alpha_1 \neq \alpha_2 \neq \dots \neq \alpha_q \neq \beta_1 \neq \beta_2 \neq \dots \neq \beta_p \neq 0
\end{aligned}$$

LM test istatistiği ise,

$$LM = (T-p-q) R^2$$

şeklinde elde edilir ve  $(p + q)$  serbestlik dereceli  $\chi^2$  dağılımına uyar. GARCH  $(p,q)$  modeli  $q = 0$  için ARCH  $(p)$  modeline denk olacaktır. Son olarak, bir çok ampirik çalışmaya göre GARCH modeli ARCH modeline göre daha cimri olup, GARCH  $(1,1)$  modeli bir çok finansal zaman serisi için en popüler yapıdır (*Poon / Granger, 2003, s.484*).

## F. DOĞRUSAL DIŞILIK

İktisatta kullanılan istatistiksel ve ekonometrik yöntemler, doğrusallık varsayımı ile aslında iktisat teorisinin bazen arka planda zimni olarak kabul ettiği bazen de direkt vurguladığı doğrusal dışı ilişkileri göz ardı etmektedir. Bu durum gerçek hayata ilişkin fenomenlerin modellenmesinde kimi zaman ciddi sınırlamalar yaratmaktadır (*Arnold, 2003, s.479*). Örneğin, ücretlerin aşağıya doğru gösterdiği yapışkanlık, ekonominin genişleme dönemlerinde işsizlik oranının azalmasındaki yavaşlığa rağmen, daralma dönemlerinde işsizlik oranındaki artışın hızlı olması ya da kısa ve uzun dönemli faiz oranlarındaki farkın bir eşik değere göre farklı hareket etmesi ve benzeri iktisat teorisinden gelen açılımlar, doğrusal modelleme araçlarıyla modelleme yapıldığında göz ardı edilmiş olmaktadır. Ancak iktisadi seriler üzerinde dinamik ayarlanma mekanizmasının çalışması kalıcı etkilerin (persistence mechanism) varlığına bağlıdır (*Mikhail / Eberwein / Handa, 2003, s.1*). Örneğin, Rodriguez (2004), petrol varil fiyatlarındaki doğrusal dışılığın göz ardı edilmesinin, petrol varil fiyatlarındaki dalgalanmaların yakalanamamasına ve buna bağlı olarak da GSMH tahminlerinin doğru olarak elde edilememesine yol açtığını belirtmektedir (*Rodriguez, 2004, s.2*). Fakat bu göz ardı ediş, bazı durumlarda tam anlamıyla yanlış model kullanmak olarak adlandırılmayabilir. Çünkü doğrusallık şeklindeki tek bir yaklaşıma karşın ortaya konulan tek bir doğrusal dışılık formu yoktur. Dolayısıyla, serinin doğrusal olup olmadığının ya da doğrusal dışı bir karakter sergileyip sergilemediğinin testinde, doğrusallık hipotezinin karşıtı hipotez tek bir doğrusal dışılık formu olmayacaktır. Bu durum da yanlış bir doğrusal dışı model seçmenin maliyeti doğrusal bir model kullanmaktan çok daha yüksek olmaktadır. Bir başka ifade ile, yanlış bir doğrusal dışı model kullanmak, doğru olmadığı bilinen doğrusal bir model kullanmaktan çok daha kötü sonuçlar verecektir. Literatürde son yıllarda yapılan çalışmalara bakıldığında, doğrusal dışı modelleri kurma ve tahmin etmede katlanılan zorluklarının, önraporlama performanslarında sağladıkları (ya da sağlayamadıkları) iyileşme ile telafi edilemeyebildiği de vurgulanmaktadır (*Clement, Franses, Swanson, 2003, s.171*). Bu nedenle doğrusal dışı modellemeye başvurulacağına, sadece serinin doğrusal dışı olup olmadığının bilinmesi değil,

hangi tip bir doğrusal dışılığın olduğunun bilinmesi de son derece önemlidir. Böylece doğrusal dışı modellerden hangisinin kullanılacağına da test edilmesi gerekmektedir. Doğrusal dışılığın doğru tespit edilip, doğru doğrusal dışı modelin kullanılmasıyla en doğru önraporlama değerlerinin elde edileceği açıktır (*Kantz / Schreiber, 2003, s.4*).

Bu çalışmada kullanılan doğrusal dışı modeller aşağıda genel özellikleriyle tartışılmıştır. Tartışmada yukarıdaki paragraftaki ciddiyetle, bu modellerin sunduğu doğrusal dışılık formu ve bu formun testi ön planda tutulmuştur. Ancak daha sofistike modellere geçmeden önce daha basit iki doğrusal dışı modelin tanıtılması uygun olacaktır.

Bu modellerden ilki *Genelleştirilmiş Otoregresif Model (GAR)* modelidir. Standart AR modelindeki otoregresif gecikmeli değişkenlerin çeşitli kuvvetlerini ve bu kuvvetlerin birbirleriyle etkileşimlerini içermektedir.. En basit haliyle bir doğrusal dışı AR modeli,

$$Y_t = \phi_1(Y_{t-1})Y_{t-1} + \varepsilon_t$$

şeklindedir. NLAR(1) yada GAR(1) şeklinde ifade elden bu model,

$$\phi_1(Y_{t-1})Y_{t-1} = f(y_{t-1})$$

olarak ele alınır ve

$$Y_t = f(Y_{t-1}) + \varepsilon_t$$

olarak gösterilir.

Modele yöneltile en temel eleştiri, modelin çok sayıda parametre içerebilmesidir. Eğer iki gecikmeden daha fazla sayıda gecikme modelde ele alınıyorsa bu eleştiri haklıdır denilebilir (*Enders, 2004, s391*). Bir GAR(2) modeli,

$$Y_t = \phi + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-1} y_{t-2} + \phi_4 y_{t-1}^2 + \phi_5 y_{t-2}^2 + \phi_6 y_{t-1}^2 y_{t-2} + \phi_7 y_{t-1} y_{t-2}^2 + \varepsilon_t$$

olacaktır. Ancak tüm bunlara rağmen modelin En Küçük Kareler (EKK) yöntemiyle tahmini yapılabileceği ve bazı durumlarda standart AR modelinin ötesinde çözümler üretebilmesi nedeniyle tercih edilebilen bir modeldir (*Rothman, 1998, s.167*). Modelin geçerliliğinin testi, GAR modelindeki katsayıların ( $\phi_i$ ) geçerliliğinin testidir. Eğer bir katsayı dahi sıfırdan farklı değil ise model geçersiz olacak ve doğrusal AR modeli geçerli model olarak kabul edilecektir.

İkinci model ise *Bilineer Modeldir (BL)*. GAR modelinin iki gecikmeden sonrası için çok sayıda parametre içerdiği şeklindeki eleştirilere karşılık olarak, yüksek derecedeki AR yapısına, çok daha cimri ARMA modelleriyle yakınsama sağlanabileceği gerçeğinden hareketle modelde MA terimlerinin AR terimleriyle etkileşimlerine izin verildiği durumda BL modeline ulaşılır. Basit bir BL modeli,

$$Y_t = \phi_0 + \phi_1 y_{t-1} + c_1 \varepsilon_{t-1} y_{t-1} + \varepsilon_t$$

şeklindedir. Aynı zamanda,

$$Y_t = \phi_0 + (\phi_1 + c_1 \varepsilon_{t-1}) y_{t-1} + \varepsilon_t$$

şeklindeki düzenlemeyle, otoregresif katsayı ( $\phi_1 + c_1 \varepsilon_{t-1}$ ) olarak yorumlanmalıdır. Bu durumda bu katsayının  $\phi_0$  ortalamaya sahip rassal bir değişken olduğu mantığıyla ve  $c_1$ 'in pozitif (negatif) olması durumunda ifade  $\varepsilon_{t-1}$  ile büyür (küçülür). Bu mantık sonuçta pozitif (negatif) şokları, negatif (pozitif) şoklardan daha kalıcı kılacaktır.

Öte yandan bu modelin tahmini en çok olabilirlik yöntemiyle yapılabilir<sup>1</sup>. Modelin geçerliliğinin testi ise basit bir Lagrange çarpanları yöntemiyle elde edilebilir. Yukarıdaki model için, öncelikle bir AR(1) modeli elde edilir. Bu

---

<sup>1</sup>  $Y_{t-i} \varepsilon_{t-i}$  değerleri direkt elde edilemeyeceğinden EKK kullanılamaz.



modelin kalıntıları ( $e_t$ ), üzerinden aşağıdaki regresyon elde edilerek,

$$e_t = \alpha_0 + \alpha_1 y_{t-1} + c_1 \varepsilon_{t-1} y_{t-1} + v_t$$

T örneklem büyüklüğü olmak üzere,  $TR^2$  istatistiği hesaplanır. Eğer,  $TR^2$  istatistiği, yukarıdaki örnek için,  $\chi^2_{(3)}$  değerini aşarsa,  $Y_t$  serisinin doğrusallığını temsil eden boş hipotez red edilip doğru modelin BL modeli olduğuna karar verilir.

### 1. Eşik Değer Otoregresif Model (TAR)

Yukarıdaki GAR ve BL modelleri doğrusal dışı bir form sunsalar da iktisat teorisinin gerektirdiği bakış açısını tam olarak yansıtamamaktadırlar. Her ne kadar BL modeli pozitif ve negatif şokları ayırabilecek argümana sahip olsa da, istatistiksel modellemenin bu anlamda sunduğu TAR modelleri daha gelişmiş modeller olarak görülmektedir. Ekonominin genişletici (daraltıcı) bir rejimden daraltıcı (genişletici) bir rejime geçişi, ilgilenilen serinin dinamik ayarlama sürecini değiştirecektir. Bu durumda veri yaratma sürecinin TAR modeli olarak ele alınması, otoregresif katsayıya, örneğin  $y_{t-1} > 0$  durumunda  $\phi_1$ ,  $y_{t-1} \leq 0$  durumunda ise  $\phi_2$  gibi ayrı değerler alma şansı tanıyan bir sürecin işletilmesini sağlar. Böylece,

$$Y_t = \begin{cases} \phi_1 y_{t-1} + \varepsilon_{1t} & \text{eğer, } y_{t-1} > 0 \text{ ise} \\ \phi_2 y_{t-1} + \varepsilon_{2t} & \text{eğer, } y_{t-1} \leq 0 \text{ ise} \end{cases}$$

şeklindeki model elde edilir. AR modelleri grubundan bir model türü olarak kabul edilen TAR modelleri için AR modelleme stratejisindeki model seçim ve tahmin prosedürü uygulanabilir (*Fu / Fu / Sun, 2004, s.37*). Bu durumda  $y_{t-1} = 0$  değeri eşik değer olmak üzere, aslında  $y_{t-1} > 0$  durumunda bir doğrusal model ve  $y_{t-1} \leq 0$  durumunda bir başka doğrusal model olmak üzere,  $Y_t$  için tüm model EKK ile

tahmin edilebilecek doğrusal dışı bir model olur (*Feng / Liu, 2002, s.1-2*). TAR modellerinin makro ekonomik serilere uygulamaları Terasvirtave Anderson (1992), ampirik finasta uygulamaları ise Franses (2000) ile başlamıştır (*Mizrach, 2006, s.6*).

Öte yandan bazı eklemeler ile model geliştirilebilir. İlk olarak iki hata teriminin varyanslarının eşitliği durumunda ( $\text{var}(\varepsilon_{1t}) = \text{var}(\varepsilon_{2t})$ ),  $I_t$  gösterge değişkeni olmak üzere model,

$$Y_t = I_t \phi_1 y_{t-1} + (1 - I_t) \phi_2 y_{t-1} + \varepsilon_t$$

şeklinde yazılabilir. Bu durumda gösterge değişkeni  $I_t$  için,  $y_{t-1} > 0$  ise  $I_t = 1$  ve  $y_{t-1} \leq 0$  ise  $I_t = 0$ , olacağı açıktır.

İkinci olarak, eşik değerin  $y_{t-1} = 0$  gibi bilinen bir değer yerine, bilinmeyen bir eşik değeri yaklaşımı getirilebilir. Bu durumda veri setinden eşik değerin tahmin edilmesi gerekmektedir. Chan (1993), ele alınan  $Y_t$  serisinde, %15'lik en düşük ve en yüksek değerin dışlanması sonrasında, elde kalan %70'lik veri setinin her bir değerine potansiyel eşik değeri muamelesi yapmayı önermektedir. Böylece eğer 100 gözlemlilik bir veri seti ile çalışıyorsak, toplamda 30 gözlemi dışlayıp geri kalan 70 gözlem için 70 ayrı modelden, elde ettiğimiz kalıntı kareler toplamının en küçüğünü veren modelin eşik değeri, kullanılacak eşik değeri olarak belirlenir.

Üçüncü bir ekleme, dinamik ayarlanma sürecinin  $y_{t-1}$ 'den değil de daha önceden başladığını düşünerek yapılabilir. Bu durumda farklı  $y_{t-d}$  değerleri kullanarak, dinamik ayarlanmanın başladığı uygun gecikmenin en optimal ve kullanışlı tahmini Akaike bilgi kriterine göre belirlenebilir (*Fu / Fu / Sun, 2004, s.38*).

Modele bir başka ekleme, ilk modelde verilen rejim sayısı üzerinden yapılabilir. Örneğin Balke ve Flomby (1997) faiz oranları dönemsel yapısı üzerine

yaptıkları çalışmada, iki eşik değeri kullanarak üç rejim içeren bir modelle çalışmışlardır.

Modelde son bir ekleme, Enders ve Granger (1998)'den gelmiştir. Eşik değeri belirlemede  $y_{t-d}$  şeklinde düzey yerine,  $\Delta y_{t-d}$  şeklinde değişimin bir eşiği göre durumu kullanılabilir. Böylece model,

$$I_t = \begin{cases} 1 & \text{eğer, } \Delta y_{t-d} > \tau \text{ ise} \\ 0 & \text{eğer, } \Delta y_{t-d} \leq \tau \text{ ise} \end{cases}$$

olmak üzere,

$$Y_t = I_t \phi_1 y_{t-1} + (1 - I_t) \phi_2 y_{t-1} + \varepsilon_t$$

şeklinde oluşturulmaktadır. Bu model eşiğin iki tarafında  $Y_t$ 'nin farklı momentumlara sahip olduğu şeklinde dizayn edildiğinden M-TAR adını alır.

Öte yandan Clements ve Smith (2001), ne kadar iyi önraporlama yapıldığının nerede tahmin yapıldığına bağlı olduğunu belirterek, ekonominin genişleme dönemlerinde doğrusal modellerinde iyi önraporlama sonuçları verdiğini, ancak eşik değeri modellerinin asıl farkının resesyon dönemlerinde yapılan önraporlamalarda ortaya çıktığını belirtmektedirler (*Clements / Smith, 2001, s.137*)

Tüm bunlara ek olarak, TAR modelleriyle ortaya çıkan en büyük yenilik aslında üzerinde halen çalışmaların yürütüldüğü ancak ana fikrin Balke ve Fomby (1997)'den çıktığı, ve zaman serisi literatürünün gelecek 5-10 yılında bir takım yenilikleri getirecek ve doğrusal dışı analizleri daha fazla içermesine yol açacak, birim kök ve doğrusal dışılığın birlikte ele alınması yaklaşımıdır. Şöyle ki, seriyeli ilişkin dinamik ayarlanma sürecinin asimetric olduğu durumda Balke ve Flomby (1997), birim kök testlerinin gücünün düştüğünü göstermiştir. Benzer bir duruma, Caner ve Hansen (1997) çalışmasında ABD işsizlik oranı verisi için yaptıkları çalışmada da rastlanmaktadır. Bu çalışmada Caner ve Hansen, eşik değerin

üzerindeki rejim için, işsizlik oranının sabit kaldığını veya arttığını ve gözlemlerin %80'inin bulunduğu bu rejimden dolayı, doğrusal yaklaşımda serinin birim köke sahipmiş gibi görüldüğünü belirtmektedirler (*Caner / Hansen, 1997, s.22*). Aslında veri yaratma süreci TAR modeli olarak ele alındığında ise seri durağandır.

Sonuçta, TAR ve M-TAR modelleri hem iktisat teorisinden gelen açılımları yakalamada hem de birim kök süreci için getirdikleri yeni bakış açısıyla doğrusal dışı modellemenin önemli bir aracıdır. Bu modellerin kullanımında, hem sundukları doğrusal dışı formun testi hem de birim kök prosedürünün işletilmesi için Enders (2004) aşağıdaki adımsal prosedürü sunmaktadır.

1.Adım: Eşik değeri biliniyorsa klasik yaklaşımla, bilinmiyorsa Chan(1993) yaklaşımıyla TAR modeli kurulur.

2. Adım: Yine eşik değerin bilindiği ya da bilinmediği durum göz önünde tutularak M-TAR modeli elde edilir. 1. ve 2. adımda elde edilen bu iki modelin Akaike bilgi kriterine göre seçimi yapılır.

3. Adım : Seçilen model üzerinden,

$$\Delta Y_t = I_t \rho_1 y_{t-1} + (1 - I_t) \rho_2 y_{t-1} + \varepsilon_t$$

$$I_t = \begin{cases} 1 & \text{eğer, } y_{t-d} > \tau \\ 0 & \text{eğer, } y_{t-d} \leq \tau \end{cases}$$

yada

$$Y_t = I_t \rho_1 y_{t-1} + (1 - I_t) \rho_2 y_{t-1} + \varepsilon_t$$

$$I_t = \begin{cases} 1 & \text{eğer, } \Delta y_{t-d} > \tau \\ 0 & \text{eğer, } \Delta y_{t-d} \leq \tau \end{cases}$$

$\rho_1 = \rho_2 = 0$  boş hipotezi F testi kullanılarak test edilir. Ancak burada F

testinin dağılımı, gözlem sayısı ve kullanılan gecikmeli değerlerden dolayı bozulacağından klasik F tablosu değil Enders ve Grenger (1998) veya Enders ve Siklos (2001)'de verilen F tablosu kullanılarak yapılmalıdır (*Wei / Su, 2007, s.5*).

#### 4. Adım :

Eğer birim kök olmadığı sonucuna varılır ise, dinamik ayarlanma sürecinin simetrik olup olmadığı testine geçilebilir. Bu durumda standart F testi ile sınanacak boş hipotez,  $\rho_1 = \rho_2$  şeklindedir.

Öte yandan Hansen (1997) bu sınama için hata terimlerinin varyansları üzerinden yapılacak F testinin dağılımının bozulduğunu çünkü,  $\rho_1$  ve  $\rho_2$ 'nin EKK tahmincilerinin yakınsama özellikleri küçük örneklerde zayıf olduğunu belirtmektedir (*Hansen, 1997, s.5*).

5. Adım: Bu son adımda kabul edilen modelin kalıntıları kontrol edilmeli ve serisel korelasyon araştırılmalıdır.

## **2. Düzgün Geçiş Modelleri (STAR)**

TAR modelinin iktisat teorisince vurgulanan rejim değişikliklerini yakalamadaki başarısına karşın, bu rejim değişikliklerini hızlı ya da başka bir ifade ile keskin bir şekilde olduğunu varsayması eleştirilebilir. Bu nedenle bazı durumlarda, istatistiksel modellemenin sunduğu STAR modelleri ailesi kullanışlı bir araçtır. STAR modelleri, rejim geçişlerinin daha yumuşak bir şekilde gerçekleşmesine izin veren modellerdir. Bu nedenle daha sofistike bir modelleme aracı olarak görülürler. TAR modellerinin STAR modellerinin özel bir durumu olduğu rahatlıkla söylenebilir. STAR modelleri iki rejim arasındaki geçiş kullandıkları geçiş fonksiyonuyla bir yumuşaklık kazandırır. Bu amaçla kullanılan iki geçiş fonksiyonu bulunmaktadır. Bu fonksiyonlar lojistik fonksiyon ve üstel fonksiyondur.

Bir NLAR modeli aşağıdaki şekilde olsun,

$$Y_t = \phi_0 + \phi_1 y_{t-1} + \beta_1 y_{t-1} f(y_{t-1}) + \varepsilon_t$$

burada  $f(\cdot)$  fonksiyonu olarak lojistik fonksiyon kullanıldığında model lojistik STAR yada LSTAR modeli adını alır ve,

$$Y_t = \phi_0 + \phi_1 y_{t-1} + \theta(\beta_0 + \beta_1 y_{t-1}) + \varepsilon_t$$

$$\theta = [1 + \exp(-\gamma(y_{t-1} - c))]^{-1}$$

şeklinde olur. Burada,  $\gamma$ , geçiş oranı ya da düzgünleştirme parametresidir.  $c$  ise, bir rejimden diğerine geçişi temsil eden eşik değeridir. Bu yapı hem modelin parametrelerin doğrusal dışılığına, hem de rejim değişikliklerinin olmasına izin verir (Rodriguez, Torra, Felix, 2004, s.3). LSTAR modeli kullandığı geçiş fonksiyonu monotonik olarak artan olduğu için, dengeden küçük ve büyük sapmalara farklı davranış gösterebilme izni vermektedir (Haug/ Siklos, 2006, s.10).

Eğer geçiş fonksiyonu olarak üstel fonksiyon kullanılırsa bu durumda,

$$\theta = [1 - \exp(-\gamma(y_{t-1} - c)^2)]$$

şeklinde olacaktır. Model bu durumda üstel STAR yada ESTAR adını alır. Geçiş fonksiyonu olarak kullanılan bu iki fonksiyonun alabileceği uç değerlerin 0 ve 1 arasında olduğu açıktır. Bu nedenle bu iki uç değer  $\theta = 0$  ve  $\theta = 1$  durumlarında  $y_t$ 'nin dinamik ayarlanma süreci yukarıdaki örnek için sırasıyla,

$$Y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t$$

$$Y_t = (\phi_0 + \beta_0) + (\phi_1 + \beta_1) y_{t-1} + \varepsilon_t$$

şeklinde olacaktır. Öte yandan  $0 < \theta < 1$  durumunda ise model  $Y_{t-1}$ 'in alacağı değerlere göre doğrusal dışı bir ayarlama sürecine sahip olacaktır. Bu süreç LSTAR modelinde asimetric doğrusal dışı bir ayarlanma süreci iken, ESTAR

modelinde ise simetrik bir doğrusal dışı ayarlanma sürecidir (*Liew / Baharumshah / Lau, 2002, s.3*). ESTAR modelindeki bu simetrik süreç, serinin dönüm noktalarının farklı periyotlarda diğerlerine göre farklı otoregresif azalmalar şeklinde ve  $Y_{t-1} = c$  etrafında gerçekleşir (*Enders, 2004, s.401*).

STAR modelinin bir dezavantajı katsayılar çarpımsal olduğu için model tahmininde EKK yönteminin kullanılamamasıdır. Parametrelerin tahmini doğrusal olmayan EKK ile gerçekleştirilir. Ancak bu iteratif bir süreç olduğundan,  $\gamma$  ve  $c$ 'nin başlangıç değerlerinin salınımlarının durduğu durumda sona erdirilir ve diğer parametrelerin tahminleri elde edilmiş olur. Burada, Enders (2004), bir başka dezavantaj olarak, doğrusal olmayan EKK yönteminde tahmin edilen parametreler için normal dağılım varsayımı yapılmamasından dolayı, t testi değerlerinin güvenilirliklerini kaybettiğini belirtmektedir.

Son olarak, STAR modellerinin geçerliliğinin doğrusal modellere karşı test edilmesi gerekmektedir. Enders (2004), Terasvirta (1994) tarafından geliştirilen test ile, hem bu modellerin doğrusal modellere karşı, hem de kendi aralarında geçerliliklerini test etmek için aşağıdaki adımsal prosedürü önermektedir. Böylece veri yaratma sürecinin doğrusal dışı olup olmadığının yanında LSTAR yada ESTAR tipi bir doğrusal dışılık olup olmadığı da anlaşılacaktır.

1.Adım : Uygun gecikme uzunluğundaki AR(p) modeli AIC ve SIC yardımıyla belirlenip tahmin edilir.

2.Adım : Birinci adımdaki modelin kalıntıları ( $e_t$ ) ve aşağıdaki değişkenler ile şu regresyon elde edilir.

$$\begin{aligned} e_t = & a_0 + a_1 y_{t-1} + \dots + a_p y_{t-p} \\ & + b_1 y_{t-1} y_{t-d} + \dots + b_p y_{t-1} y_{t-p} \\ & + c_1 y_{t-1} y_{t-d}^2 + \dots + c_p y_{t-1} y_{t-d}^2 \\ & + d_1 y_{t-1} y_{t-d}^3 + \dots + d_p y_{t-1} y_{t-d}^3 + \varepsilon_t \end{aligned}$$

T gözlem sayısı olmak üzere yukarıdaki regresyonun  $TR^2$  istatistiği elde edilir ve  $\chi^2$  tablo değeri ile karşılaştırılır. Eğer  $TR^2$  istatistiği kritik ki-kare tablo değerini aşıyorsa doğrusallık boş hipotezi red edilir, alternatifi olan düzgün geçiş modelinin geçerli olduğuna karar verilir.

3.Adım : Düzgün geçiş modellerinin doğru olduğu kabul edildikten sonra yapılması gereken, LSTAR yada ESTAR modellerinden hangisinin geçerli model olduğuna karar vermektir. Bu amaçla ikinci adımdaki regresyon üzerinden,

$$H_0 : d_1 = d_2 = \dots = d_p = 0$$

boş hipotezi test edilecektir. Eğer bu hipotez red edilirse, geçerli modelin LSTAR modeli olduğuna karar verilir. Eğer bu hipotez red edilemiyorsa geçerli model ESTAR modelidir.

Bu son adımdaki test edilecek boş hipotezin elde edilmesi aslında lojistik ve üstel fonksiyonların üçüncü dereceden Taylor serisi açılımlarının ortaya konulmasıyla bulunmuştur. Lojistik fonksiyonda,  $-\gamma(y_{t-d} - c) = m_{t-d}$  olarak kabul edilirse ve kısmi türev sonucu  $m_{t-d} = 0$  olarak alınır,

$$\begin{aligned} \frac{\partial \theta}{\partial m_{t-d}} &= 1/4 \\ \frac{\partial^2 \theta}{\partial m_{t-d}^2} &= 0 \\ \frac{\partial^3 \theta}{\partial m_{t-d}^3} &= -1/8 \end{aligned}$$

elde edilecektir. Aynı şekilde üstel fonksiyonda benzer işlemlerle,

$$\begin{aligned} \frac{\partial \theta}{\partial m_{t-d}} &= 0 \\ \frac{\partial^2 \theta}{\partial m_{t-d}^2} &= 2 \\ \frac{\partial^3 \theta}{\partial m_{t-d}^3} &= 0 \end{aligned}$$



olacaktır. Böylece, Taylor serisi açılımıyla, LSTAR modeli için,

$$Y_t = \alpha_0 + \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + (\beta_0 + \beta_1 y_{t-1}) + \dots + \beta_p y_{t-p} (\pi_1 m_{t-d} + \pi_3 m_{t-d}^3) + \varepsilon_t$$

yazılabilir.

Taylor serisi açılımıyla, ESTAR modeli için ise,

$$Y_t = \alpha_0 + \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + (\beta_0 + \beta_1 y_{t-1}) + \dots + \beta_p y_{t-p} (\pi_3 m_{t-d}^2) + \varepsilon_t$$

yazılabilir. ESTAR modelinde kübik hiçbir terimin bulunmaması,

$$H_0 : d_1 = d_2 = \dots = d_p = 0$$

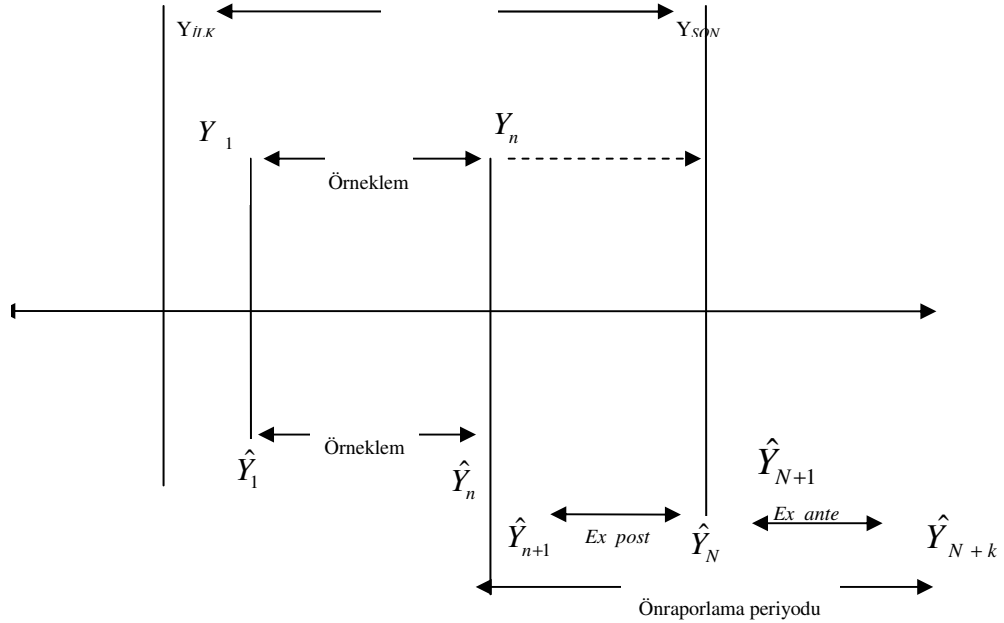
şeklindeki boş hipotezin oluşturulmasını sağlamaktadır.

Ancak hemen burada belirtilmelidir ki ilgili modele Taylor serisi açılımı ile yakınsamada çok fazla sayıda parametre gerekebilir, bu tip durumlarda testin güvenilirliği kaybolacaktır (Pena / Rodriguez, 2006, s.6).

## II. ÖNRAPORLAMA

Modern hayatın gösterdiği gelişim karşısında her alanı daha karmaşık hale gelen günlük yaşam, geleceğe bakışın daha güvenilir olmasını ve bu nedenle de geleceğe ait planların ve tahminlerin daha tutarlı bir şekilde yapılmasını gerektirmektedir. Bunun bir parçası olarak da iktisadi serilerin önraporlamaları iktisadi birimlerin politika hedeflerini oluşturmaları için önem arz etmektedir. Bu açıdan doğrusal olmayan modellemenin makro ekonomik teorideki önemi yanında, önraporlama konusunda da bazı avantajları bulunmaktadır (Nakamura, 2005, s.374) Bu başlığa kadar üzerinde durulan modellerin önraporlama performanslarının karşılıklı olarak ortaya konulması da yapay sinir ağlarının

İktisadi seriler üzerinde kullanımının hangi durumlarda daha iyi sonuçlar verebileceği sorusuna cevap niteliği taşımaktadır. Böylece yüksek önraporlama yeteneğine sahip modeller olarak tanıtilan yapay sinir ağı modellerinin, günümüzde bazıları klasikleşmiş bazıları daha modern ve karmaşık olan modelleme araçları karşısındaki performansları da değişik zaman serisi özellikleri gösteren seriler üzerinden denenmiş olacaktır. Zaman serisi modelleri önraporlama konusunda bazı önemli avantajlara sahiptir. Öncelikle zaman serileriyle yapılan önraporlamalarda önceden gerçekleşmiş değerler kullanılır. İkinci olarak, önraporlamanın doğruluğunun ölçülmesinde bazı ölçüm yöntemleri mevcuttur. Üçüncü olarak, bir kez model kurulduğunda önraporlama elde etmek için çok fazla çabaya ve zaman harcamaya gerek yoktur. Son olarak da, nokta tahmini ve aralık tahmini şeklinde önraporlama elde edilebilir (Gaynor / Kirkpatrick, 1994, s.6). Aşağıda şekil 28 yardımıyla önraporlama ile ilgili bazı kavramlar sunulmuştur.



**Şekil 3: Önraporlama ile İlgili Ana Kavramlar**

**Tarihsel Veri Seti:**  $Y_{ilk}$  ve  $Y_{son}$  arasındaki gözlemlerden oluşur.  $Y_{ilk}$  veri toplamaya başlanan ilk tarihe ait veri iken  $Y_{son}$  değerinin günümüze ait olması en iyi durumdur.

**Örnekleme Periyodu:**  $Y_1$  ile  $Y_n$  arasındaki gözlemleri kapsar. Modelin kurulacağı ve model tahmininin yapılacağı dönemdir. Modelden bu döneme karşılık elde edilecek değerler  $\hat{Y}_1$  ve  $\hat{Y}_n$  olacaktır. Aynı zamanda bu değerler örneklem içi önraporlama değerleridir.

**Önraporlama Periyodu:** Expost ve Exante olmak üzere iki ayrı periyottan oluşmaktadır. *Ex post periyot*, tarihsel veri setinin kapsadığı dönem içinde yapılan önraporlamadır. Bu dönemden elde edilen ve örneklem içi önraporlamadan elde edilen değerler ile modelin doğruluğu hesaplanabilir. *Ex ante periyot* ise, elde gerçekleşmiş herhangi bir veri yoktur. Bu nedenle doğrulayıcı bir bilgi elde edilmesi mümkün değildir. Eğer örneklem içi ve ex post periyodun performansı tatmin edici ise, ex post periyottaki gözlemlerin modele dahil edilerek modelin tekrar kurulması ile ex ante periyottaki önraporlama performansının artırılabilceği düşünülmektedir. Çünkü bu son eklenen ex post döneminin, ex ante dönemdeki gözlemler ile korelasyonu yüksektir (DeLurgio, 1998, 210).

Önraporlama yapılan modelin önraporlama için uygunluğu, modelin ürettiği önraporlama değerlerinin ( $\hat{Y}$ ) gerçek değerlere ( $Y$ ) ne kadar yakın olduğuna bağlıdır (Gaynor / Kirkpatrick, 1994, s.6). Bu durumda önraporlama hatasını ( $e$ ) ile gösterirsek,

$$e_t = (Y_t - \hat{Y}_t)$$

olacaktır. Her bir döneme ait hata ( $e_t$ ), önraporlama sonucu elde edilmiş değer

$(\hat{Y})$  etrafında rassal dalgalanmalar olacaktır ve  $\sum_{t=1}^n e_t = 0$  'dır. Bu nedenle önraporlama performansı için kullanılan ölçütler  $|e|$  yada  $e^2$  üzerinden oluşturulurlar. Aşağıda tablo 1'de önraporlama performansının değerlendirilmesinde kullanılan ölçütler sunulmuştur. Tabloda sırası ile MAE (ortalama mutlak hata), MAPE (Mutlak yüzde hata ortalaması), MSE (Ortalama kare hata), RMSE (kök ortalama kare hata) ve U (Theil'in eşitsizlik katsayısı) ölçütlerinin hesaplandıkları formüller verilmiştir.

<u>ÖLCÜT</u>	<u>FORMÜL</u>
MAE	$\frac{\sum_{t=1}^n  e_t }{n}$
MAPE	$\frac{\sum_{t=1}^n \frac{ e_t }{Y_t}}{n}$
MSE	$\frac{\sum_{t=1}^n e_t^2}{n}$
RMSE	$\sqrt{\frac{\sum_{t=1}^n e_t^2}{n}}$
U	$\frac{\sqrt{\sum_{t=1}^n [(\hat{Y}_t - Y_t)^2 / n]}}{\sqrt{\sum_{t=1}^n (\hat{Y}_t^2 / n)} \sqrt{\sum_{t=1}^n (Y_t^2 / n)}}$

**Tablo 1 : Önraporlama Performans Ölçütleri**

Eğer bir ya da birkaç tane göreceli büyük hata değeri ile çalışılacak ise, MSE ve RMSE değerleri bu durumda uygun olmayacağından MAE tercih edilir. Eğer tüm hata değerleri benzer büyüklükte ise, MSE yada RMSE tercih edilmelidir. Öte yandan, önraporlama performansları karşılaştırılan modeller, farklı frekanstan ya da farklı büyüklüklere sahip veri setlerinden elde edilmişse bu durumda MAPE kullanılmalıdır. Theil'in eşitsizlik katsayısının kullanıldığı durumda "0" değeri mükemmel uyumu verecektir ancak bu teoride mümkün bir durum olup, "0.55" in altındaki U değerine sahip modeller iyi önraporlama veren modeller olarak değerlendirilebilir (Lindberg, 1982, s.369).

Önraporlama metodu olarak başlıca iki yol kullanılabilir. Bunlardan birincisi, sadece bir önraporlama değerinin yine sadece bir periyot sonrası için elde edildiği tek adım yöntemidir. Bu metotta, örneğin model  $Y_t$  değerine kadar gözlemlerden oluşmuş ise,  $Y_{t+1}$  değerinin modelden elde edilmesidir. İkinci metod ise, çoklu adım yöntemidir ve bu durumda yine  $Y_t$  dönemine kadar gözlemlerden oluşmuş modelden en azından  $Y_{t+1}$  ve  $Y_{t+2}$  değerleri önraporlama olarak elde edilir. Eğer Ex post dönem için önraporlama yapılıyorsa bu durumda iki seçenek vardır. Bunlardan ilki, genelde kısa dönemli önraporlamalar için kullanılan ve  $Y_{t+1}$  değeri elde edildikten sonra  $Y_{t+2}$  değeri için bu bilgiyi kullanan iteratif yöntemdir. İkinci yöntem ise, her bir önraporlama yapılacak dönem için ayrı önraporlama modeli kurulmasını gerektiren bağımsız yöntemdir. Son olarak, zaman serisi modelleriyle önraporlama yapıldığında, izlenecek adımlar sırasıyla, (i) tanımlama, (ii) tahmin, (iii) diagnostik kontrolü, (iv) önraporlama değerlerinin elde edilmesi ve (v) önraporlama performansının ölçülmesi olmalıdır ( Salam / Salam / Feridun, 2006, s.144)

## İKİNCİ BÖLÜM

### YAPAY SİNİR AĞLARINA GİRİŞ VE TEMEL MODELLER

İnsanoğlu sahip olduğu medeniyeti, alet kullanma becerisi ve öğrenme yeteneğine borçludur. Alet kullanma becerisi günlük hayatın gelişimini sağlayıp pratikte de öğrenileceklerin sayısını arttırırken diğer taraftan öğrenme yeteneği ise, kullanılan aletlerin evrimini şekillendirmiştir. Alet kullanma becerisi, ve öğrenme yeteneği arasında günlük hayat üzerinden ortaya çıkan bu etkileşim geniş bir zaman süreci içerisinde makineleri ortaya çıkarmıştır. Tam da bu noktada bir ironi, insanoğluna insan gibi çalışan bir makine yaratma düşüncesinin her zaman için çekici gelmesidir. Elbetteki alet kullanma becerisinin büyük bir kısmının, öğrenme yeteneğinin tamamının ve de bu ironinin temelinde insan beyni vardır. İnsan beyninin modellenmesi üzerine ilk çalışmalar 19. yüzyılın sonlarında başlamıştır. O zamandan günümüze yapılan çalışmalarda, farklı cümlelerle de olsa iki ana vurgu ön plana çıkmaktadır.

Birinci temel vurgu paralel çalışma üzerinedir. Örnek olarak Wiggins ve Looper (1991), bu vurguyu şu şekilde vermektedir: Beyindeki nöronlar temel hesaplama üniteleridir. Bir insan beyni ortalama 100 milyar nöron içermekte ve bu hücreler birbirleriyle bağlantılı olarak paralel çalışmaktadır. Bir nöron ortalama bin başka nöronla bağlantılıdır. İnsan beynindeki nöronlar arasında yaklaşık 10 trilyon bağlantı bulunmaktadır. Temel olarak beynin çalışması da bu sinir ağları tarafından gerçekleştirilmektedir (*Wiggins / Looper, 1991, s.4*). Bu vurguyla başlayan çalışmalar ikinci olarak dijital bilgisayarlardaki mikroçip sayıları ve bunlar arasında birim zamanda gerçekleşen bağlantı sayısı üzerinden bir tartışma ile hız karşılaştırması yaparlar.

İnsan beyninin çalışma prensibi ile ilgili olarak yapılan çalışmalarda ikinci temel vurgu ise, adaptasyon yeteneği ve kendi kendini organize edebilme üzerinedir. Örneğin Werner ve Misra (1996), bu vurguyu şu şekilde yapmaktadır:

Günümüzün bilgisayar teknolojisinin ulaştığı seviyede dahi insan beyni bilgisayarlarda bulunması arzulanan birçok özelliğe sahiptir. Özellikle, özet bilgilerden genel bir fikre ulaşabilme, örüntü tanıma, hızlı bir şekilde anı hatırlama, lokal bir arıza halinde bile çalışabilme vb. şeklinde örneklendirilebilecek bu özellikler, insan beyninin modellenmesi ve çalışma prensiplerinin anlaşılması konusunda en önemli motivasyonu oluşturmaktadır (*Werner / Misra, 1996, s.286*). Bu vurguyu ön planda tutarak başlayan çalışmalar ise alıntıdan da anlaşılacağı gibi, bilgisayarların ulaştığı hıza rağmen insan beyninin sahip olduğu özelliklerin anlamı üzerinde dururlar.

Cheng ve Titterington (1994), insan beyni ve bilgisayar arasındaki işlem hızındaki farklılığın aslında bir işlem organizasyonu farklılığı olduğunu, adaptasyon farklılığının ise aslında yeni bilgi ve perspektif keşfetme ve kendini buna adapte edebilme yapısının varlığından ya da bir başka ifadeyle öğrenme yeteneğinden kaynaklandığını belirtmektedir (*Cheng / Titterington, 1994a, s.6*).

Yapay sinir ağlarına giriş amacı taşıyan bu bölümde, yapay sinir ağlarının tarihsel gelişimi, kendisine özgü terminolojisi ve yapay sinir ağlarında temel modeller ele alınacaktır. Böylece üçüncü bölümü oluşturan Çok Katmanlı Algılayıcılar ve Geri Dönüşümlü Ağlar içinde bir alt yapı hazırlanmış olacaktır. Bilindiği gibi istatistiksel modelleme konusu başlıca üç başlıkta ele alınabilir, bunlar sırasıyla yoğunluk tahminleri, sınıflandırma problemleri ve regresyondur (*Jordan / Bishop, 1996, s.1*). Yoğunluk tahminleri başlığı altında ele alınan analizlere Temel bileşenler analizi, Faktör analizi, Kümeleme analizi, Sınıflandırma problemleri başlığı altında ele alınan analizlere Diskriminant analizi, Lojistik regresyon v.b, örnek olarak verilebilir. Bu çalışmanın konusu makro ekonomik zaman serileri analizinde yapay sinir ağlarının kullanımı olduğundan geniş yapay sinir ağı modelleri ailesinin regresyon tipi kullanımı ele alınacaktır.

Bu konunun biraz daha açılması gerekirse, yapay sinir ağı terminolojisi başlığı altında anlatılacağı gibi, parametre tahminlerinin bir başka deyişle nöronlar arası bağlantıların ağırlıklarının belirlenmesi eğitim ya da diğer adıyla öğrenme

olarak adlandırılır. Yapay sinir ağı literatüründe öğrenme teriminin anlamı örneklemeden alınan bilginin aslında henüz elde olmayan gözlemler için genelleştirilmesidir. Bu nedenle yapay sinir ağı bazen öğrenen makine olarak da adlandırılır (*Hammer / Villmann, 2003, s.59*). Başlıca üç ana tip öğrenme stratejisi vardır. Bunlar öğretmenli öğrenme, öğretmensiz öğrenme ve destekleyici öğrenme stratejileridir. Öğretmenli öğrenme stratejisinde, girdi vektörü ve karşılığında sistemin oluşturması beklenen çıktı vektörü sisteme verilir. Sistemden beklenen ilgili ağırlık değerlerini bu girdi – çıktı setine göre (bağımlı – bağımsız değişken ilişkisine dayanarak) hesaplayarak, verilen girdiye karşılık doğru çıktıyı üretmesidir. Bu stratejiyi kullanan başlıca yapay sinir ağı modelleri: Çok katmanlı algılayıcılar, Geri dönüşümlü ağılar, Hopfield ağıları sayılabilir. Öğretmensiz öğrenme stratejisini kullanan ağılarda ise sisteme sadece girdiler verilir ve ağdan veri kümesinin sahip olduğu dağılımdan çıktıları üretmesi beklenir (*Oja, 2002, s.189*). Direkt olarak veri setinin gösterdiği dağılım üzerinden kümeleme yapılmaktadır. Bu açıdan, bu öğrenme stratejisini kullanan ağılar, yoğunluk tahmini olarak adlandırılan kümeleme analizi , faktör analizi v.b. analizlere benzerlik gösterirler. Destekleyici öğrenmede ise sisteme yine sadece girdiler verilirken üretmesi beklenen çıktılar sisteme sunulmamaktadır (*Ghahramani,2004, s.3*). Bir başka ifade ile, girdi çıktı ikilisi bulunmamakta ancak üretilen çıktının doğru ya da yanlış olduğu sisteme çevre tarafından, çevre durumundaki değişimle belirtilmektedir (*Kaelbling / Littman / Moore, 1995, s.239*) Bu öğrenme stratejisini kullanan ağılara örnek olarak Doğrusal vektör parçalama modeli (LVQ) ağı verilebilir. Bu ağılar içerisinde hem istatistik biliminin yaklaşımı çerçevesinde hem de kullanım alanı ve sıklığı yönüyle bakıldığında çok katmanlı algılayıcılar en önde gelen ağı yapısı olmaktadır.

Sonuç olarak bu çalışmanın ilerleyen başlık, alt başlık ve bölümlerinde ele alınacak ağılar öğretmenli öğrenme stratejisini kullanan ağılar olacağından, öğrenme ya da eğitim ile öğretmenli öğrenme kastediliyor olacaktır.



## I. YAPAY SINİR AĞLARINDA TARİHSEL GELİŞİM

1890 yılında William James “Psychology (Briefer Course)” adlı yayınıyla insan beyninin yapısı ve fonksiyonlarıyla ilgili ilk çalışmayı yapmıştır. Onun oluşturduğu genel konsept sinir ağı çalışmalarında halen kullanılmaktadır (*Wiggins / Looper, 1991, s.1*). 1911 yılında insan beyninin nöronlardan oluştuğu fikri genel kabul görmüştür. 1940’lardan önce bazı yapay sinir ağı çalışmaları olsa da, bu alanda ilk ciddi çalışmanın 1943 yılındaki Warren McCulloch ve Walter Pitts tarafından yapıldığı kabul edilir. Matematiksel işleyişine temel modeller başlığı altında değinilecek olan bu model basit bir ikili eşik değer ünitesi içermektedir. Öte yandan her ne kadar basit bir model olarak ifade edilse de, McCulloch - Pitts yapay sinir hücresi, temelde dijital bilgisayarların çalışma prensiplerine eşdeğer olarak gösterilmektedir (*Abu-Mostafa, 1986, s.7*). 1950’lerin sonuna gelindiğinde iki yeni gelişme yaşanmıştır, bunlar sırasıyla basit algılayıcıların geliştirilmesi ve ADALINE ünitesinin ortaya konulmasıdır. Basit algılayıcı, Rosenblatt tarafından örüntü sınıflandırma amacıyla geliştirilmiştir (*Wendemuth, 2006, s.1*). ADALINE ise Widrow ve Hoff’un çalışmaları sonucu ortaya konulmuş ve doğrusal örüntü tanımlama amacıyla geliştirilmiştir (*Öztemel, 2003, s. 68*). Bu iki modelin matematiksel işleyişi de yine temel modeller başlığı altında sunulacaktır.

Yapay sinir ağı çalışmaları tarihinde Marvin Minsky ve Seymour Papert’in 1969 yılında yayınladıkları ‘Perceptrons’ adlı kitap önemli bir yer tutmaktadır. Minsky ve Papert bu çalışmalarında XOR problemi olarak adlandırılan ve doğrusal olmayan bir ilişkiyi gösteren probleme basit algılayıcıların çözüm getiremediğini göstermişlerdir. Oysa Rosenblatt geliştirdiği basit algılayıcı modellerinin sonlu sayıda iterasyonla ilgili sınıflandırma problemini çözebileceğini yada sinir ağı terminolojisindeki tanımıyla örüntü tanımlanmasını yapabileceğini (öğrenebileceğini) belirtmektedir. Bu sonuç, Rosenblatt’ın haksız olduğunu göstermiş ve yapay sinir ağlarına olan ilgiyi ve bununla ilişkili olarak finansal desteği azaltmıştır. 70’lerde ve 80’lerin başında hemen hepsi Asya yada Avrupada bulunan sınırlı sayıda araştırmacı yine az sayıda çalışmayla yapay sinir

ađı konusunda alıřmalarda bulunmuřtur ve bu arařtırmacılar genellikle nörobiyolojist yada matematikilerden oluřmaktadır (*Wiggins / Looper, 1991, s.1*). Öte yandan Amari, 1970'leri pratikte karanlık periyot olarak adlandırırken teorik alıřmalar için ise o denli karanlık olmadığını belirtmektedir (*Amari, 1994, s. 32*). Ayrıca, bu yıllarda ilginin azlığı yalnızca Minsky ve Papert'ın alıřmalarını ortaya koyduđu olumsuzluklardan deđil aynı zamanda bilgisayar teknolojisinin 60 ve 70'li yıllarda gösterdiđi hızlı gelişimin, arařtırmacıların kafasında sembol işlemlerinin bilgisayarda ok daha hızlı ve güçlü yapılabileceđi olgusunu da yaratmış olmasındandır (*Cheng / Titherington, 1994a, s.51*). Aslında 70'li yıllarda yapılan alıřmalara bakıldığında Amari bir yönüyle de haklıdır. ünkü 80'lerde geliştirilecek birçok yeniliđin temelleri bu yıllarda atılmıştır. 1972'de ađırışmalı bellek üzerine ilk alıřmalar Kohonen tarafından yapılmış, kendi kendini organize edebilen özellik haritalarının (SOM), üzerine ilk alıřmalara 1973'te Von der Malsburg tarafından başlanmıştır. Yine, XOR probleminin özümünü verecek Geriye yayılım algoritması üzerine ilk alıřmalara Werbos tarafından 1974 yılında başlanmıştır. 1976 yılında Grosberg tarafından ortaya konulan Adaptif Rezonans Teorisi ađları ve 1982'de Kohonen'in SOM modelini geliřtirmesi bunu takip etmiştir. Bu dönemden sonra Hopfield'ın 1982 yılındaki alıřmaları sonucunda içeriđi ile adreslenebilen hafıza (Content-addressable Memory) ve 1985 yılında Tank ile beraber yayınladıkları 'Optimizasyon Problemlerinde Kararların Nöral Hesaplamaları' adlı makaleyle yapay sinir ađlarına olan ilgiyi tekrar yavaş yavaş artmaya başlamıştır. Bu alıřmada Hopfield ve Tank yapay sinir ađlarının genelleřtirilebileceđini ve o zaman için geleneksel bilgisayar programlama yaklařımıyla özümlemesi zor olan gezgin satıcı probleminin yapay sinir ađları ile özümlenebileceđini göstermektedir (*Warner / Misra, 1996, s. 286*). Yine aynı yıllarda McClelland, Rumelhart ve PDP (Paralel Dađıtılmış İşlemler) arařtırma grubu Werbos'un başlattıđı alıřmaları geliřtirerek ok katmanlı algılayıcılarda geriye yayılım algoritmasının kullanımını ve bunun sonucunda da XOR probleminin özümünü gerekleřtirmişlerdir. XOR probleminin özümünden sonra dikkatler tekrar yapay sinir ađı alıřmaları üzerine toplanmıştır ve bu nedenle XOR problemi yapay sinir ađı alıřmalarında önemli bir kilometre taşıdır (*Öztemel, 2003, s.76*). Geriye yayılım algoritması ile ok

katmanlı algılayıcılar bir çok alanda uygulama imkanı bulmuştur. 1988 yılında radial temelli fonksiyon modelleri ve 1989'da olasılıksal sinir ağlarından sonra 1991 yılında genelleştirilmiş regresyon sinir ağları modelleri geliştirilmiştir.

1991 yılından günümüze yapılan çalışma ve uygulama sayısı oldukça fazlaştığından konumuzu ilgilendiren kısmıyla yani iktisadi zaman serileri alanında yapılan belli başlı çalışmalara değinerek devam etmek daha uygun olacaktır.

Yapay sinir ağlarının zaman serisi analizinde ciddi bir şekilde kullanıldığı çalışmalardan ilki 1988 yılında White tarafından yapılmıştır. White makalesinde yapay sinir ağlarını etkin piyasa hipotezini test etmek amacıyla kullanmıştır. İlgili yapay sinir ağı modelinde kaç katman kullanılacağı ve katmanlardaki sinir hücrelerinin sayısının ne olacağı gibi araştırmacının bilgi tecrübe ve becerisine bırakılmış yada diğer bir ifade ile subjektivitenin hakim olduğu ve bir takım sorulara tam bir yanıtın olmayışı, bu yıllarda araştırmacıları yapay sinir ağı modellerine biraz soğuk bakmasını beraberinde getirse de, 1989-93 yılları arasında Cybenko (1989), Funahashi (1989), Hornik (1989, 1991, 1993) ve Hornik, Stinchcombe and White (1989, 1990), tarafından yapılan teorik çalışmalar bu anlamdaki şüpheleri gidermiştir (*Kuan, 2006, s.1*). Bu tarihten sonra da finansal ve ekonomik temelli çalışmalar hızla artmıştır. Her bir çalışmaya değinemeyeceğimizden literatürde sıklıkla atıf alan çalışmalar şu şekilde özetlenebilir. Bosargei (1993), White'ın çalışmasının üzerine giderek yapay sinir ağı modelleriyle önemli ölçüde doğrusal dışılık tespit ettiği Standart and Poors 500, Ham petrol, Yen- Dolar paritesi ve Nikkei endekslerini başarılı şekilde tahmin etmiştir. Hisse senetleri piyasaları üzerine benzer çalışmalar ve başarılı tahminler Zeidenberg (1995), Refenes (1995), Heiemstra (1996), ve yine aynı yıllarda Hoefhe (1996) ve Helmestien (1996) tarafından da yapılmıştır. Ramazan Gençay'ın (1994) ileri beslemeli ağların küçük veri setlerinde zaman serilerindeki gürültüyü filtreleme yeteneklerini ölçmek üzere yaptığı çalışma bu alandaki önemli çalışmalar arasında gösterilmektedir. Kaastra (1996) tarım ekonomisi alanında tahminleme de yapay sinir ağlarını kullanmıştır. Cheng (1996) ise ABD

hazine bonolarının getiri tahminlerinde başarılı bir çalışma yapmıştır. Donalson ve Kamstra (1997) çok katmanlı algılayıcıya dayanan yarıparametrik doğrusal dışı GARCH modeli ile hisse senedi getiri volatilitelerini tahmin etmiştir. Franses ve Draisma (1997) mevsimsel bileşenin nasıl ve ne zaman değiştiğini ortaya koymak için yapay sinir ağları ile çalışan bir metot geliştirmişlerdir. Olasılıksal sinir ağları ile iflas tahminleri üzerine Yang (1997) tarafından yapılan başarılı bir çalışmada bunlara eklenebilir. Diğer taraftan, Elkateb (1998) elektrik talebi tahmini çalışması ve Franses ile Griensven (1999) tarafından döviz kurları için al-sat sinyali tahmini amacıyla yaptıkları çalışmalarda literatürdeki önemli çalışmalar arasında gösterilebilir. Kapaetnios (2000), ARCH yapısının testi için sinir ağı modeli kullanmış ve diğer taraftan Heinmann (2000), rasyonel bekleyişlerin uyarlayıcı öğrenilmesi üzerine sinir ağı modelleri geliştirmiştir. Qi (2001), resesyon ve konjonktür tahmininde yapay sinir ağı modeli kullanmıştır. Tkacz (2001), milli gelir tahmininde, Chen (2001) enflasyon tahmininde, Tseng (2001), sanayi üretim endeksi tahmininde yapay sinir ağı modellerini kullandığı görülmektedir. Bu değinilen çalışmalar yapay sinir ağlarının finansal ve ekonomik temelli çalışmalarda da güvenle kullanılabilceğini kanıtladıklarından dolayı 2001 yılı sonrasında yapay sinir ağı çalışmaları ekonomi biliminde de hızla artmış, günümüzde yapay sinir ağları ekonomi ve finans konularında kullanışlı bir istatistiksel araç olarak görülmeye başlanmıştır. Sonuçta, McClelland'ın yapay sinir ağları geldi – gitti, şimdi yine aramızda ve kalacağı benziyor yakıştırması (McClelland, 1994, s.42) gerçekleşmiştir.

Bu anlamda yapay sinir ağlarının gittikçe artan bir oranda ekonomi ve finans alanında kullanılmasının arkasında yatan başlıca iki önemli neden belirtilebilir. Bunlarda birincisi, regresyon yaklaşımının aksine yapay sinir ağlarının tahmin edilecek parametrelere doğrusallık ile ilgili herhangi bir kısıt getirmemesidir. İkinci önemli neden ise, yapay sinir ağı modellerinin yine regresyonun aksine, veri kümesinin sahip olduğu dağılım üzerine herhangi bir varsayımda bulunmamasıdır. Böylece daha yüksek dereceden karmaşıklığın yakalanmasında modelleme aracı çok daha esnek bir yapıdadır. Yapay sinir ağlarının özellikle ekonomik önraporlama alanında kullanımında, parametrelere

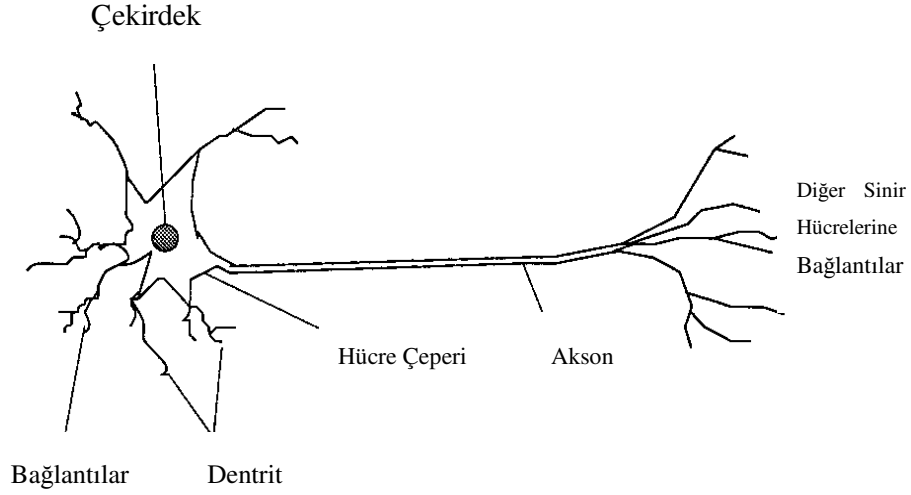
doğrusallıkla ilgili herhangi bir kısıt getirilmemesi ve veri setinin dağılımına ilişkin herhangi bir varsayımda bulunulmaması regresyon modellerine karşılık yapay sinir ağı modellerinin açık bir avantajıdır (*Tal / Nazareth, 1995, s.71*).

## **II. YAPAY SİNİR AĞLARINDA TERMİNOLOJİ VE TEMEL MODELLER**

Yapay sinir ağlarının kendine özgü terminolojisi ve özellikle öğrenmelerini öğretmenli öğrenme stratejisine göre gerçekleştiren ağların yapısının anlaşılabilmesi için temel modellerin işleyişinin ortaya konulması gerekmektedir. Bu amaçla bu başlık altında sırasıyla McCulloch – Pitts yapay sinir hücresi, Tek Katmanlı Algılayıcı ve ADALİNE ünitelerinin, çalışma ve öğrenme prensipleri ile yapay sinir ağları alanının kendine özgü terminolojisi birlikte sunulacaktır. Öte yandan temel modeller olarak tanımlanabilecek bu modellere geçmeden önce, yapay sinir ağlarında temel yapı ve elemanların sunulması gerekmektedir.

### **A. YAPAY SİNİR AĞLARINDA TEMEL YAPI VE ELEMANLAR**

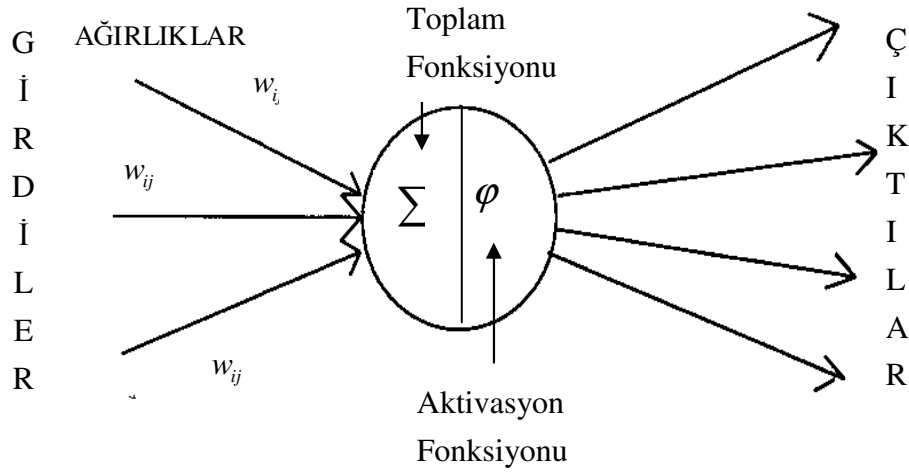
Önceki başlık altında belirtildiği gibi yapay sinir ağları insan beyninin işleyişinden esinlenerek ortaya konulmuş modellerdir. İnsan beyni sinir hücrelerinin ve bu hücreler arasındaki bağlantıların işleyişiyle çalışmasını gerçekleştirir. İnsan beyninin avantajı, yüksek derecede paralel hesaplama yapabilen bir yapıya sahip olma ve bununla ortaya çıkan bilgi işleme yeteneğidir (*Abraham, 2005, s.901*). Bunun yanında elbette insan beyninin nasıl çalıştığına cevabı tek bir sinir hücresinde bulunamaz. Ancak yapay sinir ağları alanında kullanılan terminolojinin büyük bir kısmı biyolojiden alınmıştır (*Francis, 2001, s56*). Bu nedenle tek bir sinir hücresinden başlayarak devam edilmesi önem arz etmektedir. Aşağıdaki şekilde tek bir biyolojik sinir hücresinin yapısı görülmektedir



**Şekil 4: Biyolojik Sinir Hücresinin Yapısı**

Biyolojik bir sinir hücresinin yapısı yukarıda şekil 4’te ana hatlarıyla gösterilmiştir. Şekilden de anlaşılacağı gibi, bir biyolojik sinir hücresinde başlıca dört ana eleman vardır. Bunlar Dentrit, Çekirdek, Akson ve Bağlantı elemanlarıdır. Biyolojik sinir hücresinin işleyişi şu şekilde gerçekleşmektedir: Bir hücreden diğerine bağlantı elemanlarıyla gelen elektrik sinyali, girdi kanalları olarak adlandırılabilir dentrit tarafından alınır. Alınan sinyal çekirdek tarafından işlenerek bir çıktıya dönüştürülür. Bu yeni elektrik sinyali akson yardımıyla dentritlere ve oradan da bağlantı elemanlarına geçirilir. Böylece bir sinir hücresinin çıktısı bu son bağlantı elemanlarının sayesinde başka bir sinir hücresinin girdisi olacak şekilde yol alır. Bu işleyiş algılamayı yapacak ağ boyunca paralel olarak tekrarlanır. Basitçe anlatılan bu biyolojik işlemler bütünü yapay sinir ağları modellerinin geliştirilmesine dayanak oluşturmuştur (Roberts, 1989, s.481).

Yapay bir sinir hücresinin yapısı ise aşağıda şekil 5’de görülmektedir. Yapay sinir hücresinde başlıca beş eleman vardır. Bunlar, girdiler, ağırlıklar, toplama fonksiyonu, transfer fonksiyonu ve çıktıdır.



**Şekil 5: Yapay Sinir Hücresinin Yapısı**

**i. Girdiler** : Girdiler yapay sinir hücresine dışarıdan alınan bilgilerdir. Bu girdi, kullanılan yapay sinir ağı modeline ve yapay sinir hücresinin bulunduğu katmana göre örneklemeden, başka bir sinir hücresinden veya yapay sinir hücresinin kendisinden gelebilir. Bir kümenin kesikli değerleri yada gerçek sayı olabilirler (Shachmurove, 2004, s.12).

**ii. Ağırlıklar** : Biyolojik sinir hücresindeki bağlantı elemanlarının yapay sinir hücresindeki karşılığıdır.  $W_{ij}$  gösterimine göre, j. hücre yada girdi ile i. hücre arasındaki bağlantının kuvvetini göstermektedir. Öte yandan ağırlık değerinin sayısal olarak büyük yada küçük veya – yada + olması bu bağlantının önemli yada önemsiz olduğu anlamına gelmeyeceği gibi, bazı durumlarda bu değerın sıfır olması yapay sinir ağı için en önemli olay da olabilir (Öztemel 2003, s.49).

**iii. Toplam Fonksiyonu** : Toplam fonksiyonunun amacı, hücreye gelen net girdi değerinin hesaplanması şeklinde özetlenebilir. Değişik toplam fonksiyonları kullanılabileceği gibi, en basit toplam fonksiyonu şu şekildedir.

$$\text{Net Girdi}_i = \sum_{j=1} w_{ij} x_j$$

Bu ifadenin anlamı, her bir j. yapay sinir hücresinden gelen çıktı değerinin ( $x_j = y_j = \text{girdi}_i$ )  $w_{ij}$  ağırlığı ile ağırlıklandırılıp toplamlarının alınmasıdır.

Öte yandan ağırlıklandırma işleminden sonra her bir değer çarpımının alınması yada yine ağırlıklandırma işleminden sonra en büyük girdinin veya en küçük girdinin net girdi olarak tek başına seçilmesi, pozitif olanların yada negatif olanların net girdi olarak alınması gibi çalışan toplam fonksiyonlarının da kullanıldığı yapay sinir hücreleri vardır (*Öztemel, 2003, s.50*).

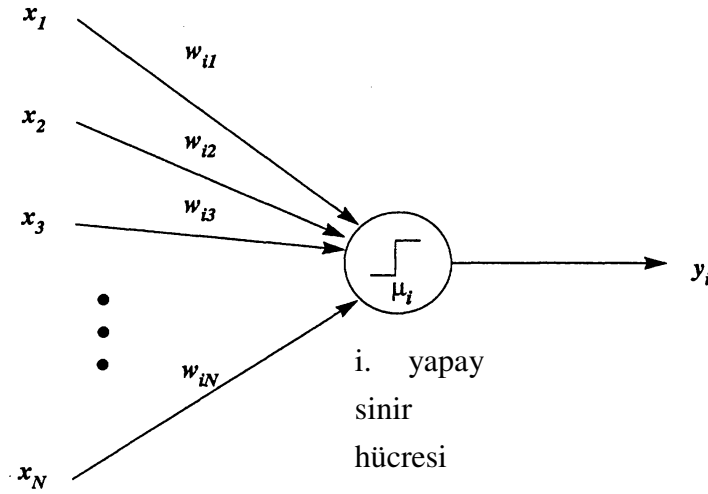
**iv. Aktivasyon Fonksiyonu** : Biyolojik sinir hücresinin çekirdeğinde girdiyi işleyip çıktıya dönüştürmesi işlemi, yapay sinir hücresinde herhangi bir toplam fonksiyonu ile oluşturulan net girdinin bir aktivasyon fonksiyonundan geçirilip çıktıya dönüştürülmesi şeklinde gerçekleştirilir. Araştırmacılar tarafından bugüne kadar çok sayıda farklı fonksiyon aktivasyon fonksiyonu olarak kullanılmış olsa da en fazla tercih edilen aktivasyon fonksiyonları, eşik değer ve lojistik fonksiyonlardır (Warner / Mısra, 1996, s.287). Aslında yapay sinir ağının potansiyel hesaplama gücünü ortaya koymak için doğrusal olmayan aktivasyon fonksiyonu kullanmak gerekmektedir (*Gonzalez / Castro, 2001, s.5*). Bunun yanı sıra kullanılan öğrenme algoritmasına göre belirli bazı aktivasyon fonksiyonlarının seçilmesini gerektirecek durumlar söz konusu olabilir. Örneğin daha sonra ayrıntılı şekilde tartışılacak geriye yayılma algoritmasının kullanıldığı durumda, algoritmanın hesaplama yaklaşımından dolayı aktivasyon fonksiyonunun türevlenebilir olması gerekmektedir.



**v. Çıktılar** : Aktivasyon fonksiyonundan elde edilen değer yapay sinir hücresinin çıktısını oluşturur. Kullanılan yapay sinir ağı modeline göre bu çıktı bir yada birden fazla sayıda olabilir (*Cheng ve Titterington, 1994a, s.7*).

## B. McCULLOCH - PITTS MODELİ

İnsan beyninin modellenmesi açısından ilk çalışmaların sinir hücresi modellemesi üzerine yapıldığı daha önce belirtilmişti. Bu alandaki ilk başarılı çalışmanın 1943 yılında McCulloch ve Pitts tarafından yapıldığı kabul edilir ve oluşturdukları model basit bir ikili eşik değer ünitesi içermektedir (Hetz / Palmer / Krogh, 1991, s.1). Aşağıda şekil 6'da McCulloch-Pitts modeli görülmektedir.



**Şekil 6: McCulloch – Pitts Yapay Sinir Hücresi**

McCulloch ve Pitts modelinin matematiksel formülasyonu şu şekilde verilebilir;

$$y_i = \varphi \left( \sum_j (w_{ij} x_j - \mu_i) \right)$$

Burada,

$y_i$  , i. yapay sinir hücresinin çıktısıdır.

$w_{ij}$  j. yapay sinir hücresinden i. yapay sinir hücresine gelen girdi için ağırlık

$x_i$  , i. yapay sinir hücresinin girdisi (aynı zamanda bir önceki katmandaki j. yapay sinir hücresinin çıktısı,  $y_j$  yada  $çıkıtı_j$ )

$\mu_i$ , i. yapay sinir hücresinin eşik değeri

$\phi$  , ise aktivasyon fonksiyonudur.

McCulloch ve Pitts modelinde aktivasyon fonksiyonu,

$$\phi(\text{net girdi}) = \begin{cases} 1 & , \quad \text{netgirdi} \geq 0 \text{ ise} \\ 0 & , \quad \text{diğer durumlarda} \end{cases}$$

şeklinde bir adım fonksiyonudur.

Bu modelin çalışması şu şekilde gerçekleşir;

1) Girdiler toplam fonksiyonu yardımıyla,  $\text{Net Girdi}_i = \sum_{j=1} w_{ij} \text{çıkıtı}_j$  olarak düzenlenir.

2) Eğer bulunan net girdi değeri eşik değerden büyükse,  $(\sum_j w_{ij} x_i \geq \mu_i)$

aktivasyon fonksiyonuna sıfırdan büyük bir değer gider, sıfırdan büyük bir değer de kullanılan bu adimsal aktivasyon fonksiyonundan geçirildiğinde çıktı 1 olarak bulunur. Bu durum yapay sinir hücresinin aktif olduğu anlamına gelmektedir. Ayrıca yapay sinir hücresinin ateş ettiği (fire) şeklinde de adlandırılmaktadır (Warner / Mısra, 1996, s. 285).

3) Eğer bulunan net girdi eşik değerden küçükse,  $(\sum_j w_{ij}x_i < \mu_i)$  aktivasyon fonksiyonunda işlenecek değer sıfırdan küçüktür ve bu durumda aktivasyon fonksiyonu 0 değerini üretir.

Bilgisayarlarda yapılan hesaplamalar bir küme McCulloch – Pitts yapay sinir hücresi ile başarılabilir (Abu- Mostafa 1986, s.6).

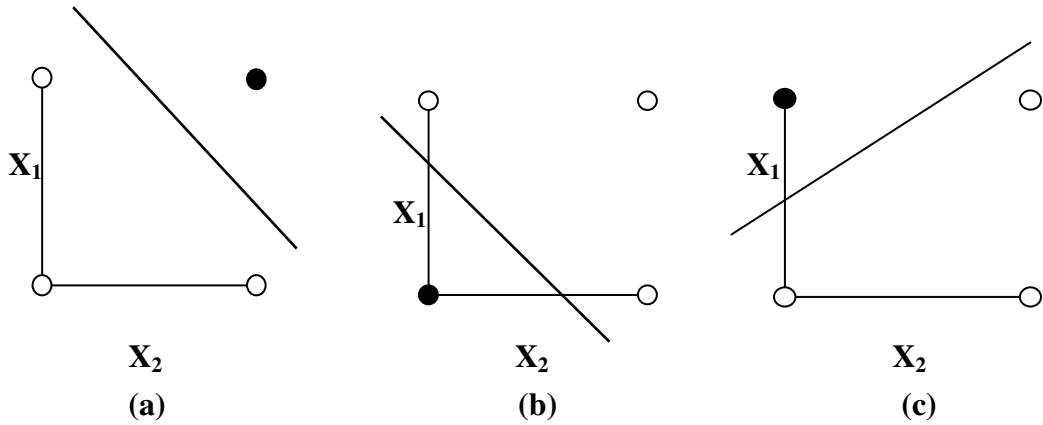
### C. BASİT ALGILAYICI MODELİ

Daha önce de değinildiği gibi Basit algılayıcı modeli 1958 yılında Rosenblatt tarafından geliştirilmiştir. Modelin temel amacının örüntü tanıma yada sınıflandırma olduğu söylenebilir. Bu ağda kullanılan yapay sinir hücreleri aynen McClulloch-Pitts modelindeki yapay sinir hücrelerine benzer hücrelerdir (*Hagan / Demuth, 1999, s.4-2*) Rosenblatt tarafından birçok çeşidinin ilk kez dizayn edilip kullanıldığı bu modelin en basit hali, sunulan girdi vektörüne karşılık doğru çıktı vektörünün, hataların azaltılması yönünde ağırlıkların güncellenerek eğitimi ile elde edildiği tek katmanlı ve tek hücreden oluşan ağdır ve ismi de algı anlamına gelen “perception” kelimesinden gelmektedir (*Veelentruf, 1995, s.10*). Bu şekildeki bir basit algılayıcının öğrenme algoritması bir örnek üzerinde açıklanacaktır. Ancak öncelikle daha önce tarihçe bölümünde de değinilen ve yapay sinir ağı çalışmalarına sekte vuran Minsky ve Papert (1969) yayınladıkları ‘Perceptons’ adlı çalışma ve bu çalışma sonucu değinilen problemi hatırlamada yarar vardır: Rosenblatt’a göre “basit algılayıcılar her türlü örüntü tanıma problemi için sonlu sayıda iterasyonla çözüm oluşturacak uygun bir araçtır”. Ancak çözebildikleri problemler için hızlı ve güvenilir ağlar olmalarına rağmen, XOR problemi olarak adlandırılan ve doğrusal olmayan bir ilişkiyi ortaya koyan problemi çözememektedirler. Burada XOR problemini tanımlamadan önce, ilk olarak daha basit ve *doğrusal* bazı örüntü yapılarını ortaya koymak gerekmektedir.

Mantık literatüründeki, VE, VEYA, VE DEĞİL problemlerinin tablo olarak gösterimi aşağıda Tablo 2’deki gibidir.

**Tablo 2: “VE”, “VEYA”, “VE DEĞİL” Problemleri**

<u>VE</u>			<u>VEYA</u>			<u>VE DEĞİL</u>		
$X_1$	$X_2$	Y	$X_1$	$X_2$	Y	$X_1$	$X_2$	Y
1	1	1	1	1	1	1	1	0
1	0	0	1	0	1	1	0	1
0	1	0	0	1	1	0	1	0
0	0	0	0	0	0	0	0	0



**Şekil 7: (a) Ve - (b) Veya - (c) Ve değil Problemleri**

Şekil 7’de birinci sütunda, verilen  $X_1$  ve  $X_2$  değerlerinin “ve” işlevi ile işleme alınmasıyla ilgili Y değerleri oluşmaktadır. Aynı mantık ikinci ve üçüncü sütun içinde geçerli olup, ikici sütundaki  $X_1$  ve  $X_2$  değerleri “veya”, üçüncü sütundaki  $X_1$  ve  $X_2$  değerleri ise “ve değil” işlevi ile işleme alınmış ve ilgili Y değerleri elde edilmiştir. Şekil 7 (a), (b) ve (c)’de Y değerlerinden farklı olanı içi dolu olarak gösterilmiştir. Şekil 7’den de anlaşılacağı gibi bu problemlerde bir

dođru fonksiyonu yardımıyla sınıflandırma yapılabilir.

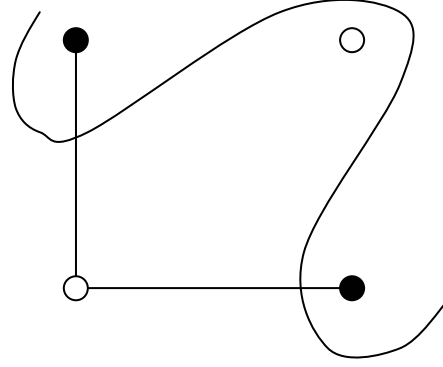
Basit algılayıcılar Rosenblatt tarafından geliştirilen algoritma sayesinde bu dođru fonksiyonlarını oluşturabilmekte ve böylece bu problemlere çözüm getirebilmektedirler.

XOR problemini ise ařađıdaki gibi tanımlanabilir.

### XOR

$X_1$	$X_2$	$Y$
1	1	0
1	0	1
0	1	1
0	0	0

*Tablo 3: XOR Problemi*

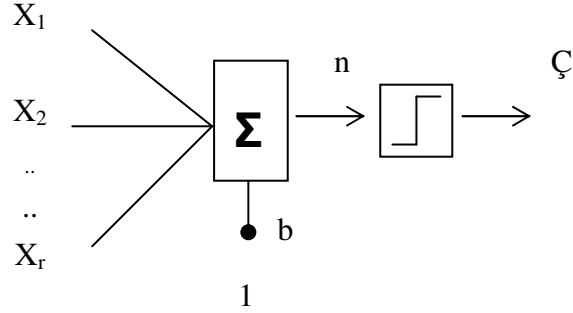


*Şekil 8: XOR Problemi*

Şekil 8'te  $Y$  deđeri 0 ise ii boş ve 1 ise ii dolu olarak gsterilmiřtir. Şekilden de anlaşılabilir gibi, bu durumda sınıflandırma yapabilmek iin bir dođru fonksiyonunun oluşturulması iře yaramamaktadır. XOR problemi dođrusal olmayan bir fonksiyon yardımıyla sınıflandırma yapılarak çzlebilir. Rosenblatt tarafından geliştirilen algoritma böyle bir fonksiyon oluşturamamaktadır. XOR problemi basite bu şekilde özetlenebilir.

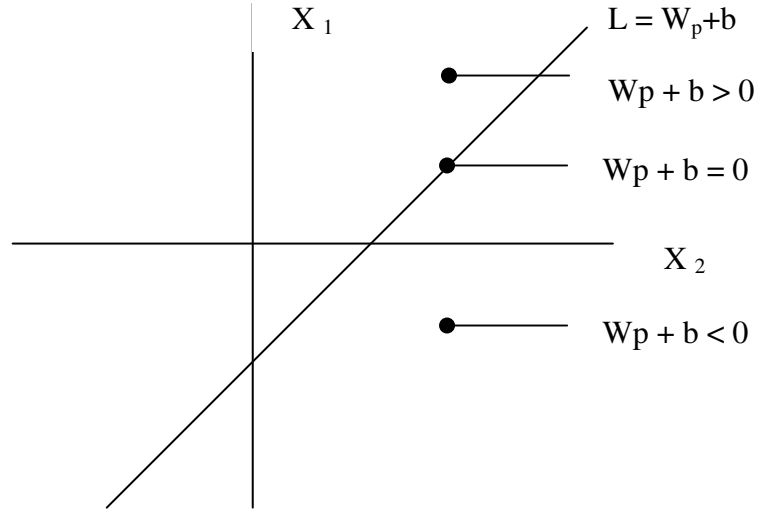
XOR problemi ve bu problemle ilgili olarak dođrusal olmayan sınıflandırma problemlerini bir tarafa bırakırsak, basit algılayıcılar çzebildikleri problemler iin hızlı ve güvenilir ađlardır (*Hagan, 1995, s.4-5.*). Bunlara ek olarak basit algılayıcıların alıřma prensibinin anlaşılması, ikinci blmde ele alınacak

daha karmaşık ağ yapılarının anlaşılabilmesi için önemlidir. Eşik değer transfer fonksiyonu kullanan bir basit algılayıcı modeli aşağıda şekil 9’da gösterildiği gibidir;



**Şekil 9 : Basit Algılayıcı Modeli**

Basit algılayıcı modelinde her bir girdi değeri uygun bir  $w_r$  ağırlığı ile ağırlıklandırılarak eşik değer fonksiyonuna yollanır. Eşik değer fonksiyonu eğer net girdi 0’den büyükse 1, aksi durumda 0 değeri üretir. Aslında eşik değer fonksiyonu girdi uzayını iki bölgeye ayırarak basit algılayıcıya girdi vektörlerini sınıflandırma kabiliyeti vermektedir. Aşağıdaki Şekil 10’da basit algılayıcı tarafından çözüm getirilen teorik bir sınıflandırma problemi sunulmuştur.



**Şekil 10 : Teorik Bir Sınıflandırma Problemi Çözümü**

Şekil 10'da iki sınıflandırma alanı basit algılayıcının ürettiği sınıflandırma fonksiyonu olan karar sınır çizgisi (L) ile belirlenmektedir.  $W_p + b$  şeklinde bir fonksiyon olan karar sınır çizgisi b yan değerine göre kayar. Karar sınır çizgisi yan ve ağırlık değerlerinin değişmesiyle çözüm alanında her yere kayabilir. b yan değeri olmayan bir karar sınır çizgisi orijinden geçer. Karar sınır çizgisi fonksiyonuna b şeklinde bir yan değerinin eklenmesi, orijinin aynı tarafında olan iki girdi vektörünün sınıflandırılması gereken problemlere çözüm getirmek içindir.

Basit algılayıcı öğrenme kuralı olarak bilinen, Rosenblatt tarafından geliştirilmiş ve tek katmanlı basit algılayıcıların eğitiminde kullanılan algoritma şu şekildedir.

- 1) Basit algılayıcı ünitesine gelen net girdi,  $\sum_{i=1}^r w_i x_i$  şeklinde hesaplanır. Burada  $w_i$  ağırlık değeri araştırmacı tarafından subjektif olarak belirlenir.

2) Net girdi değeri eşik fonksiyonundan geçirilir ve çıktı bulunur.

$$\varphi(\text{net girdi}) = \begin{cases} 1 & , \quad \text{netgirdi} > 0 \text{ ise} \\ 0 & , \quad \text{diğer durumlarda} \end{cases}$$

3) Eğer elde edilen çıktı değeri ile beklenen çıktı değeri aynı ise sorun yoktur. Ancak çıktı değeri ile beklenen çıktı değeri farklı ise bu farklılık iki şekilde ortaya çıkabilir.

a) Beklenen çıktı 0 iken ağın ürettiği değer 1 dir. Bu durumda başlangıçta araştırmacı tarafından atanan ilk ağırlık değeri azaltılmalıdır. Bu azaltma  $w_y = w_i - \eta x$  şeklinde olacaktır. Buradaki  $\eta$  değeri **öğrenme katsayısı** olarak adlandırılır ve sabit bir değer olarak araştırmacı tarafından subjektif belirlenir.

b) Beklenen çıktı 1 iken ağın ürettiği değer 0 dir. Bu durumda başlangıçta araştırmacı tarafından atanan ilk ağırlık değeri artırılmalıdır. Bu artırma  $w_y = w_i + \eta x$  şeklinde olacaktır.

4) Bu işlemler tüm örnekler doğru sınıflandırılana kadar tekrarlanır.

Son olarak mantık literatüründeki VE probleminin basit algılayıcı öğrenme algoritmasıyla nasıl çözüldüğünün ortaya koyulması yararlı olacaktır.

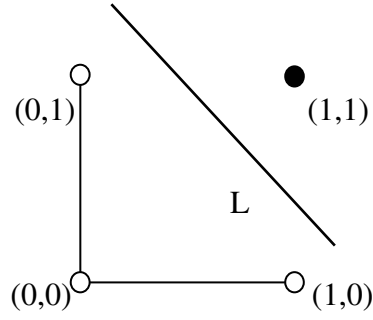
### VE

X <sub>1</sub>	X <sub>2</sub>	Y
1	1	1
1	0	0
0	1	0
0	0	0

**Tablo 4: “VE” Problemi**

Şekil 7- (a) bir kez daha hatırlanacak olursa, X<sub>1</sub> ve X<sub>2</sub> değerlerinin “ve” işlevi ile işleme alınmasıyla oluşan Y değerleri yukarıdaki tablo 4’deki gibidir. Bu Y değerlerinden farklı olanın içi dolu olarak gösterilmesiyle oluşan sınıflandırma problemi de aşağıdaki gibidir.





**Şekil 11: “VE” Problemi**

Şekildeki L doğrusu karar sınır çizgisi olarak bilinmekte ve basit algılayıcı öğrenme algoritmasıyla bu doğru fonksiyonuna benzer bir fonksiyon elde edilmek istenmektedir.

Birinci örnek :  $X_1= 1$  ve  $X_2=1$   $Y=1$

İkinci örnek :  $X_1= 1$  ve  $X_2=0$   $Y= 0$

Üçüncü örnek:  $X_1= 0$  ve  $X_2=0$   $Y= 0$

Dördüncü örnek:  $X_1= 0$  ve  $X_2=1$   $Y= 0$

Ağırlıklar :  $w_1= 1$  ve  $w_2= 2$

Eşik değeri :  $\Phi =1$

Öğrenme katsayısı:  $\eta = 0,5$

olarak ele alınırsa;

**- 1. örnek ağı gösterildiğinde;**

$$\text{net girdi} = \sum_{i=1}^2 w_i x_i = 1.1 + 2.1 = 3$$

$3 > \Phi = 1$  olduğundan ağı çıktısı  $\Ç = 1$  dir. Beklenen çıktı değeri de 1 olduğundan sorun yoktur.

**- 2. örnek ağı gösterildiğinde;**

$$\text{net girdi} = 1.1 + 2.0 = 1$$

$1 \leq \Phi = 1$  olduğundan  $\Ç = 0$  ve beklenen çıktıda 0 olduğundan sorun yoktur.

**-3. örnek ağı gösterildiğinde;**

$$\text{net girdi} = 1.0 + 2.0 = 0$$

$0 \leq \Phi = 1$  olduğundan  $\Ç = 0$  ve beklenen çıktıda 0 olduğundan sorun yoktur.

**- 4. örnek ağı gösterildiğinde;**

$$\text{net girdi} = 1.0 + 2.1 = 2$$

$2 > \Phi = 1$  olduğundan  $\Ç = 1$  ancak beklenen çıktı 0 olduğundan ağırlık değerleri değiştirilmelidir. Ağı çıktısı 1 beklenen çıktı 0 olduğundan algoritma gereği  $w_y = w_i - \eta x$  şeklinde bir ayarlama gereklidir.

$$w_{y1} = w_1 - \eta x_1 = 1 - 0,5.0 = 1$$

$$w_{y2} = w_2 - \eta x_1 = 2 - 0,5.1 = 1,5$$

olarak yeni ağırlık değerleri belirlenir.

**- 1. örnek tekrar ağı gösterildiğinde;**

net girdi =  $1.1 + 1,5.1 = 2,5 > \Phi = 1$   $\zeta = 1$  ve beklenen çıktı da 1 olduğundan sorun yoktur.

**- 2. örnek tekrar ağı gösterildiğinde;**

net girdi =  $1.1 + 1,5.0 = 1 \leq \Phi = 1$   $\zeta = 0$  ve beklenen çıktı da 0 olduğundan sorun yoktur.

**- 3. örnek tekrar ağı gösterildiğinde;**

net girdi =  $1.0 + 1,5.0 = 0 \leq \Phi = 1$   $\zeta = 0$  ve beklenen çıktı da 0 olduğundan sorun yoktur.

**- 4. örnek tekrar ağı gösterildiğinde;**

net girdi =  $1.0 + 1,5.1 = 1,5 > \Phi = 1$   $\zeta = 1$  ve beklenen çıktı 1 olduğundan ağırlık değerleri değiştirilmelidir. Ağırlık çıktısı 1 beklenen çıktı 0 olduğundan algoritma gereği  $w_y = w_i - \eta x$  şeklinde bir ayarlama gereklidir.

$$w_{y1} = w_1 - \eta x_1 = 1 - 0,5.0 = 1$$

$$w_{y2} = w_2 - \eta x_1 = 1,5 - 0,5.1 = 0,5$$

**- 1. örnek tekrar ağı gösterildiğinde;**

net girdi =  $1.1 + 0,5.1 = 1,5 > \Phi = 1$   $\zeta = 1$  ve beklenen çıktı da 1 olduğundan sorun yoktur.

**- 2. örnek tekrar ağı gösterildiğinde;**

net girdi =  $1.1 + 0,5.0 = 1 \leq \Phi = 1$   $\zeta = 0$  ve beklenen çıktı da 0 olduğundan sorun yoktur.

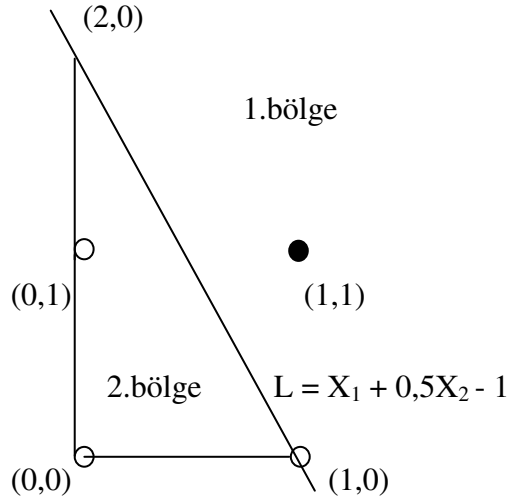
**- 3. örnek tekrar ağı gösterildiğinde;**

net girdi =  $1.0 + 0,5.0 = 0 \leq \Phi = 1$   $\zeta = 0$  ve beklenen çıktı da 0 olduğundan sorun yoktur.

**- 4. örnek tekrar ağı gösterildiğinde;**

net girdi =  $1.0 + 0,5.1 = 0,5 \leq \Phi = 1$   $\zeta = 0$  ve beklenen çıktı da 0 olduğundan sorun yoktur.

Sonuçta, basit algılayıcı öğrenme algoritmasıyla  $w_1=1$  ve  $w_2=0,5$  olarak ağırlık değerleri elde edilmiş ve tüm örnekler doğru sınıflandırılarak öğrenme tamamlanmıştır. Bulunan fonksiyon,  $L = X_1 + 0,5X_2 - 1$  şeklinde olup karar sınır çizgisini oluşturur. Çözüm aşağıda şekil 12’de sunulmuştur.



**Şekil 12 : “VE” Probleminin Basit Algılayıcı İle Çözümü**

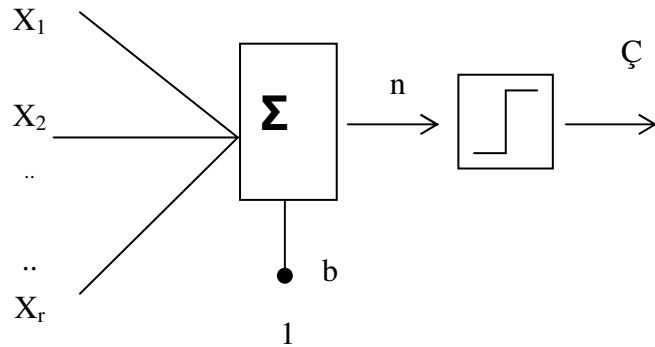
(2,0) noktasıyla  $X_1$  ve (1,0) noktasıyla  $X_2$  eksenlerini kesen karar sınır çizgisi (L) içi dolu olarak gösterilen farklı Y değerini içi boş olarak gösterilen diğer Y değerlerinden ayırmıştır. Bu arada, karar sınır çizgisinin üzerinden geçtiği

(1,0) noktasının kullanılan eşik değer fonksiyonunun özelliğinden dolayı ikinci bölgede yer aldığı belirtilmesinde yarar vardır.

Son olarak doğrusal olmayan problemlerin çözümünde başarısız olmasının yanı sıra bazı durumlarda bir girdi vektörünün, ağırlıkları aşağıya, bir diğerinin de yukarıya çekmesi, öğrenmenin unutulması şeklinde yorumlanmakta ve bu sorunda doğru sınıflandırmayı verecek fonksiyonun bulunması için iterasyon sayısını arttırabilmektedir (Öztemel 2003, s.68). Ancak, eğer doğrusal bir çözüm varsa ve ağırlıkların güncellenmesi sadece yanlış sınıflandırma sonucu yapılıyorsa basit algılayıcı modelinin sonlu sayıda iterasyonla bu çözüme yakınsayacağı rahatlıkla söylenebilir. Bunlarında ötesinde, Engel ve deBroeck (2001) herşeye rağmen basit algılayıcının çok ilginç ve değerli bir model sistemi olduğunu ve öğrenmenin istatistiksel mekaniğinde bir hidrojen atomu görevini üstlendiğini belirtmektedirler (Engel / deBoreck, 2001, s.4).

#### D. ADAPTİF DOĞRUSAL ELEMAN AĞI

Rosenblatt'ın basit algılayıcı modelini geliştirdiği yıllarda ondan tamamen bağımsız bir şekilde fakat aynı temellerle dayanan Adaline modeli Widrow ve Hoff tarafından oluşturulmuştur. Aşağıda şekil 10'da da görüldüğü gibi Adaline yapısı, basit algılayıcı modeli ile aynıdır.



Şekil 13 : ADALINE Ünitesi

Adaline ve basit algılayıcı arasındaki temel fark kullandıkları öğrenme algoritmasının farklı olmasıdır. Öte yandan adından da anlaşılacağı gibi (**AD**Aptive **LIN**er **E**lement) Adaline doğrusal sınıflandırma problemlerini çözmek amacıyla geliştirilmiştir. Bu nedenle lineer aktivasyon fonksiyonu kullanmaktadır. Adaline da kullanılan öğrenme kuralı Widrow- Hoff öğrenme kuralı, Delta öğrenme kuralı yada En küçük kareler öğrenme kuralı olarak adlandırılmaktadır. Kural, ortalama kare hatayı minimize etme prensibine göre işlemektedir, bu da karar sınırını örneklemden elde edilen örüntüye mümkün olduğunca yakınlaştırmakta ve basit algılayıcı öğrenme kuralına göre daha güçlü bir öğrenme kuralı olarak yorumlanmasını sağlamaktadır (*Haykin, 1999, s.118*).

Hata ( $e$ ) ve ortalama kare hata da ( $\varepsilon = \frac{1}{2}e^2$ ) şeklinde ifade edilirse, ve beklenen çıktı ( $d$ ) ağıın ürettiği çıktı da ( $y$ ) ise,

$n$ . iterasyonda ,  $e(n) = d(n) - y(n)$  olacaktır. Burada,  $n$ . İterasyonda ağıın ürettiği çıktı  $y = \sum w(n)x(n)$  şeklindedir. ortalama kare hatanın ağırlık vektörüne göre türevi,

$$\frac{\partial \varepsilon}{\partial w} = e(n) \frac{\partial e(n)}{\partial w}$$

olacaktır.

Öte yandan  $\frac{\partial e(n)}{\partial w}$  ise,  $e(n) = d(n) - \sum w(n)x(n)$  olduğundan dolayı,

$$\frac{\partial e(n)}{\partial w} = -x(n)$$

şeklindedir.

Bu durumda gradient vektörü,  $g(n) = -x(n)e(n)$  olarak bulunmuş olur. Böylece, (n+1). İterasyona başlamadan önce ağırlıkların güncellenmesi

$$w(n+1) = w(n) + \eta x(n)e(n)$$

şeklinde formüle edilebilir.

İyi bir yakınsama sağlanacağından emin olmak için, iteratif bir çözüm arayışı ile tüm örneklerin hesaplamalarının doğru olduğu durum belirlenmeye çalışılır (*Wendemuth, 2006, s.3*).

Bu durumda tek yapay sinir hücresi kullanan ve aktivasyon fonksiyonu olarak eşik değer fonksiyonu kullanan bir Adaline ünitesinin öğrenme kuralı algoritması şu şekilde yazılabilir:

1) Adaline ünitesine gelen net girdi,  $\sum_{i=1}^r w_i x_i + \phi$  şeklinde hesaplanır.

Burada  $w_i$  ağırlık değeri ve  $\phi$  eşik değeri başlangıçta araştırmacı tarafından subjektif olarak belirlenir.

2) Hesaplanan net değer,  $\sum_{i=1}^r w_i x_i + \phi \geq 0$  ise çıktı 1 olarak,  $\sum_{i=1}^r w_i x_i + \phi < 0$

ise çıktı -1 olarak belirlenir.

3) Eğer beklenen çıktı ile gerçekleşen çıktı değeri arasında fark varsa bu hata, (e) hata, (d) beklenen çıktı ve (y) Adaline ünitesinin ürettiği çıktı olmak üzere  $e = d - y$  şeklinde elde edilir.

4) Amaç, bu hatayı minimize edecek ağırlıkların elde edilmesi olduğundan, yeni ağırlık değerleri aşağıdaki şekilde bulunur.

$$W_y = W_1 + \eta e x_i$$

Burada  $\eta$  değeri öğrenme katsayısı olup yine araştırmacı tarafından

subjektif olarak belirlenir.

5) Net girdi hesaplanırken kullanılan eşik değerde aşağıdaki şekilde yenilenir.

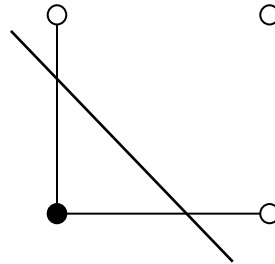
$$\phi_y = \phi + \eta e$$

Bir önceki başlık altında “VE” problemi basit algılayıcı ile çözülmüştü, burada da “VEYA” problemi Adaline ile çözülecektir.

### VEYA

$X_1$	$X_2$	$Y$
1	1	1
1	0	1
0	1	1
0	0	0

*Tablo 5 : “VEYA” Problemi*



*Şekil 14: “VEYA” Problemi*

Örnek değerleri Tablo 5 ve geometrik gösterimi de Şekil 14’de verilen problemin çözümü Adaline ile şu şekilde yapılabilir.

Birinci örnek :  $X_1=1$  ve  $X_2=1$   $Y=1$

İkinci örnek :  $X_1=1$  ve  $X_2=0$   $Y=1$



Üçüncü örnek:  $X_1=0$  ve  $X_2=1$   $Y=1$

Dördüncü örnek:  $X_1=0$  ve  $X_2=0$   $Y=0$

Ağırlıklar :  $w_1=1$  ve  $w_2=2$

Eşik değeri :  $\Phi = 0,5$

Öğrenme katsayısı:  $\eta = 0,5$

olarak ele alınırsa,

**- 1. örnek ağı gösterildiğinde;**

$$NET = [1 \ 2] \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0,5$$

$NET = 1+2+0,5 = 3,5 > 0$  olduğundan  $y = 1$  olacaktır ve bu değeri beklenen çıktı ile aynıdır.

**- 2. örnek ağı gösterildiğinde;**

$$NET = [1 \ 2] \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0,5$$

$NET = 1 + 0,5 = 1,5 > 0$  olduğundan  $y = 1$  olacaktır ve bu değeri beklenen çıktı ile aynıdır.

**- 3. örnek ağı gösterildiğinde;**

$$NET = [1 \ 2] \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 0,5$$

$NET = 2 + 0,5 = 2,5 > 0$  olduğundan  $y = 1$  olacaktır ve bu değeri beklenen çıktı ile aynıdır.

**- 4. örnek ağı gösterildiğinde;**

$$NET = [1 \ 2] \begin{bmatrix} 0 \\ 0 \end{bmatrix} + 0,5$$

NET = 0,5 > 0 olduğundan y = 1 olacaktır. Ancak beklenen çıktı -1 olduğu için ağırlık değerleri ve eşik değeri yenilenmelidir. Bu işlemler için,

$$W_y = W_1 + \eta e x_i \quad \text{ve} \quad \phi_y = \phi + \eta e \quad \text{formülleri kullanılacaktır.}$$

Önce ağırlıklar için,

$$W_y = W_1 + \alpha e x_i$$

$$W_y = [1,2] + 0,5 (-1,-1) [0,0]$$

$W_y = [1, 2]$  olarak bulunur burada ağırlık değerleri değişmemiştir.

Şimdi de eşik değeri için,

$$\phi_y = \phi + \eta e$$

$$\phi_y = 0,5 - 1$$

$\phi_y = -0,5$  değeri bulunur.

**- 1. örnek ağı tekrar gösterildiğinde;**

$$NET = [1 \ 2] \begin{bmatrix} 1 \\ 1 \end{bmatrix} - 0,5$$

NET = 1+2+0,5 = 2,5 > 0 olduğundan y = 1 olacaktır ve bu değer beklenen çıktı ile aynıdır.

**- 2. örnek ağı tekrar gösterildiğinde;**

$$NET = [1 \ 2] \begin{bmatrix} 1 \\ 0 \end{bmatrix} - 0,5$$

$NET = 1 + 0 - 0,5 = 0,5 > 0$  olduğundan  $y = 1$  olacaktır ve bu değer beklenen çıktı ile aynıdır.

**- 3. örnek ağı tekrar gösterildiğinde;**

$$NET = [1 \ 2] \begin{bmatrix} 0 \\ 1 \end{bmatrix} - 0,5$$

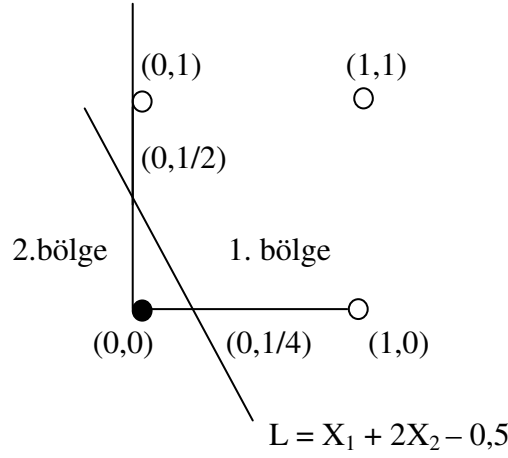
$NET = 0 + 2 - 0,5 = 1,5 > 0$  olduğundan  $y = 1$  olacaktır ve bu değer beklenen çıktı ile aynıdır.

**- 4. örnek ağı tekrar gösterildiğinde;**

$$NET = [1 \ 2] \begin{bmatrix} 0 \\ 0 \end{bmatrix} - 0,5$$

$NET = 0 + 0 - 0,5 = -0,5 < 0$  olduğundan  $y = 0$  olacaktır ve bu değer beklenen çıktı ile aynıdır.

Sonuçta, Adaline öğrenme algoritmasıyla  $w_1=1$  ve  $w_2=2$  olarak ağırlık değerleri ve elde edilmiş ve tüm örnekler doğru sınıflandırılarak öğrenme tamamlanmıştır. Bulunan fonksiyon,  $L = X_1 + 2X_2 - 0,5$  şeklinde olup karar sınır çizgisini oluşturur. Aşağıda şekil 15’de çözüm sunulmuştur.



**Şekil 15 : “VEYA” Probleminin ADALINE İle Çözümü**

Şekilden de anlaşılacağı gibi  $X_1=0$  ,  $X_2=0$  değerlerinin “veya” işlevi ile işleme alınmasıyla üretilen  $Y = 1$  değeri, diğer durumlarda elde edilen  $Y = 0$  değerlerinden,  $L = X_1 + 2X_2 - 0,5$  doğrusu ile ayrılmıştır.

Öğrenme oranı  $\eta$  için araştırmacı tarafından subjektif olarak belirleneceği daha önce belirtilmişti. Ancak Haykin (1996), öğrenme oranının belirlenmesi için aşağıdaki ölçütü sunmaktadır.

$$0 < \eta < \frac{2}{\lambda_{\max}}$$

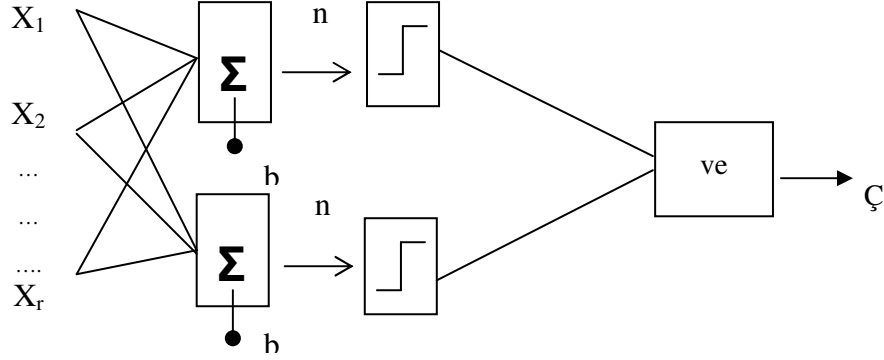
burada,  $\lambda_{\max}$  girdilerin korelasyon matrisi  $R_x$ ’in en büyük özköküdür. Eğer bu değer hesaplanamıyorsa, bu durumda  $\text{iz}(R_x)$  kullanılarak,

$$0 < \eta < \frac{2}{\text{iz}(R_x)}$$

şeklinde öğrenme oranı belirlenebilir (Haykin, 1999, s.131).

Son olarak, birden fazla Adaline ünitesinin bir araya getirilmesiyle Madaline ağlarının elde edilebileceğini ve bu modellerinde yine doğrusal

sınıflandırma problemlerinde kullanılabileceğinin belirtilmesinde yarar vardır. İki Adaline ünitesinin birlikte kullanıldığı bir Madaline ağı şekil 16’da sunulmuştur.



**Şekil 16 : İki Adaline Ünitesinden Oluşan Madaline Ağı**

Burada en önemli konu şeklin son kısmında bulunan son işlemcinin ne olduğudur. Eğer “ve” işlemcisi kullanılmış ise sadece iki Adaline ünitesinin çıktılarının 1 olması durumunda Madaline ağının 1 diğer durumlarda -1 değeri üreteceği açıktır. Eğer “veya” işlemcisi kullanılmış ise iki Adaline ünitesinden sadece bir tanesinin 1 değeri üretmesinin Madaline ağının çıktısının 1 olması için yeterli olduğu söylenebilir (Öztemel, 20003, s.73).

## ÜÇÜNCÜ BÖLÜM

### YAPAY SİNİR AĞI MODELLERİNİN ZAMAN SERİSİ ANALİZİNDE KULLANIMI

Zaman serisi analizi konu başlığının içerisinde değerlendirilebilecek ve yapay sinir ağı modellerinin kullanıldığı bilimsel çalışmalarda kullanılan ağlar, çok katmanlı algılayıcılar ve geri dönüşümlü ağlardır. Bu nedenle, tez çalışmasının bu üçüncü bölümünde ele alınacak yapay sinir ağı modelleri, çok katmanlı algılayıcılar (MLP) ve geri dönüşümlü ağlar (RNN) olacak ve ele alınacak bu iki ağ genel özellikleriyle sunulacaktır. Bu ağların, zaman serisi analizi yaklaşımında başlıca konu başlıklarındaki (ki bu başlıklar zaman serileri modellemesi çabalarının tarihsel hikayesini oluşturduklarından ilk bölümde tartışılmış konulardır) performanslarının değerlendirilmesi uygulama bölümünde ele alınacaktır. MLP ve RNN modellerinin ele alınacağı bu bölüme giriş MLP ve RNN'in kullandıkları öğrenme kuralı öğretmenli öğrenme stratejisine dayandığı için istatistiksel öğrenme teorisinin bu geniş başlığının bir tanıtımı ile başlamaktadır. Bu bölümde ikinci olarak MLP tanıtılıp, ağların kullanımında en önemli konu başlıkları olan tahmin süreci ve uygun ağ yapısının seçimi sırasıyla 3. ve 4. alt başlıklar olarak sunulmuştur. Beşinci alt başlık olarak da RNN'lerin MLP'den farklı olarak ne işe yarayabileceği üzerinde durulmuştur. Son bir alt başlık olarak, uygulama bölümü için kullanışlı bir el haritası hazırlamak amacıyla yapay sinir ağlarıyla zaman serisi analizine bakış adımsal bir prosedür olarak sunulmuştur.

## I. ÖĞRETMENLİ ÖĞRENME STRATEJİSİ

Yapay sinir ağı modelleri günümüzde bir çok bilimsel amaçlı çalışmada kullanılmaktadır. Özellikle, örüntü tanıma, tahmin, doğrusal olmayan sistem modelleme, ilişkilendirme ve örüntü eşleştirme, sınıflandırma, veri filtreleme, optimizasyon ve tabii ki zaman serisi analizi konularında son on yılda popülaritesi artan bir modelleme aracıdır. Bu nedenle, ekonomi, finans, mühendislik, tıp vb. bilim dallarının, bu teorik konu başlıklarını günlük hayatta karşılaşılan problemlerin çözümünde kullandıkları uygulamaya dönük çalışmalarda yapay sinir ağları daha da sıklıkla karşımıza çıkmaktadır. Öte yandan yapay sinir ağı modelleri başlığı aslında oldukça geniş bir başlıktır ve bir çok çeşit yapay sinir ağı modeli bulunmaktadır (*Kaashoek / van Dijk 2001, s.2*). Birinci bölümde de değinildiği gibi, kullanılan öğrenme stratejisine göre başlıca üç değişik ağ modeli vardır. Bunlar sırasıyla Öğretmenli Öğrenme Stratejisini kullanan ağlar, Öğretmensiz Öğrenme Stratejisini kullanan ağlar ve Destekleyici Öğrenme stratejisini kullanan ağlardır.

Öğretmenli öğrenme stratejisi, İstatistiksel Öğrenme Teorisine dayanmaktadır. İstatistiksel öğrenme teorisinin klasik istatistikten farklı olduğu nokta; klasik istatistiksel ve ekonometrik modellemede ele alınan konu ile ilgili olarak her zaman doğru modelin bilindiği varsayımıyla hareket edilip, amaç parametre tahminlerinin en doğru şekilde elde edilmesi ile sınırlanmışken, istatistiksel öğrenme teorisine göre ise, doğru modelin bilinmediği ve amacın öncelikle doğru model yapısının ortaya konulması olduğudur (*Pednault, 1997, s.1*). Bu nedenle bir grup model arasından en doğru modelin seçimi için bazı performans tahminlerine ihtiyaç duyulmaktadır. Bu yapılırken kullanılacak öğrenme algoritmasının özellikleri üzerine istatistiksel bir çatı oluşturmak istatistiksel öğrenme teorisinin amacıdır (*Bousquet / Boucheron / Lugosi, 2003, s.175*). Diğer taraftan istatistiksel öğrenme, bir çok açıdan çok değişkenli fonksiyon yakınsamasına eşdeğer bir metottur (*Lawrence / Tsoi / Back, 1996, s.1*).

İstatistiksel öğrenme teorisi ve kullandığı kavramlar bir örnek yardımıyla şu şekilde açıklanabilir. Bir araştırmacı, bir ülkedeki finansal krizler üzerine çalışsın. Araştırmacı ilgili ülkenin  $n$  tane makro göstergesi hakkında aylık olarak veri toplasın. Bu durumda her özellik  $X_i \subseteq \mathbb{R}$  değişkeniyle gösterilirse her bir ay  $n$  boyutlu bir vektörle  $(x=x_1, x_2, \dots, x_n)$  tanımlanabilir. Böylece  $X$  girdi uzayı olarak ele alınır, her bir ay bu uzayda  $(X \subseteq \mathbb{R}^n)$  bir nokta ile temsil edilmektedir. Diğer yandan araştırmacı için olası iki çıktı vardır. Bunlar, ilgili ayı güvenli kılabilecek kriz olmadan geçebilecek bir ay ( $y=1$ ) ve ilgili ayı güvensiz kılabilecek finansal bir krizle yüz yüze kalınacak kriz ayı ( $y = -1$ ) durumlarıdır. Çıktı uzayı  $Y$  olarak ele alınır, araştırmacı sonlu sayıda (ele alınan ay sayısı  $t$  olsun) bir örnekleme  $S = \{(x_i, y_i)\}_{i=1}^t$  çalışacaktır. Sonuçta araştırmanın amacı, kriz olan aylar ile olmayan ayları, ilgili makro değişkenleri kullanarak bu örneklem üzerinden ayrı sınıflara atamak ise araştırmacı  $t$  boyutlu bir örneklem uzayı olan  $S$ 'de  $f = X \rightarrow Y$  şeklinde bir fonksiyon bulmalıdır. Elbetteki bu örnek araştırmacının ilgili makro değişkenlerin gecikmeli değerlerini de kullanabileceği şekilde genişletilebilir. Diğer bir deyişle her bir ayı temsil eden vektörde makro değişkenlerin önceki aylara ait değerleri de yer alabilir. Böylece  $x$  vektörüyle temsil edilen her bir yeni ay finansal bir krizin olmayacağı ( $y = 1$ ) yada finansal krizin olabileceği ( $y = -1$ ) şeklinde sınıflandırılabilir. Bu şekilde bir harita(sınıflandırıcı) bulma konusunda araştırmacının her bir  $f$  için bir riski vardır. Buradaki risk ilgili ayın bir kriz riski taşımaz iken, kriz riski taşıyor ya da kriz riski taşıırken, taşıyor şeklinde atama olasılığıdır.  $X$  uzayından rassal olarak seçilen bir ayın herhangi bir sınıfa atanma olasılığı  $P_{xy}$  olarak kabul edilirse ve atanacak doğru sınıf  $y$ , atama yapılan sınıf  $\hat{y}$  olarak belirlenirse atama maliyeti ya da kayıp fonksiyonu (loss function),

$$L(y, \hat{y}) = \begin{cases} 0 & y = \hat{y} \text{ (} y=f(x) \text{) ise} \\ 1 & y \neq \hat{y} \text{ (} y \neq f(x) \text{) ise} \end{cases}$$

olacaktır. Bu durumda,  $L(y, \hat{y})$  'nin  $P_{xy}$  'ye göre beklenen değeri (beklenen kayıp) yanlış sınıflandırma olasılığıdır. Daha genel olarak fonksiyonel risk,



$$R(f) = \int_{xy} L(y, f(x)) P_{xy}(x, y) dx dy$$

şeklindedir ve araştırmacı tarafından minimize edilmesi gereken değerdir. Ancak araştırmacı sonlu sayıda bir örneklem üzerinden çalışabildiğinden (S) (eğitim kümesi), sadece ampirik risk (eğitim hatası) ulaşılabilir bir değerdir ve

$$R_{emp}(f) = \frac{1}{t} \sum_{(x, y_i) \in S} L(y_i, f(x))$$

şeklindedir (*Herbitch vd., 1999, s.2*).

Sonuçta araştırmacı  $R_{emp}(f)$  fonksiyonunu yani ampirik riski ya da diğer bir ifade ile eğitim hatasını minimize edebilecek bilgiye sahiptir. Ancak burada iki problem vardır. Bunlardan birincisi ampirik risk minimize edilirken araştırmacı neredeyse sonsuz sayıda f fonksiyonu kullanabilir. İkincisi ise aşırı uyum diye adlandırılan ve ampirik riskin ( $R_{emp}(f)$ ) minimizasyon sonucu sıfıra çok yakın bir değer olarak elde edilmesi ile araştırmacının asıl ilgilendiği beklenen risk ( $R(f)$ ) ile arasındaki farkın artması durumudur.

Bu iki problemin aşılabilmesi için istatistiksel öğrenme teorisinin yaklaşımı şu şekildedir. İlk olarak beklenen risk  $R(f)$  ve ampirik risk arasındaki uzaklığa olasılıksal bir sınır koyulur. Böylece fonksiyon uzayında ampirik risk minimize edilirken aşırı uyum olgusu kontrol altında tutulabilir (*Evgeniou / Pontil/ Poggio, 2000, s10*). Bu olasılıksal sınır şu şekilde bir eşitsizlik formu ile mümkündür

$$R(f) < R_{emp}(f) + \Phi\left(\sqrt{\frac{h}{t}}, p\right)$$

burada t örneklem hacmidir. h fonksiyon uzayının kapasitesi olarak adlandırılır ve bu uzayın karmaşıklığının niceliksel bir ölçüsüdür.  $\Phi$  ise  $\frac{h}{t}$  oranı ve p olasılığının artan bir fonksiyonudur ve şu şekilde tanımlanmaktadır,

$$\Phi = \sqrt{\frac{h(\log(\frac{2t}{h}) + 1) - \log(\frac{p}{4})}{t}}$$

Bu formülüzasyona göre ampirik risk minimizasyonu yapılacak fonksiyon uzayının kapasitesi yüksek ve örneklem sayısı az ise ampirik risk ve beklenen risk arasındaki fark fazla olabilecektir. Diğer bir ifade ile aşırı uyum problemi ile karşılaşılması olası bir durumdur.

Fonksiyon uzayı ( $\mathbb{F}$ ) genelde çok geniş olduğundan daha küçük bir hipotez uzayında ( $\tilde{h}$ ) çalışmak düşünülebilecek iyi bir yaklaşımdır. Öte yandan, yukarıdaki eşitsizlik formu aslında bize iyi bir genelleştirme sağlayabilecek alternatif bir metod sunmaktadır Bu metod ampirik riski doğrudan minimize etmek yerine hipotez uzayının karmaşıklığı ve ampirik risk minimizasyonu arasında en iyi değiş tokuşu sağlamak üzerine kurulu Yapısal Risk Minimizasyonu Metodudur (*Evgeniou / Pontil / Poggio, 1999, s10*).

Yapısal risk minimizasyonu metodunun arkasında yatan temel düşünce birbirlerini kapsayan şekilde sıralanmış hipotez uzaylarına dayanır ( $\tilde{h}_1 \subset \tilde{h}_2 \subset \dots \subset \tilde{h}_m$ ) (*Bousquet / Boucheron / Lugosi, 2003, s.179*). Her bir hipotez uzayı bir öncekinden daha yüksek kapasiteye sahiptir, diğer bir ifade ile daha karmaşıktır ( $\tilde{h}_1 \leq \tilde{h}_2 \dots \leq \tilde{h}_m$ ). Burada m karmaşıklığın bir ölçüsü olarak sınıflandırmayı yapacak fonksiyon için örneğin,  $Y = \cos(m\pi + 2)$  şeklinde düşünülebilir. Bu durumda, daha önce belirtilen eşitsizlik formunun sağ tarafıyla sınırlanmış olan yapısal riskin minimizasyonu ile üzerinde çalışılan fonksiyon uzayının karmaşıklığı arasında en iyi değiş tokuşu veren hipotez uzayında aşırı uyumdan kaçınarak optimal minimizasyon elde edebilir (*Evgeniou / Pontil / Poggio, 2000, s11*).

Sonuçta, İstatistiksel öğrenme teorisi bir örneklemden hareketle öğrenme konusunda üç adım içeren strateji uygulamaktadır. Bu durum (Devroye vd., 1996), (Vapnik, 1998), (Evgeniou / Pontil / Poggio, 2000) de şu şekilde verilmektedir.

(1) Gerçek çıktı  $y$  iken tahmin edilen çıktının  $f(x)$  olduğu ve bu tahmin edilen çıktının  $L(y, f(x))$  şeklinde elde edildiği bir kayıp fonksiyonu tanımlamak,

(2) Kapasitesi fonksiyon uzayının karmaşıklığı şeklinde dizilmiş ve birbirlerini kapsayan hipotez uzaylarının tanımlanması,

(3) Ampirik riskin her bir fonksiyon uzayında tahmin edilip sonuçlar arasından minimizasyon ve fonksiyon uzayının kapasitesi arasında en iyi değiş tokuşu veren çözümün seçilmesi şeklindedir.

Sonuç olarak, öğrenme, bir fonksiyon minimizasyonu haline getirilmiştir.

Öte yandan, bir regresyon problemi için tanımlanacak kayıp fonksiyonu ise,

$$L(y, f(x)) = (y - f(x))^2$$

şeklinde bir kuadratik kayıp fonksiyonu olacaktır. Dikkat edilecek olursa bu durumda kayıp fonksiyonu tahmin edilen değer doğru olup olmadığıyla değil gerçek değerden ne kadar uzak olduğuyula ilgilenmektedir. Öte yandan bu kez fonksiyon uzayında beklenen kuadratik kaybı minimum yapan fonksiyon aranmakta ve fonksiyon uzayının kapasitesi ile en düşük kuadratik kayıp arasında değiş tokuş koşulları araştırılmaktadır. Bu durumun sınıflandırma problemlerinden bir farklılığı da, kuadratik kaybın sınırlandırılmış olmamasıdır. Çünkü  $y$  değerleri için bir alt ve aslında çoğu zaman bir üst değer yoktur. Bu ilk bakışta büyük bir problem olarak görülmesi de, diğer bir ifade ile,  $f(x) - y$  farkının çok büyük olma olasılığı düşük olsa da, böyle bir durumun maliyeti büyük olacaktır. Bu problemten kaçınmak için  $y$  değerlerinin bir aralık içerisinde yer aldığı varsayımının getirilmesi gerekmektedir. Öte yandan daha sonrada değinileceği gibi lojistik yada hiperbolik tanjant transfer fonksiyonu kullanıldığında girdi ve çıktılarının sırasıyla  $(0,1)$  yada  $(-1,1)$  arasında değerler alması gerektiğinden, zaten uygulamada bir dönüşüm kullanılması gerekmektedir. Bu nedenle, bu varsayım aslında bir kısıt yaratmamaktadır (*Antony / Barlett, 1999, s.232*).

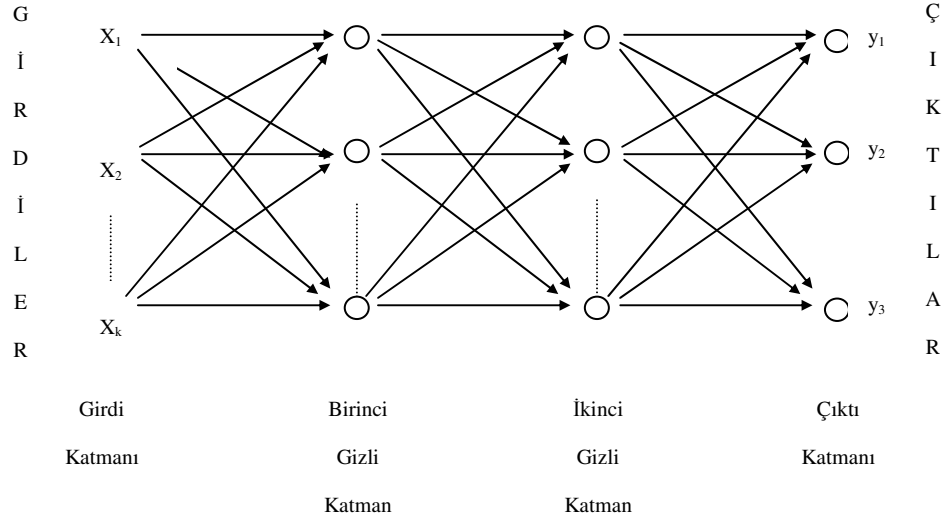
Burada birkaç noktaya değinilmesi bu çalışmanın esasına ve devamına ışık tutması bakımından önem taşımaktadır. Aslında daha çok mühendislik alanında matematiksel bir yaklaşım çerçevesinde şekillenmiş olan yapay sinir ağı modelleme yaklaşımı, iktisadi veriler üzerinde istatistiksel bir araç olarak kullanılmaya başlandığında literatürde bir dizi tartışmayı da beraberinde getirmiştir. Mühendislik alanının iktisada oranla göreceli olarak çok daha geniş veri setiyle çalışabilme fırsatı ve sadece elde olan probleme çözüm getirme anlayışı yapay sinir ağlarının iktisadi verilere istatistiksel bir araç olarak uygulanması konusunda güvensizlik yaratmıştır. Breiman (1994), istatistiğin konuya evrensel bir metodoloji getirmek ve değiştirmeden her ortamda kullanabilmek eğilimiyle yaklaşması gerekliliğini belirtmektedir ve çalışmalarında bu şekilde yaklaşımlarda bulunmayanları problem orijinli çalışmakla tenkit etmektedir (*Breiman, 1994, s.39*). Tibshirani (1994), Yapay sinir ağı ile çalışanların kullandıkları metotların istatistiksel özellikleri üzerine daha çok düşüncelerinin gerekliliğini vurgulamaktadır (*Tibshirani, 1994, s.48*). Aslında konuyu sadece yapay sinir ağlarıyla kısıtlamak yersizdir. İstatistik ve matematik disiplinleri arasında bu tip tartışmalara rastlamak mümkündür. Box (1996), istatistiğin matematiğin bir alt dalı olmak yerine bilim dalı olduğu tartışmasını yaparken istatistiğin amaçlarından birinin de yeni bilgi (knowledge) keşfetme olduğunu belirtir (*Box, 1996, s.4*). Öte yandan Hand (1998) ise, matematiğin istatistik içinde sadece bir bileşen olduğunu ve istatistiğin diğer iki önemli bileşenin sırasıyla, araştırılmakta olan alanın iyi anlaşılması gerekliliği ve gerekli hesapsal yöntemlere yatkınlık olduğunu belirterek, istatistiksel pratiğin bu üç bileşenden birinin eksikliği durumunda her hangi bir anlam taşımayacağını söylemektedir. Ayrıca bugünün istatistiğinin geçmişten farklı olduğunu ve geleceğin istatistiğinin de bugünkünden farklı olacağını vurgulayarak, bunun sadece kullanılan istatistiksel teknikler için değil aynı zamanda istatistiksel uygulamalar içinde doğru olduğunu belirtmekte ve veri madenciliği, yapay sinir ağları gibi konularda çalışanların büyük çoğunluğunun kendilerini istatistikçi olarak adlandırmadıkları örneğini vermektedir (*Hand, 1998, s.246-249*). Bu nedenlerle bu bölümde, birinci bölümde tek katmanlı algılayıcı ve Adaline üniteleri tartışılırken sadece matematiksel bir fonksiyon elde edilmesi gayretine ek

olarak, transfer fonksiyonlarının özellikleri, yapay sinir ağlarının evrensel yakınsayıcı özellikleri, ağ seçimi için çeşitli istatistiksel teknikler gibi konulara da yeri geldikçe değinilecektir.

Son olarak, öğretmenli öğrenme stratejisi kullanılarak dizayn edilen bir öğrenen sistemde cevaplanması gereken başlıca üç soru vardır. Bunlardan ilki öğrenen sistemin yakınsama özelliklerinin ne olduğudur. İkincisi, tahmin özelliklerinin ne olduğudur. Üçüncüsü ise kullanılan öğrenme algoritmasının etkinliğinin olup olmadığıdır (*Antony / Barlett, 1999, s.1*). İzleyen başlık altında Çok Katmanlı Algılayıcılar genel özellikleriyle anlatılırken aslında bu soruların cevapları tartışılacaktır.

## **II. ÇOK KATMANLI ALGILAYICILAR**

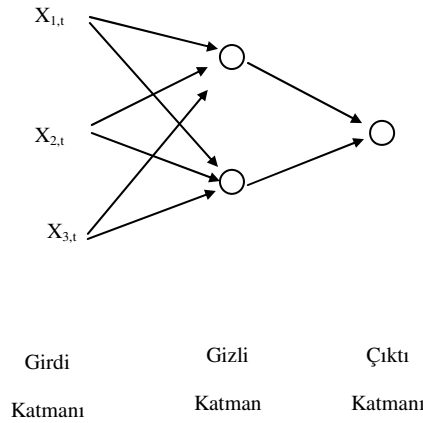
Çok katmanlı algılayıcılar yapay sinir ağları modelleri ailesi içerisinde en çok kullanılan ağ tipidir. Aslında tek katmanlı algılayıcının genelleştirilmiş bir hali olarak da düşünülebilir (*Baron, 1994, s.33*). Bu ağlarda girdi sinyali katmanlar boyunca ileriye doğru işlenerek çıktı üretilmektedir. Aşağıda şekil 17’de çok katmanlı bir algılayıcı örneği olarak üç katmanlı bir algılayıcıya yer verilmiştir.



**Şekil 17 : Üç Katmanlı Algılayıcı**

Şekil 17’de sinir hücreleri daire olarak gösterilmiştir. Birimler arası bağlantılar tek yönlü oklarla gösterilmiş olup sinyal akışı bir ileri katmana doğrudur. Katmanın kendi içerisinde yada geriye doğru bir sinyal akışı yoktur. Şekildeki her ok bir ağırlık değerine sahiptir. Öte yandan bu ağ adlandırılırken girdi katmanı sayılmayıp geriye kalan katmanlar üzerinden adlandırılma yapılır. Sonuçta, girdi katmanı, iki ara katman ve çıktı katmanından oluşan böyle bir ağa üç katmanlı algılayıcı denir. Çıktı katmanı istatistik literatüründe bağımsız değişkenlere ( $y$ ), girdi katmanı ise bağımlı değişkene karşılık gelir( $x$ ), ara katmanlar dediğimiz ve girdi katmanı ile çıktı katmanı arasında kalan katmanlara gizli(hidden) katman adı verilmektedir (*Jordan / Bishop, 1996, s.2*). Ağa doğrusal dışı yapıyı analiz etme yeteneği bu ara katmanlar ve bu katmanlarda yer alan nöronlar tarafından verilmektedir (*Koskivaara, 2000, s.1069*). Diğer taraftan bu gizli katmanda yer alan nöronlarda da gizli üniteler olarak adlandırılır. Son olarak çok katmanlı algılayıcıların literatürdeki bir diğer adı ileri beslemeli ağlardır. İleri beslemeli ağ (FNN) adı genelde zaman serisi çalışmalarında çok katmanlı ağa verilen addır ve çok katmanlı algılayıcı daha genel kapsamlı bir ad olarak kullanılır

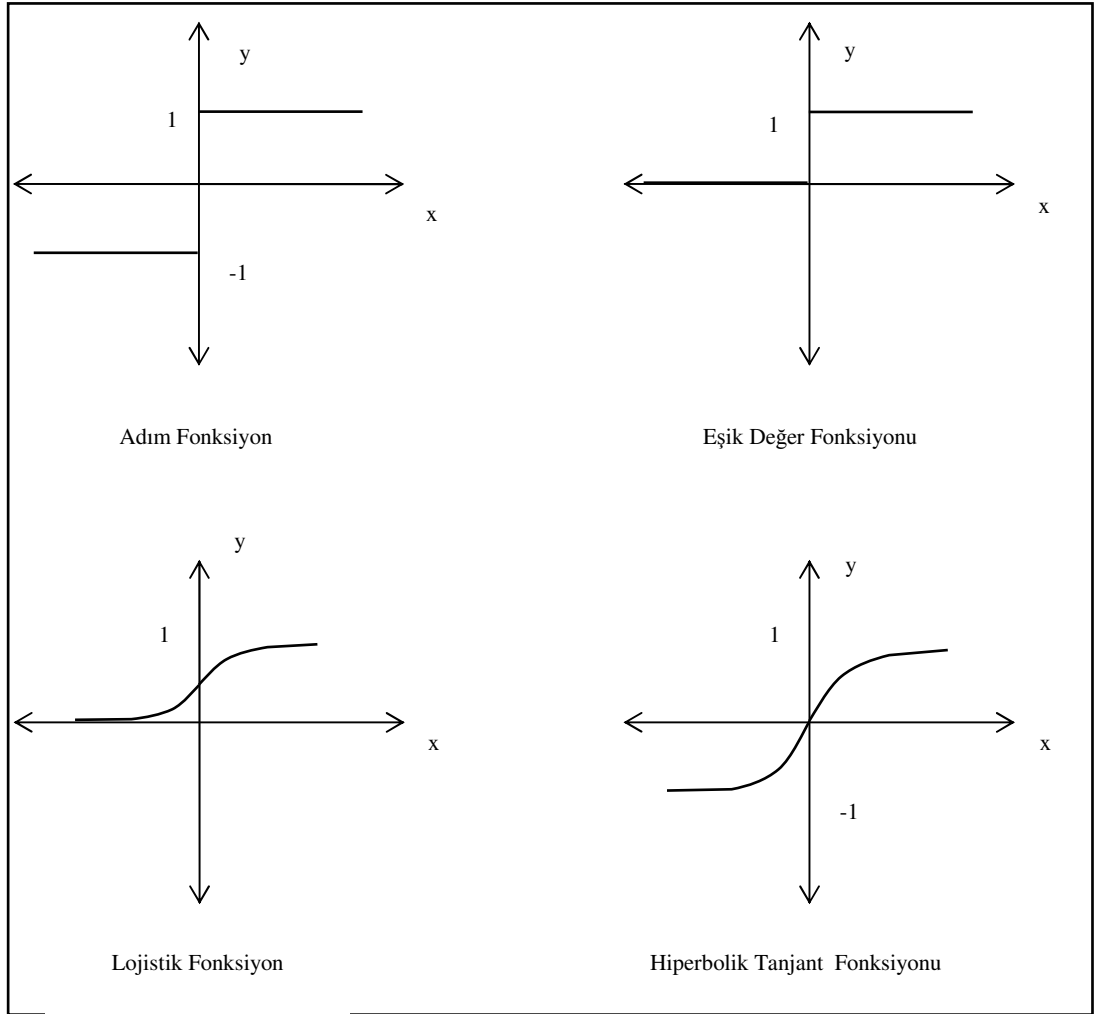
Zaman serisi analizinde sıklıkla kullanılan bir ileri beslemeli ağ yapısı aşağıda şekil 18’de sunulmuştur. Girdi katmanı hesaba katılmadan ve gizli ünite sayısı belirtilerek adlandırılma yapılacağından bu ağ iki katmanlı iki gizli üniteli ileri beslemeli bir ağ olarak adlandırılabilir. Öte yandan bazı çalışmalarda ise sırasıyla katmanlardaki ünite sayıları verilerek de adlandırılma yapılmaktadır. Bu durumda ağa verilecek ad 3-2-1 ileri beslemeli ağ olacaktır.



**Şekil 18 : İki Katmanlı İki Gizli Üniteli İleri Beslemeli Ağ**

Öte yandan adlandırılma konusu bir tarafa bırakılırsa asıl önemli olan nokta ağın veri setini yada bir başka ifade ile bilgiyi nasıl işlediğidir. Şekilden de açıkça görüldüğü gibi bu tip bir ağda bilgi sırasıyla üç katmandan ileri doğru geçirilmektedir. İlk katman olan girdi katmanının görevi bilgiyi bir sonraki katman olan gizli katmana aktarmaktır. Burada veri herhangi bir işlemden geçirilmemektedir. Gizli katman olarak adlandırılan ikinci katmanda yer alan üniteler yada sinir hücreleri sahip oldukları transfer fonksiyonlarından veriyi geçirerek elde ettikleri değerleri bir sonraki katman olan çıktı katmanına yollarlar. Çıktı katmanı kendisine gelen bu bilgiyi kendi transfer fonksiyonundan geçirerek çıktıyı üretmiş olur. Elde edilen çıktı istenilen çıktı değeriyle karşılaştırılarak hata payı ağ boyunca ağırlıklara geri yayılır. Düzenlemeler yapılarak veri seti ağa tekrar sunulur başta anlatılan süreç tekrarlanarak çıktı tekrar elde edilir ve istenilen çıktı değeriyle karşılaştırılır. Bu işlem istenilen çıktıya kabul edilir hata değeriince yaklaşılan dek sürdürülür (Wang / Zhang, 1997, s133). Bu işlem için

kullanılan en yaygın kabul görmüş algoritma geri yayılım algoritmasıdır. Bu işlem esnasında veri setin tamamı değil bir bölümü kullanılmaktadır. Bu kullanılan bölüme eğitim seti adı verilmektedir. Eğitim setindeki bilgi ile eğitilen ağ geri kalan ve test seti olarak adlandırılan veri ile test edilir.Yapay sinir ağlarında kullanılan transfer fonksiyonları genelde doğrusal olmayan fonksiyonlardır. Doğrusal olmayan fonksiyonların tercih nedeni doğrusal fonksiyonlarda çıktının girdi ile orantılı olmasıdır (Yurtoğlu,2005;17). Öte yandan doğrusal dışı fonksiyonların kullanımı ağın doğrusal dışı problemlere, ilgili bilgiyi tüm ağa yayarak (bir anlamda doğrusal dışılığı da ağa dağıtmış olarak) çözüm getirmesini sağlamaktadır (Csaji, 2001, s.4). Literatürde sıklıkla kullanılan transfer fonksiyonları, adım, eşik, lojistik ve hiperbolik tanjant fonksiyonlarıdır.



**Şekil 19 :Sıklıkla Kullanılan Transfer Fonksiyonları**



Adım fonksiyon,  $\varphi(x) = \begin{cases} 1, & x \geq 0 \text{ ise} \\ -1, & x < 0 \text{ ise} \end{cases}$  şeklindedir.

Eşik değer fonksiyonu,  $\varphi(x) = \begin{cases} 1, & x \geq 0 \text{ ise} \\ 0, & x < 0 \text{ ise} \end{cases}$  şeklindedir.

Lojistik fonksiyon,  $\varphi(x) = \frac{1}{1 + \exp(-x)}$  şeklindedir.

Hiperbolik tanjant fonksiyonu,  $\varphi(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)}$  şeklindedir.

Bu fonksiyonlar arasında lojistik fonksiyonu ve hiperbolik tanjant fonksiyonu daha sıklıkla kullanılmaktadır. Hangi tip transfer fonksiyonu kullanılacağına karar vermede çok keskin kurallar olmasa da *Klimasauskas, (1991)*, ağın ortalama davranışı öğrenilmesi isteniyorsa lojistik, ortalamadan sapmanın öğrenilmesi isteniyorsa hiperbolik tanjant fonksiyonları transfer fonksiyonu olarak kullanılabilirliğini ileri sürmektedir. Diğer taraftan ileri beslemeli ağlar için, transfer fonksiyonları üzerinden tanımlanan çok önemli bir konu da bu ağların evrensel yakınsayıcı özellikleridir. ***Evrensel yakınsama teoremi***, ileri beslemeli bir ağın yeterli sayıda gizli katman ve bu katmanlarda yeterli sayıda nöron ile doğru ayarlanmış parametreler ve bir transfer fonksiyonu yardımıyla seçilen herhangi bir fonksiyona yakınsayabileceğini belirtmektedir. Bu konuya daha sonra ağ seçimi başlığı altında değinilecek olmasına rağmen burada, transfer fonksiyonlarını ilgilendiren kısmı üzerinde durmak gerekmektedir.

Burada,  $\varphi(\cdot)$  herhangi bir transfer fonksiyonunu,  $f(x)$  isteğe bağlı seçilmiş bir fonksiyonu,  $(A_n f)(x)$  ağı oluşturacağı fonksiyonu,  $\varepsilon$  yakınsama hatasını,  $w_i, b_i$  sırasıyla ağırlıklar ile yan değerlerini ve  $n$  ise, gizli katmandaki nöron sayısını göstermektedir.

$$(A_n f)(x) = \sum_{i=1}^n w_i \varphi(x + b_i) \quad \text{şeklinde tek katmanlı bir ileri beslemeli ağ için}$$

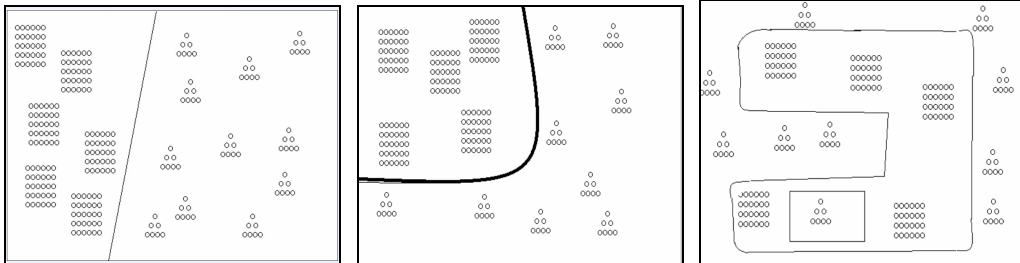
yakınsama, matematiksel olarak,

$$|f(x) - A_n f(x)| = \sqrt[p]{\int_x |f(x) - A_n f(x)|^p dx} \quad \text{şeklinde tanımlanabilir. Bu}$$

tanım  $p = 2$  durumunda ortalama kare hataya karşılık gelecektir (Kuan, 2006, s.10). Tek kısıt transfer fonksiyonu için;

$$\lim_{x \rightarrow -\infty} \varphi(x) = a ; \lim_{x \rightarrow +\infty} \varphi(x) = b \quad \text{ve } a \neq b \text{ olmasıdır.}$$

Aslında ileri beslemeli bir ağ transfer fonksiyonlarının katmanlar olarak düzenlendiği bir yapıdır (Gençay, Selçuk, Whitcher, 2002, s.279). İlgilenilen problemin karmaşıklığı ölçüsünde katman sayısı değişecektir. Aşağıda şekil 20’de bu konuya açıklık getirilmektedir.



(a)

(b)

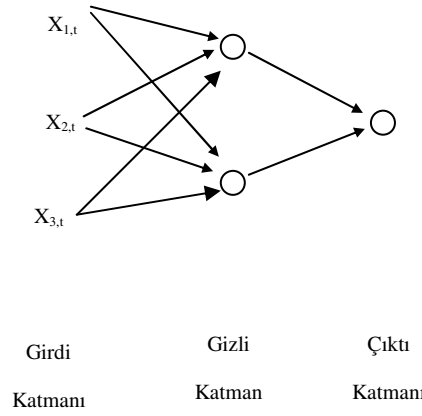
(c)

**Şekil 20 : Problem Karmaşıklığı**

Şekil 20’de açıkça görüldüğü gibi üçgen ve dikdörtgenlerden oluşan uç ayrı düzlemde sınıflandırma yapılabilmesi için a kısmında bir doğru fonksiyonuna, b kısmında bir konvekse ve c kısmında isteğe bağlı ayarlanmış bir fonksiyona ihtiyaç duyulmaktadır. Bu fonksiyonlardan (a) kısmındaki fonksiyon

tek katmanlı ağlar tarafından oluşturulabilir. Bununla ilgili olarak basit algılayıcı ve ADALINE üniteleriyle benzer fonksiyonlar birinci bölümde elde edilmiştir. (b) kısmında kullanılacak fonksiyonun elde edilmesi için iki katmanlı bir algılayıcıya gerek duyulmaktadır, (c) kısmındaki ayırıcı fonksiyonun elde edilmesi için ise üç katmanlı bir algılayıcı gereklidir (*Wiggins / Looper, 1991.s.27*).

Şekil 21’de iki katmanlı ve iki gizli üniteli bir ileri beslemeli ağ verilmiştir. Bu ağ ayrıca 3-2-1 ileri beslemeli ağ olarak da adlandırılabilir.

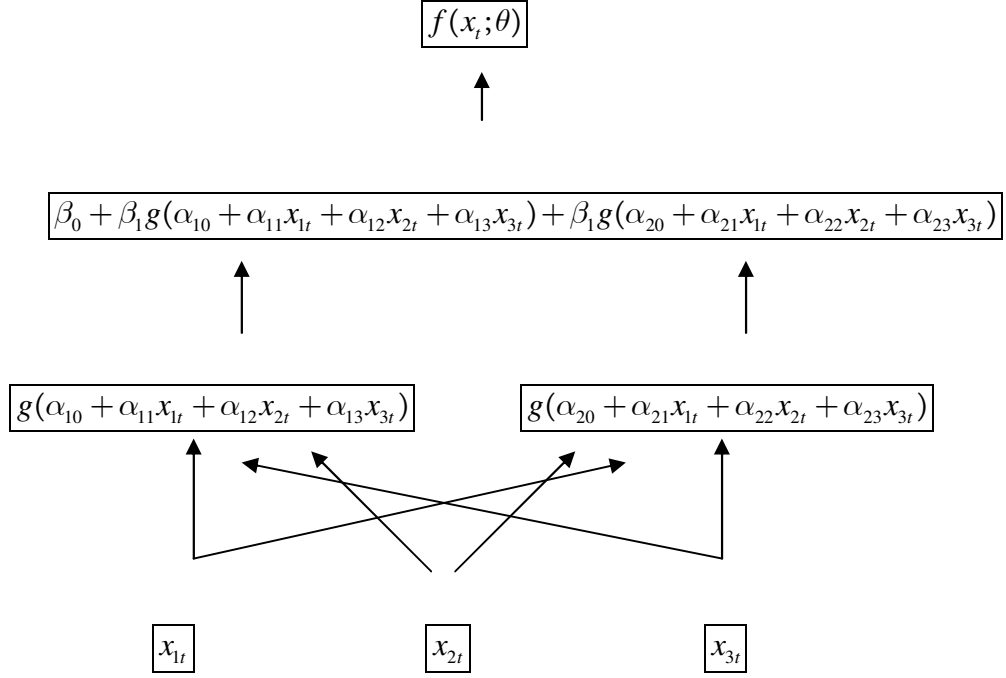


**Şekil 21 : 3-2-1 İleri Beslemeli Ağ**

Bu ağ yapısı matematiksel olarak,

$$f(x_t, \theta) = \beta_0 + \beta_1 g(\alpha_{10} + \alpha_{11}x_{1t} + \alpha_{12}x_{2t} + \alpha_{13}x_{3t}) + \beta_2 g(\alpha_{20} + \alpha_{21}x_{1t} + \alpha_{22}x_{2t} + \alpha_{23}x_{3t})$$

şeklinde yazılabilir. Anlaşılır olması bakımından, şekil 22 sunulmuştur.



**Şekil 22 : 3-2-1 İleri Beslemeli Ağ (Fonksiyon Gösterimi)**

Burada,

$x_t : (x_{1t}, x_{2t}, x_{3t})$  t zamanına ait girdilerdir.

$g$  : bilinen transfer fonksiyonlarından herhangi biridir.

$\alpha_{1i} : (\alpha_{10}, \alpha_{11}, \alpha_{12}, \alpha_{13})$  birinci transfer fonksiyonunun parametreleridir.

$\alpha_{2i} : (\alpha_{20}, \alpha_{21}, \alpha_{22}, \alpha_{23})$  ikinci transfer fonksiyonunun parametreleridir.

$b_j : (\beta_0, \beta_1, \beta_2)$  kesme ve eğim parametreleridir.

$f(x_t, \theta)$  : ağın çıktısı olup, girdiler ve ağ parametrelerine  $[\theta = (\alpha, \beta)]$  göre değer alır.

### III. TAHMİN

Verili bir ağ yapısı ve belirlenmiş transfer fonksiyonları için en önemli konu ağırlık tahminlerinin elde edilmesidir. Bu anlamda en sık kullanılan algoritma geriye yayılım algoritmasıdır. Bununla beraber doğrusal olmayan en küçük kareler yöntemi de geniş kullanım alanı bulunan ve bazı üstünlüklere sahip olan bir yöntemdir. Zaten şekil 18 de verilen yapı, doğrusal dışı regresyondan çok da farklı değildir (*Biennenstock ve Geman, 1994, s.37*).  $f(x_t, \theta)$  yada  $\hat{y}$  olarak gösterebileceğimiz ağın çıktısı ağ parametrelerini temsil eden,  $\theta = (\alpha, \beta)$  'ya göre değer alır. Burada  $\alpha$  ve  $\beta$  şeklinde gösterilen  $\theta$ , nöronlar arası bağlantı ağırlıkları ( $w$ ) ve yan değerleri ( $b$ )'ler den oluşan bir vektördür. Ağın eğitilmesi ve böylece yakınsamada bu vektörde yapılan güncellemelerle sağlanmaktadır. Ağın eğitilmesi adı altında yapılan iş aslında,

$$\theta^* = \arg \min E(y_t - f(x_t, \theta))^2$$

şeklinde  $\theta$  vektörünün,  $\theta^*$  ile gösterilen vektöre yakınsamasıdır. Bu tahminin tutarlı bir şekilde elde edilebileceği farklı algoritmalar kullanılabilir. Her biri doğrusal dışı minimizasyon yapıldığı için iterasyon mantığıyla çalışmaktadır (*Madsen, Nielsen ve Tingleff, 2004, s.6*). Bu algoritmalar için ortak hareket noktası, her hangi bir  $r$ . iterasyon için kalıntıların yada hataların gradientini veren aşağıdaki eşitliktir.

$$\nabla Q(\theta^r) = \frac{\partial \nabla Q(\theta^r)}{\partial \theta}$$

Yeni bir iterasyon ( $r+1$ ) için, tahmin,

$$\theta^{r+1} = \hat{\theta}^r - \lambda A(\theta^r)^{-1} \nabla Q(\theta^r)$$

şeklinde elde edilir. Burada  $\lambda$  adım uzunluğu,  $A(\theta^r)$  matrisi ise  $\theta$ 'nın temsil ettiği parametrelerin bir fonksiyonudur. Zaten farklı algoritmalarda bu matrisin farklı seçimlerinden türemektedir.

## A. GERİ YAYILIM ALGORİTMASI

Çok katmanlı bir algılayıcının eğitiminde kullanılan bir çok algoritma(makine) olmasına rağmen en çok rağbet gören algoritma Geri Yayılım Algoritmasıdır. Diğer bir adı da Genelleştirilmiş Delta Algoritmasıdır. Genel mantığı, hata kareleri ortalamasının dereceli düşürme yöntemiyle minimize edilmesi üzerine kurulmuştur. Bu azaltma işlemi, nöronlar arasındaki bağlantı değerlerinin yani ağırlıkların ayarlanması şeklinde gerçekleştirilmektedir (*Curram ve Minger, 1994, s.441*). Algoritmanın mantığı şu şekilde çalışır.

n. iterasyonda çıktı katmanındaki j. nöronun ürettiği çıktı değeri  $y_j$  olsun. Bu durumda bu değer,

$$y_j = \varphi(v_j(n))$$

şeklinde hesaplanmaktadır. Burada,  $\varphi(\cdot)$  transfer fonksiyonu,  $v_j(n)$  ise toplam fonksiyonu kullanılmasıyla, n. iterasyonda j. nöronun i. nörondan gelen çıktı değeri ile ağırlık değerlerinin ( $w_{ji}$ ) çarpılıp toplanmasıyla oluşan net girdiyi göstermektedir.  $v_j(n)$  matematiksel olarak,

$$v_j(n) = \sum_{i=0}^m w_{ji}(n) y_i(n)$$

şeklinde gösterilebilir. Öte yandan eğer j. nöron ilk gizli katmanda yer alıyorsa,  $y_i(n)$  değeri aslında girdi katmanındaki değere yani  $x_i(n)$ 'e eşit olacaktır. Bu durumda eğer beklenen çıktı değeri  $d_j(n)$  olarak ifade edilirse, sistemin ürettiği çıktı değeri ile beklenen çıktı arasındaki fark hatayı oluşturacaktır ve,

$$e_j(n) = d_j(n) - y_j(n)$$

olarak gösterilebilir. Eğer bu hata hata kareler toplamı cinsinden ifade edilirse,  $\varepsilon_j(n) = \frac{1}{2} e_j^2(n)$  değeri elde edilir. Eğer çıktı katmanında birden fazla sayıda,

örneğin N tane nöron yer alıyorsa bu durumda ortalama kare hata değeri,

$$\varepsilon_j(n)_{ort} = \frac{1}{N} \sum_1^N \varepsilon_j(n)$$

olacaktır (Haykin, 1999, s.161). Geri yayılım algoritması, ağırlık değerlerini, kısmi türevden yararlanarak, ağırlık hata fonksiyonunu minimize edene kadar değiştirecektir (Ng / Leung / Luk, 1999, s.15). Zincir kuralına göre hatanın ağa geri yayılımı,

$$\frac{\partial \varepsilon(n)}{\partial w_{ji}} = \frac{\partial \varepsilon(n)}{\partial e_j(n)} = \frac{\partial e_j(n)}{\partial y_j(n)} = \frac{\partial y_j(n)}{\partial v_j(n)} = \frac{\partial v_j(n)}{\partial w_{ji}}$$

şeklinde olacaktır. Bu gradientteki her bir terim için aşağıdaki eşitlikler dikkate alınarak,

$$\frac{\partial \varepsilon(n)}{\partial e_j(n)} = e_j(n), \quad (\varepsilon_j(n) = \frac{1}{2} e_j^2(n) \text{ olduğundan})$$

$$\frac{\partial e_j(n)}{\partial y_j(n)} = -1, \quad (e_j(n) = d_j(n) - y_j(n) \text{ olduğundan})$$

$$\frac{\partial y_j(n)}{\partial v_j(n)} = \varphi'(v_j(n)), \quad (y_j = \varphi(v_j(n)) \text{ olduğundan})$$

$$\frac{\partial v_j(n)}{\partial w_{ji}} = y_i(n), \quad (v_j(n) = \sum_{i=0}^m w_{ji}(n) y_i(n) \text{ olduğundan})$$

sonuçta,

$$\frac{\partial \varepsilon(n)}{\partial w_{ji}} = -e_j(n) \varphi'(v_j(n)) y_i(n)$$

elde edilir. Bu durumda daha önceden belirlenmiş bir öğrenme katsayısı ( $\eta$ ) ile ağırlıklara uygulanacak düzeltme ( $\Delta w_{ji}$ ),

$$\Delta w_{ji} = - \eta \frac{\partial \varepsilon(n)}{\partial w_{ji}}$$

olacaktır. Burada (-) işareti, hata uzayında, global minimum olması umulan, lokal minimum değerine ulaşılmaya çalışıldığını göstermektedir (Azoff, 1994, s.150).

Son olarak, lokal gradient ( $\delta_j(n)$ ) değeri,

$$\delta_j(n) = e_j(n) \varphi'(v_j(n))$$

şeklindedir. Lokal gradient kullanılarak ağırlıklara uygulanacak düzeltme,

$$\Delta w_{ji} = - \eta \delta_j(n) y_i(n)$$

şeklinde gösterilebilir.

Öte yandan tüm bu anlatılanlar j. nöronun çıktı katmanında yer aldığı durum için düzenlenmiştir. Ancak ağırlık geri kalan katmanlarındaki nöronlarında ağırlık değerlerinin de öğrenme işlemi boyunca ayarlanması gerekmektedir. Bir başka ifade ile hata enerjisi diye tanımlanan ortalama kare hata tüm ağa yayılmalı ve ilgili düzeltmeler yapılmalıdır. Çünkü, bu ara katmanlardaki nöronların, çıktı katmanındaki hatadan paylarını almaları gerekmektedir (Öztemel, 2003, s.80). Bu durumda gizli katmanlardaki nöronların bağlantılarının ağırlık değerlerinin düzenlenmesi şu şekilde bir yol izlenerek yapılmaktadır.

Gizli katmanda yer alan bir k. nöron için lokal gradient, çıktı katmanında yer alan j. nöron için oluşturulmuş lokal gradientden hareketle şu şekilde yazılabilir,

$$\delta_k(n) = - \frac{\partial \varepsilon(n)}{\partial y_k(n)} \frac{\partial y_k(n)}{\partial v_k(n)}$$

bu eşitlik daha açık bir şekilde yazılacak olursa,



$$\delta_k(n) = -\frac{\partial \varepsilon(n)}{\partial y_k(n)} \varphi'_k(v_k(n))$$

şeklindedir. Bu durumda,  $\frac{\partial \varepsilon(n)}{\partial y_k(n)}$  kısmi türevinin elde edilmesi gerekmektedir.

Daha önce çıktı katmanındaki bir  $j$ . nöron için,  $\varepsilon_j(n) = \frac{1}{2} e_j^2(n)$  olduğu belirtilmiştir. Bu eşitliğin  $y_k(n)$ ' e göre türevinin alınması ile,

$$\frac{\partial \varepsilon(n)}{\partial y_k(n)} = \sum_j e_j \frac{\partial e_j(n)}{\partial y_k(n)}$$

elde edilir. Bir kez daha zincir kuralına başvurulduğunda,

$$\frac{\partial \varepsilon(n)}{\partial y_k(n)} = \sum_j e_j \frac{\partial e_j(n)}{\partial v_k(n)} \frac{\partial v_k(n)}{\partial y_k(n)}$$

elde edilecektir. Burada  $\frac{\partial e_j(n)}{\partial v_k(n)}$  kısmi türevinin daha önceki açıklamalar

doğrultusunda  $-\varphi'_j(v_j(n))$  olduğu açıktır. Ayrıca  $v_j(n) = \sum_{k=0}^m w_{jk}(n) y_k(n)$  olduğu bir kez daha hatırlatılacak olursa, bu durumda,

$$\frac{\partial v_j(n)}{\partial y_k(n)} = w_{jk}$$

olacaktır. Böylece,

$$\frac{\partial \varepsilon(n)}{\partial y_k(n)} = -\sum_j e_j(n) \varphi'_j(v_j(n)) w_{jk}(n)$$

elde edilebilir. Öte yandan, bu durumda  $k$ . nöronun gizli katmanda yer alması durumunda kullanılacak lokal gradient,

$$\delta_k(n) = \varphi'(v_k(n)) \sum_j \delta_j(n) w_{jk}(n)$$

şeklindedir (Haykin, 1999, s.166). Eğer lokal gradient kullanılarak  $\frac{\partial \varepsilon(n)}{\partial y_k(n)}$  kısmi türevi yazılacak olursa,

$$\frac{\partial \varepsilon(n)}{\partial y_k(n)} = \delta_k(n) w_{jk}(n)$$

şeklinde elde edilebilir. Sonuçta geri yayılım algoritması n. iterasyonda ağırlık düzeltilmesi ( $\Delta w_{ji}$ ) için şu kuralı sunmaktadır,

$$\Delta w_{ji} = - \eta \delta_j(n) y_i(n)$$

bu eşitliğin kullanımında dikkat edilmesi gereken nokta, her bir iterasyon sonucunda ağırlıkları güncellenecek nöronun hangi katmanda yer aldığıdır. Eğer ağırlıkları güncellenecek nöron çıktı katmanında yer alıyorsa, bu durumda lokal gradient  $\delta_j(n)$ ,

$$\delta_j(n) = e_j(n) \varphi'(v_j(n))$$

şeklinde ele alınmalıdır. Eğer gizli katman olarak adlandırılan ara katmanlardaki bir nöronun ağırlıkları güncellenecekse, bu durumda lokal gradient,

$$\delta_k(n) = \varphi'(v_k(n)) \sum_j \delta_j(n) w_{jk}(n)$$

şeklinde kullanılmalıdır.

Önceki bölümde de belirtildiği gibi geri yayılım algoritmasının kullanıldığı durumda transfer fonksiyonlarının türevlenebilir olması gerektiği açıktır.

Son olarak *Öğrenme Oranı ve Momentum Katsayısının* bu başlık altında tartışılması gerekmektedir. Öğrenme oranı nöronlar arasındaki bağlantıların ağırlık katsayılarındaki değişimin belirlenmesinde yada bir başka ifade ile ağırlık değerlerinin güncellenmesinde önemli bir rol oynamaktadır. Bu oranın çok küçük olması öğrenme zamanını uzatmakta, gereğinden büyük seçilmesi ise salınımın çok fazla artmasına ve böylece de yakınsamada bir kararsızlığın ortaya çıkmasına neden olmaktadır (Hagan / Demuth 1995, s.9-5). Bu nedenle Rumelhart (1986), algoritmayı modifiye ederek ağırlık güncellemesi için kullanılan eşitliğe momentum terimi ( $\alpha$ ) eklemiştir. Bu modifikasyonla ağırlık değişimi,

$$\Delta w_{ji}(n) = \alpha \Delta w_{ji}(n-1) + \eta \delta_j(n) y_j(n)$$

şeklinde bir eşitliğe dönüşmüştür.

Bu eşitlik Genelleştirilmiş Delta Kuralı adını alıp,  $\alpha = 0$  durumunda Delta Kuralını verir.  $\Delta w_{ji}(n)$ 'e göre çözümlendiğinde,

$$\Delta w_{ji}(n) = \eta \sum_{t=0}^n \alpha^{n-t} \frac{\partial \varepsilon(t)}{\partial w_{ji}(t)} \text{ elde edilir.}$$

Bu ilişki sayesinde Haykin (1999) şu çıkarımları yapmaktadır,

i) Zaman serilerinde yakınsama gereğince momentum katsayısı  $0 \leq |\alpha| \leq 1$  değerleri arasında olmalıdır. Teorik olarak momentum katsayısı negatif değerler alabilse de pratikte her zaman pozitif değerler almalıdır.

ii) Ardı ardına yapılan iterasyonlarda  $\frac{\partial \varepsilon(t)}{\partial w_{ji}(t)}$  kısmi türevi aynı işareti elde ediyorsa,  $\Delta w_{ji}(n)$  giderek daha büyük oranda güncelleniyor olur. Bunun anlamı momentum katsayısının öğrenme sürecini hızlandırdığıdır.

iii) Ardı ardına yapılan iterasyonlarda  $\frac{\partial \varepsilon(t)}{\partial w_{ji}(t)}$  kısmi türevi farklı işaretli

olarak elde edilmekteyse,  $\Delta w_{ji}(n)$  küçük miktarda ayarlama maruz kalıyordur. Bu momentum katsayısının işaret deęişimleri yoluyla ayarlamanın yönü üzerinde stabilize edici bir etkiyle algoritmayı düzenledięi anlamına gelir (*Haykin, 1999, s.170*).

Sonuç olarak, öğrenme oranının küçük seçilmesi iterasyon sayısını artırmakta ancak hata uzayında lokal minimumlara takılmayı engellemektedir (*Chan / Wong / Lam, 2000, s.1*). Shadra ve Patil (1992), yaptıkları deneysel çalışmalarda, en iyi performans  $\eta = 0,1$  ve  $\alpha = 0,5$  olarak belirlenen durumda elde edildiğini belirtmektedirler (*Shadra / Patil, 1992, s.493*). Geri yayılım algoritmasının kaç iterasyon sonra durdurulması gerektiğine dair çok genel kabul görmüş bir kurallar bütünü yoktur. Ayrıca iterasyon sayısının çok küçük seçilmesi yakınsama için sorun teşkil edebilmekte iken çok büyük seçilmesi ise, algoritmayı felce uğratmakta yada sürekli bir kararsızlığa sürüklemektedir (*Loh / Fong, 1993, s.592*). Pratikte birçok çalışmada ortalama kare hata ölçütündeki deęişim 0,01'e düştüğü durumda algoritmanın durdurulma işlemi gerçekleştirilmektedir.

Özet olarak, geri yayılım algoritmasının iki safhasının olduğunu bunların ileri hesaplama ve geri hata yayılımı olduğunu ve bu iki safhada da tüm hesaplamaların ağırlık ve yan deęerleri üzerinden doğrusal olarak gerçekleştiğini ve bir algoritmanın hesapsal karmaşıklığının ölçüsünün yapılan toplama çarpma ve depolama işlemleri olduğunu bu nedenle de hesapsal karmaşıklığın iterasyonlarla hesapsal etkinliği sağladığını belirterek, geri yayılım algoritmasının tüm bu nedenlerden dolayı hesapsal etkinliğe sahip bir algoritma olduğunun söylenmesi gerekmektedir (*Anthony / Bartlett, 1999, s.301*).

## B. DOĞRUSAL OLMAYAN EN KÜÇÜK KARELER

Daha önce tahmin konusuna girişte değinildiği gibi,  $A(\theta^r)$  şeklinde verilen matrisin farklı yollarda oluşturulması yaklaşımı, doğrusal dışı en küçük kareler başlığı altında çeşitli algoritmaların oluşturulmasına yol açar. Bu algoritmaların kendi aralarında ve geri yayılım algoritmasına karşı bazı üstün ve zayıf yönleri bulunmaktadır. Yukarıda tartışılan geri yayılım algoritmasının da bu başlıkta söz edilecek algoritmalara karşı hem belirgin bir üstünlüğü hem de zaafı  $A(\theta^r)$  matrisi yerinde birim matris kullanmasıdır. Öte yandan, geri yayılım algoritması ağırlıkların güncellenmesi için n. iterasyonda,

$$\Delta w(n) = -\eta \frac{\partial \varepsilon}{\partial w}$$

eşitliğini kullanmaktaydı. Bu güncellemede aslında eldeki tek bilgi gradient vektörü  $\frac{\partial \varepsilon}{\partial w}$ 'den gelmektedir. Ancak hata uzayı hakkında lokal bir bilgi olan bu bilgi algoritma için hesapsal kolaylık ve hatta etkinlik sağlarken diğer taraftan yakınsama hızı için bir handikap oluşturmaktadır. Ayrıca, başlangıç ağırlıklarına güvendiğinden ve lokal bilgi kullandığından lokal minimumlara takılabilmektedir (Yu, Efe, Kaynak, 2002, s.252). Aslında bu sorunun iki nedeni bulunmaktadır. Bunlardan birincisi, aktivasyon fonksiyonunun türevlenmesinden kaynaklanmaktadır. Eğer çıktı, aktivasyon fonksiyonunun ekstrem değerlerine yakın olursa (0 veya 1 yada -1 veya 1), aktivasyon fonksiyonunun türevlenmiş değeri ile üretilecek hata sinyali (ağın çıktısı gerçek değerden çok farklı olsa bile) küçük olacaktır (Ng / Leung / Luk, 1996, s409). İkinci neden ise, sadece birinci türevle ilgilenildiğinden, herhangi bir optimizasyon probleminde optimumu bulmak için sağlanması gereken *yeterli ve gerekli koşul* ilkelerinden sadece gerekli koşulla (birinci türevin sifıra eşitliğiyle) ilgilenilmektedir. Böylece lokal minimumlara takılma problemi yaşanabilmekte ve özellikle geniş ölçekli problemlerde bu önemli bir soruna yol açabilmektedir. Aslında momentum katsayısı ile genelleştirilmiş delta kuralının ve böylece hata uzayı hakkında ikinci bir bilginin kullanılması da bu iki nedenden dolayıdır.

*Newton Rapson* algoritması, geri yayılım algoritması sunulurken kullanılan ve ortalama kare hatayı tanımlayan  $\varepsilon_{av}$  matrisinin ağırlık vektörüne göre ikinci türevi alınarak elde edilen Hessian (H) matrisini bu amaçla kullanmaktadır (*Madsen / Nielsen / Tingleff, 2004, s.35*). Diğer bir ifade ile,

$$A(\theta^r) = H = \frac{\partial^2 \varepsilon_{av}}{\partial w^2}$$

şeklindedir.

Doğrusal olmayan en küçük kareler yönteminde Newton metodu Hessian matrisi ile bu bilgiye ulaşmaktadır. Böylece, Hessian matrisinin tersi ve gradient vektörünün çarpımı ile elde edilen ağırlık güncellemesi,

$$\Delta w(n) = \left( \frac{\partial^2 \varepsilon_{av}}{\partial w^2} \right)^{-1} \left( \frac{\partial \varepsilon_{av}}{\partial w} \right)$$

şeklinde elde edilir. Hata uzayını geri yayılım algoritmasının ilerisinde kuadratik bir fonksiyon ile ele almayı başaran bu yöntemin de bazı handikapları vardır. Çünkü bu matris tersi alınabilir bir matris olmalıdır. Ancak H matrisinin tekil olması tehlikesi her zaman vardır. Bu durumda Newton algoritması başarısızlığa uğrayacaktır (*Madsen / Nielsen / Tingleff, 2004, s. 24*).

Bu tip problemlerden kaçınmanın bir yolu  $H^{-1}$  matrisinin hesaplanmayıp doğrudan tahmin edilmesi mantığı ile hareket eden *Quasi Newton* metodunun kullanılmasıdır. Bu yaklaşımda gradient vektörü cinsinden bir yön vektörü tanımlanarak ağırlıkların güncellenmesinde (n+1). iterasyonda bu yön vektörünün kullanılması esas alınmaktadır (*Hagan / Demuth / Beale, 1995, s.9-15*).

$$w(n+1) = w(n) + \eta(n)s(n)$$

şeklindeki bu ağırlık güncelleme eşitliğinde s(n), n. iterasyon için kullanılan yön vektörünü gösterir. s(n),

$$s(n) = -S(n) \frac{\partial \varepsilon_{av}}{\partial w}$$

şeklindedir. Burada  $-S(n)$  düzenlenmesi  $(n-1)$ . iterasyondan sonra yapılmış pozitif tanımlı (karakteristik kökleri pozitif) bir matristir. Sonuçta gradient vektörü ve ağırlık vektörünün birlikte iterasyonlarının gerçekleştirilmesiyle elde edilmektedir. Şöyle ki,

$$q(n) = \frac{\partial \varepsilon_{av}(n+1)}{\partial w(n+1)} - \frac{\partial \varepsilon_{av}(n)}{\partial w(n)}$$

ve

$$\Delta w(n) = w(n+1) - w(n)$$

olmak üzere,

$$q(n) \cong \left( \frac{\partial q(n)}{\partial w} \right) \Delta w(n)$$

olacaktır. Böylece, Hessian matrisin tersi olan,  $\left( \frac{\partial^2 \varepsilon_{av}}{\partial w^2} \right)^{-1}$  değerini hesaplamak yerine ,

$$H^{-1} = [\Delta w_1, \Delta w_2, \Delta w_3, \dots, \Delta w_n][q(0), q(1), q(2), \dots, q(n)]^{-1}$$

şeklindeki tahmini kullanılarak yapılır (*Haykin, 1999, s. 243*). Öte yandan Oren(1973), Quasi Newton yaklaşımını yön vektörü hesabında çok karmaşık ve kısıtlayıcı varsayımlarda bulunduğu gerekçesiyle eleştirmektedir (*Oren, 1973, s.426*). Ancak, Irwin, Lightbody ve McLoone yaptıkları ortak deneysel çalışmada quasi Newton metodunun yakınsama için daha hızlı olduğunu belirtmişlerdir (*Irwin /Lightbody / McLoone,1994, s.2*).

Son olarak doğrusal dışı en küçük kareler yöntemi için bir birlerine yakın iki algoritma, *Gauss Newton* ve *Levenberg Marquardt* algoritmalarıdır. Bu iki

algoritmada Quasi Newton algoritması gibi bir taraftan hata uzayı hakkında daha çok bilgi kullanmaya çalışırken diğer taraftan ise Hessian matrisin tersini almaktan kaçınılmaktadırlar. Bunu yaparken de Jacobian (J) matrisi kullanılmaktadır. Bu matris de yine kare hata olarak tanımlanan  $\varepsilon_{av}$  matrisinin ağırlık vektörüne göre birinci türevi ile elde edilir. Burada aslında Hessian matrisi,

$$H = J^T J$$

şeklinde yakınsanabileceği mantığıyla hareket edilmektedir (*Madsen / Nielsen / Tingleff, 2004, s. 22*). Böylece Gauss Newton metodu, Newton Rapson metoduna göre bir üstünlük olarak ikinci türevlerle uğraşmayı engelleyebilmektedir. Ancak elbetteki asıl problem olan Hessian matrisinin tersinin hesaplanması uğraşından kurtulunamamıştır. Çünkü  $J^T J$  şeklinde oluşturulan ve Hessian matrisine yakınsayan bu yeni matrisin de tersi alınamayabilir. İşte Levenberg Marquardt algoritması,  $\mu$ , öğrenme parametresi ve I birim matris olmak üzere,

$$G = H + \mu I$$

şeklinde bir modifikasyonla yakınsanan H matrisini aynı zamanda tersi alınabilir bir matris haline de getirir. Şöyle ki,  $(\lambda_1, \lambda_2, \dots, \lambda_n)$  ve  $(z_1, z_2, \dots, z_n)$  sırasıyla H matrisinin özdeğer ve özvektörlerini göstermek üzere,

$$Gz_i = (H + \mu I)z_i = Hz_i + \mu z_i = \lambda_i z_i + \mu z_i = (\lambda_i + \mu)z_i$$

olduğundan, H matrisi ve G matrislerinin özvektörleri aynıdır. Öte yandan G matrisinin özdeğerleri  $(\lambda_i + \mu)$  olacağından,  $\mu$ 'nün artışı G matrisini pozitif belirli kılacak ve G matrisi tersi alınabilir bir matris olacaktır (*Madsen / Nielsen / Tingleff, 2004, s. 25*). Böylece, algoritmanın işleyişi şu şekilde olur. İlk olarak, hatalar üzerinden hata kareler toplamı hesaplanır. Geri yayılım algoritmasında olduğu gibi, zincir kuralından yararlanılarak, hata vektörünün ağırlık vektörüne göre birinci kısmi türevi alınırken kullanılacak ara kısmi türev olan hata vektörünün net girdi vektörüne göre birinci kısmi türevi Marquardt duyarlılığı



adını alır. Marquardt duyarlılıkları ile J matrisi elemanları elde edilir. Bu adımdan sonra,

$$\Delta w_{t+1} = w_t - [J^T J + \mu I]^{-1} J^T \varepsilon_t$$

şeklinde güncelleme değeri elde edilir.  $\mu=0$  durumunda algoritma Gauss Newton algoritmasına dönüşecektir (*Wiliamovski / Chen, 1999, s.1725*). Levenberg Marquardt algoritması ileri beslemeli ağlar için en hızlı algoritma olarak bilinmektedir (*Qi, 2001, s.386*). Ancak, her bir iterasyonda  $J^T J$  matrisinin hesaplanıp tersinin alınması gerekliliği özellikle büyük ağ yapıları için hesapsal karmaşıklığı arttırmakta, bu nedenle algoritma pratik bir algoritma olarak görülmemektedir (*Wiliamovski vd., 2001, s.1178-1179*).

#### IV. AĞ SEÇİMİ

Daha önce de belirtildiği gibi, oluşturulan ağın yapısı, ağın öğrenme performansı, öğrendiklerini genelleştirme performansı ve örneklem dışı tahmin performansı üzerinde son derece önemlidir. Öte yandan ağın ne tip bir yapıda olacağı konusunda da çok kesin kurallar bulunmamaktadır. Burada ağın yapısı ile kast edilen, ağın kaç katmandan oluşacağı ve katmanlardaki nöron sayısının ne olacağı hususlarıdır. Bu konulara açıklık getirilmesi için bu başlık altında öncelikle katman sayısı, evrensel yakınsama ve genelleştirme konuları ele alınacak daha sonrada teorik bilgi kriterleri ve bazı genel yaklaşımlar tartışılacaktır.

Yapay sinir ağları üzerine yazılan ilk yayınlara bakıldığında ve bunlara getirilen kritikler değerlendirildiğinde büyük kuşku yaratan konular yapay sinir ağlarının yakınsama yeteneklerinin ne olduğu, katman sayısı ve katmanlardaki nöron sayılarının belirlenmesinde bir rehber olmayışı konuları olduğu görülmektedir. Bir başka ifade ile ağın dizaynı ve konfigrasyonun subjektif oluşu, ele alınan problem, araştırmacının bilgi beceri ve yeteneği ile bu konfigrasyonun

şekillenmesi sıkıntı yaratmaktadır. Ağ dizaynının bir bilimden çok sanata benzetilmesi bu yüzdendir (*Zhang / Patuwo / Hu., 1998, s.42*). Öte yandan daha öncede değinildiği gibi problem çözme tabanlı mühendislik çalışmaları için bu bir problem teşkil etmemekte ancak istatistik bilimi için büyük bir sorun olmaktadır. Ancak, 1989-1992 yılları arasında bir dizi çalışma bu alandaki şüpheleri büyük ölçüde gidermiştir. Cybenko (1989), Hornik(1989), Hornik vd. (1990) ve Hornik (1991) çalışmalarıyla ileri beslemeli ağların yeterli sayıda katman ve katmanlarda yeterli sayıda nöron içermesi durumunda uygun biçimde ayarlanmış parametreler ile isteğe bağlı seçilmiş herhangi bir fonksiyona yakınsayabileceği ortaya konulmuştur. Aynı araştırmacıların 1991-1993 yılları arasındaki bir dizi çalışmasıyla tek ve iki katmanlı ağlarında yeterli sayıda nöron ve düzgün bir şekilde ayarlanmış parametre ile isteğe bağlı seçilmiş herhangi bir fonksiyona yakınsayacağı ispatlanmıştır. Yine Hornik (1991) ve Baron (1991) de isteğe bağlı seçilmiş fonksiyonların türevlerine de ileri beslemeli ağların yakınsadığını ortaya koymuştur. Tüm bu gelişmeler, bilinmeyen bir fonksiyon ve türevlerine, veriden elde edilen bilgi ile evrensel yakınsama kabiliyeti ileri beslemeli ağlar için çok önemli bir teorik sonuçtur (*Gençay / Selçuk / Witcher, 2003, s.274*). Özellikle transfer fonksiyonlarının yakınsaması tartışılırken verilen keyfi bir fonksiyona yakınsama yaklaşımı, bu kez  $y$  bağımlı ve  $x$  bağımsız değişkenleri üzerinden düşünülecek olursa, klasik istatistiksel yaklaşımla bilinmeyen koşullu ortalama,

$$M(x) = E(y|x) \text{ için,}$$

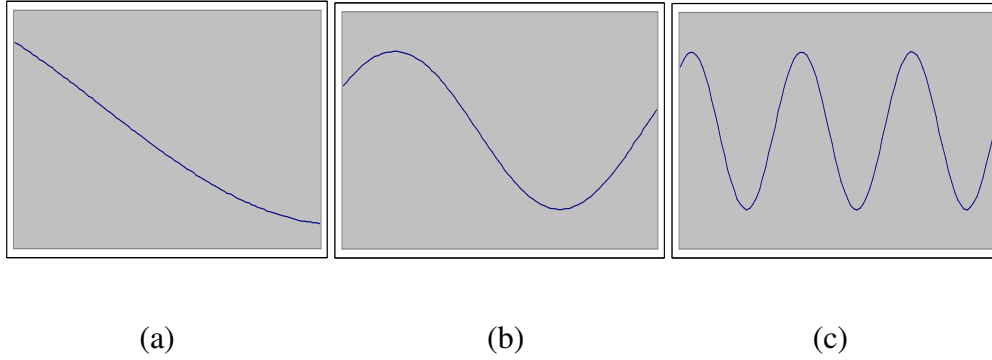
$$E|M(x) - f(x, \theta)|^2 < \varepsilon$$

olacaktır (*Kuan / Liu, 1995, s.349*). Yukarıda sayılan çalışmaların yapay sinir ağları için ortaya koyduğu bu gerçek, bu ağların yakınsama oranının, yakınsama için kullanılan parametrelerdeki artış oranından küçük olduğu ve kullanılması gereken parametre sayısının artışının doğrusal olduğudur (*Baron, 1991, s.1069*). Bu sonucun başlıca iki önemi vardır. Bunlardan ilki, yakınsamanın yapay sinir ağları ile sağlanabilmesi iken ikincisi ise diğer tekniklerin yakınsama için ihtiyaç duyduğu parametre sayısının üstel olarak artmasından dolayı yapay sinir ağlarının

bu konuda bir adım öne çıkmasıdır (Kuan / Liu, 1995, s.349).

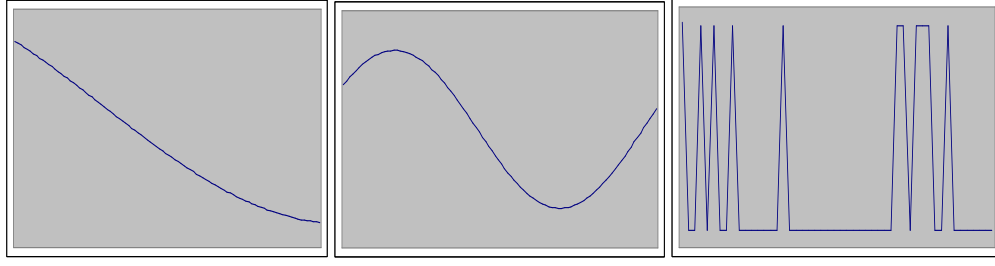
Bu gelişmeler sonucu ileri beslemeli bir ağ için katman sayısına karar vermede herhangi bir problem yoktur. Tek katmanlı bir ağ ile dahi yeterli sayıda nöron kullanarak isteğe bağlı seçilmiş herhangi bir fonksiyona yakınsama *teorik* olarak gerçekleştirilebilmektedir. Bu durumda karar verilecek problem katmanlardaki nöron sayısının ne olacağıdır. Bu başlık altında sadece gizli katmanlardaki nöron sayısının belirlenmesi üzerine yapılacak tartışma, bu çalışmanın uygulaması zaman serileri analizi üzerine yapılmış olduğundan dolayı, zaman serisi analizi ışığında açıklığa kavuşturulacak konular olup, yapay sinir ağları ile yeniden bakış başlığı altında ele alınacaktır.

Aşağıda Şekil 23'ün a, b ve c kısımlarında sırasıyla,  $y = \cos(x+2)$ ,  $z = \cos(3x+2)$ ,  $h = \cos(9x+2)$ , fonksiyonlarının  $[-1,1]$  aralığındaki çizimleri sunulmuştur. Açık ki bu üç fonksiyonun ilgili aralıktaki karmaşıklığı  $x$ ' in katsayısı arttıkça artmaktadır.



**Şekil 23 : Belirli Bir Aralıkta Fonksiyon Karmaşıklığı**

Aşağıda Şekil 24'de 1-3-1 ileri beslemeli bir ağ eğitilerek her bir fonksiyonun modellenmesiyle elde edilmiş çizimler sunulmuştur.



(a)

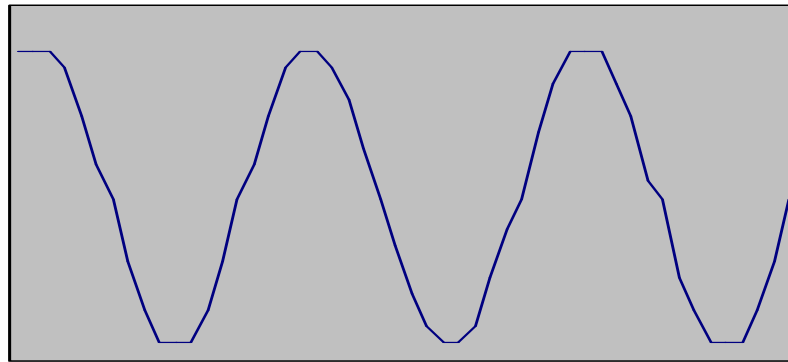
(b)

(c)

**Şekil 24 : 1-3-1 Ağ ile Çeşitli Fonksiyonlara Yakınsama**

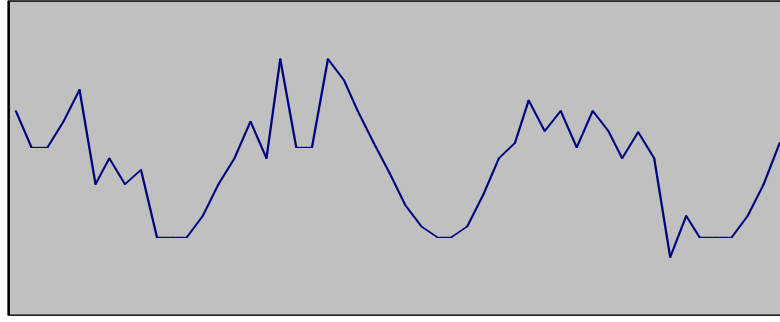
Şekil 24'den de açıkça görülebileceği gibi fonksiyon uzayının karmaşıklığının artmasıyla 1-3-1 ileri beslemeli ağın yakınsama kabiliyeti özellikle şeklin (c) kısmında ortadan kalkmaktadır.

Aşağıda Şekil 25'de 1-3-1 ileri beslemeli ağ ile yakınsama sağlanamamış olan  $h = \cos(9x+2)$  fonksiyonunun 1-5-1 ağı ile modellenmesi sonucu elde edilen çizim sunulmuştur



**Şekil 25 : 1-5-1 Ağı ile  $h = \cos(9x+2)$  Fonksiyonuna Yakınsama**

Yukarıda verilen bu örnekler gizli katmandaki nöron sayısının artmasıyla daha kompleks fonksiyonlara yakınsamanın başarılabilceğini ortaya koymaktadır. Diğer taraftan aşağıda şekil 26 ise aynı fonksiyona yani,  $h = \cos(9x+2)$  fonksiyonuna 1-10-1 ileri beslemeli ağı ile yakınsamayı göstermektedir.



**Şekil 26 : Genelleştirme Problemi**

1-10-1 ileri beslemeli ağı, Şekil 26'da 1-5-1 ileri beslemeli ağı ile sağlanan yakınsamadan çok daha kötü bir performans ortaya koymaktadır. Bunun nedeni ağın genelleştirme problemi yaşamasıdır. 1-10-1 ileri beslemeli ağı 21 ağırlık ve 11 yan değeri olmak üzere toplam 32 tane ayarlanabilir parametre içermektedir. Ağ eğitim setindeki örnek üzerinde oldukça başarılı performans gösterip test setindeki veriye uyum sağlayamamaktadır. Bu probleme aşırı uyum adı verilmektedir. Test setindeki yeni veriye uyum sağlayamama problemi yaşayan böyle bir ağ tahminleme içinde kullanılamayacak bir ağıdır. Çünkü gizli katmandaki nöron sayısını arttırmak sadece ve sadece eğitim setine uyumu yükseltmekte ancak ağın genelleştirme yeteneğini geliştirmemektedir (*Schiffmann / Werner, 1994, s.35*). Bu örnek açıkça göstermektedir ki yakınsamayı sağlayan

daha küçük ağırlar varken büyük ağırlarla çalışmaktan kaçınılmalıdır.<sup>2</sup>

Sonuçta katman sayısı problemi teorik olarak aşılmasına rağmen katmanlarda yer alacak nöronların sayısının belirlenmesi için kriter bulma zorunluluğu vardır.

## A. TEORİK BİLGİ KRİTERLERİ

Ağın dizaynının oldukça önemli bir konu olduğu ve sadece gizli katmandaki nöron sayısının artırılmasıyla ağın yakınsama özelliğinin ortaya çıkmasını beklemenin yanlış olduğu açıktır. Fonksiyon uzayındaki karmaşıklığın ağ ile yakalanması yanında aşırı uyum probleminden kaçınılması gerekmektedir. Bu nedenle dene gör yöntemi gibi subjektifliğe açık bir yaklaşım yerine belli başlı kriterlerle bu konunun açıklığa kavuşturulması önem arz etmektedir. Böylece farklı araştırmacıların farklı ağ seçimi yada farklı ağa güvenme karmaşası engellenebilir.

### 1. Schwarz Bilgi Kriteri (SIC)

G. Schwarz (1978) tarafından geliştirilmiş, istatistik ve ekonometri literatüründe sıklıkla kullanılan bir bilgi kriteridir. Aşağıdaki şekilde hesaplanmaktadır.

$$SIC = \log\left[\frac{1}{N} \sum_{T=1}^{\mu} (y_t - \hat{y}_t)^2\right] + \frac{w}{N} \log N$$

Burada,  $y_t$  hedeflenen değer,  $\hat{y}_t$  ağ tarafından hesaplanan değer,  $w$  modeldeki parametre sayısı ve  $N$  gözlem sayısıdır. En küçük SIC değerini veren

---

<sup>2</sup> Benzer bir örnek  $y = 1 + \sin(p \pi / 4)$  fonksiyonu için Hagan, Demuth, Beale "Neural Network Desing" (1995,11-21)'de bulunabilir.

model tercih edilecek model olacaktır. Dikkat edilirse SIC aslında iki terimden oluşmaktadır. Bunlar  $[\frac{1}{N} \sum_{T=1}^{\mu} (y_t - \hat{y}_t)^2]$  şeklindeki ortalama kare hatayı veren birinci terim ve  $\frac{w}{N} \log N$  şeklindeki ikinci terimlerdir. Bu ikinci terim daha az parametre ile basit tahmin modelinin daha iyi olduğunu söylemektedir. Eğer iki modelde aynı sayıda parametreye sahipse bu durumda ortalama kare hatası küçük olanın seçilmesi gerektiği açıktır (*Gençay / Selçuk / Witcher, 2003, s.295*).

## 2. Akaike Bilgi Kriteri (AIC)

H. Akaike (1973) tarafından ortaya konulmuş ve yine istatistik ile ekonometri literatürlerinin sıklıkla kullandığı bir bilgi kriteridir. Aşağıdaki şekilde hesaplanmaktadır.

$$AIC = \log\left[\frac{1}{N} \sum_{T=1}^{\mu} (y_t - \hat{y}_t)^2\right] + \frac{2}{w}N$$

Buradaki parametreler SIC'daki parametreler ile aynıdır. Bunun yanı sıra AIC'de ki terim içermekte ve birinci terim yine ortalama kare hatayı göstermektedir. En küçük AIC değerine sahip olan model tercih edilecek olan modeldir. Eğer aynı sayıda parametreye sahip iki model varsa yine ortalama kare hatası küçük olan model seçilecektir.

Bu iki bilgi kriterinden SIC, AIC'ye göre daha sert bir ceza uygulamaktadır, özellikle büyük örneklerde cimri bir model seçmek için daha uygun bir kriter olarak tanınır (*Arifovic / Gençay, 2001, s.575*). Qi ve Zhang (2001), SIC'ın AIC'ye göre daha güvenilir olduğunu ileri sürmelerine rağmen, yapay sinir ağlarında ağ seçimi için iki kriterinde tek başlarına aslında yeterince güvenilir olmadığını belirtmektedirler.

## B. DİĞER YAKLAŞIMLAR

Yukarıdaki bu bilgi kriterleri dışında bir dizi teknik daha bu amaçla kullanılmaktadır. Bunlar içerisinde önde gelen üç yaklaşım: Çapraz Geçerlilik, Erken Durdurma Metodu ve bazı Hata Kriterlerini kullanmaktadır<sup>3</sup>.

### 1. Çapraz Geçerlilik

Bu uygulamada veri seti üç parçaya ayrılır. Bunların adları sırasıyla eğitim, geçerlilik ve test örnekleridir. Verili bir ağ ile önce eğitim setindeki verilerle ağ eğitilir. İkinci olarak, geçerlilik setindeki veri ile değerlendirme yapılarak aşırı uyum olgusuna karşı ağ test edilir. Son aşamada ise test örnekleme ile ağın genelleştirme özellikleri ölçülür. Aslında veri setinin bu şekilde parçalanması, optimal ağ yapısına karar vermeyi ve buna bağlı olarak önraporlama performansını etkilemekte olmasına rağmen, bu parçalama için genel kabul görmüş oranlar söz konusu değildir (*Zhang / Patuwo / Hu, 1997, s.50*). Sıklıkla kullanılan ayırım %70 eğitim, %20 geçerlilik ve %10 test seti ayırımıdır. Öte yandan literatürde M – yarışması olarak bilinen ve 1982 yılında yapılan önraporlama yarışmasında aylık veriler için son 18, çeyrek yıllık veriler için son 8 ve yıllık veriler için son 4 gözlem, test seti için ayrılmıştır.

Bu metodun birçok farklı varyasyonu bulunmaktadır. Nakamura (2005) ve Gonzalez (2000) ağ seçimi için en uygun yöntem olarak bu yöntemi önermektedirler. Bu yaklaşım genelde veri seti yeterince geniş olduğunda başvurulan bir yöntemdir. Zaten kimi araştırmacılar tarafından geçerlilik seti veriyi boşa harcamak anlamına da gelmektedir (*Ripley, 1994, s.46*).

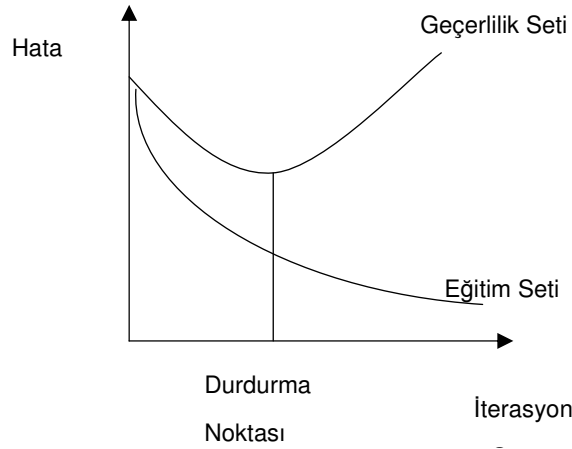
---

<sup>3</sup> Bunlar dışında burada ele alınmayan ve ayrı bir uzmanlık gerektiren genetik algoritmalar ve bayesgil tekniklerden de yararlanılmaktadır.



## 2. Erken Durdurma Metodu

Bu metodun mantığı, arzulanan genelleştirme özelliklerine sahip ağı elde edilmesi için en uygun zamanda eğitimin sonlandırılması amacıyla eğitimin öğrenme eğrisinin oluşturulup gözlemlenmesi üzerine kurulmuştur. Eğitim setindeki hata iterasyonlar boyunca azalacaktır. Ancak aşırı uyumun gerçekleştiği durumda eğitim setindeki hata düşmeye devam ederken geçerlilik setindeki hatanın artmaya başlayacağı açıktır. Buna göre geçerlilik setindeki hatanın artmaya başladığı noktadan önceki iterasyonda eğitim kesilirse bu durumda elde edilen ağırlıklarla optimal ağ dizaynı yakalanmış olur (*Gonzalez, 2000, s.11*). Aşağıda şekil 27’de bu durum gösterilmiştir.



**Şekil 27 : Erken Durdurma Metodu**

Öte yandan bu metot üzerinden daha önce yukarıda tartışılan durdurma kriteri konusunda da bazı çıkarımların yapılabileceği açıktır. Erken durdurma metodu optimal ağ dizaynı ve kullanılan tahmin algoritmasının işleyiş mantığı arasında karşılıklı etkileşime vurgu yapmaktadır. Özellikle çapraz geçerlilik yöntemi ile birlikte kullanıldığında hem optimal ağ dizaynını yakalamak için hem de aşırı uyum problemine yakalanmamak için en gelişmiş yöntemdir.

### 3. Hata Kriterlerinin Kullanımı

Genellikle ortalama kare hata (MSE), ortalama karekök hata (RMSE) ve Ortalama mutlak yüzde hata (MAPE) ölçütleri ağ seçimi için kullanılan hata ölçütleridir. Aşağıdaki şekilde hesaplanırlar.

$$MSE = \frac{\sum_{t=1}^n e_t^2}{n}$$

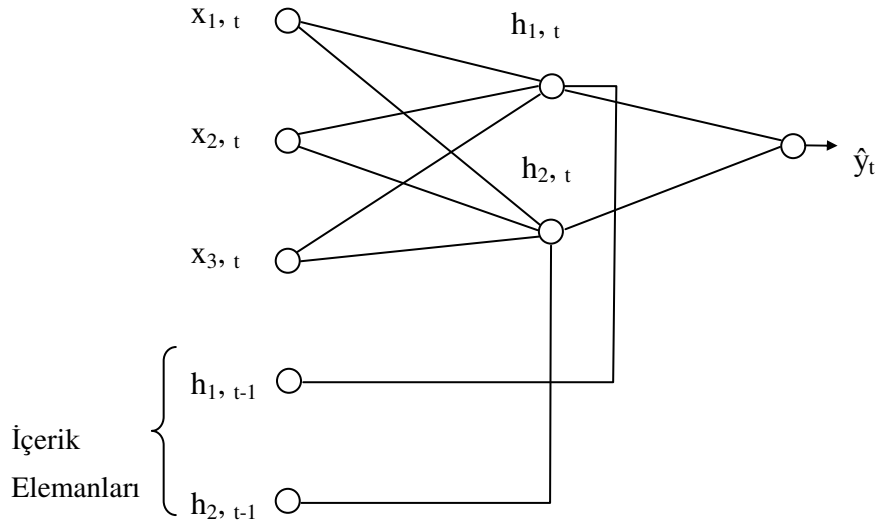
$$RMSE = \left( \frac{\sum_{i=1}^T (y_i - \hat{y}_i)^2}{T} \right)^{1/2}$$

$$MAPE = \frac{1}{T} \sum_{i=1}^T \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

Öte yandan, bu ölçütlere göre seçilen ağ yapıları genellikle zaman serisindeki dönme noktalarını yakalayamamaktadır. Ayrıca bu ölçütler modelde yer alacak parametre sayısının fazla olmasını cezalandırmamaktadır. Qi ve Zhang (2001), MSE, RMSE ve MAPE ölçütleri tek başına ağ seçimi için kullanılmayacağını belirtirlerken, Kaastra ve Boyd (1996), yapay sinir ağlarının ortalama mutlak yüzde hatayı (MAPE) minimize edemeyeceğini ileri sürmektedir. Bu eleştirilere rağmen MSE ve RMSE ölçütü, ağ seçimi için literatürde en yaygın kullanılan kriterlerdir. Temelde, Qi ve Zhang (2001)'in asıl eleştirisi hem ağ seçiminin hem de önraporlama performansının beraberce, bu yukarıda sayılan kriterlerle (AIC, SIC, MAPE, RMSE) yakalanamadığı üzerinedir. Bu nedenle, bu yaklaşımların ötesinde bazı kriterlere ihtiyaç duyulduğunu ileri sürmekte ve öğrenme – genelleştirme ikilemi üzerine geniş araştırmalar yapılması gerektiğini belirtmektedirler (Qi / Zhang, 2001, s. 678).

## V. GERİ DÖNÜŞÜMLÜ AĞLAR

Buraya kadar ele alınan çok katmanlı algılayıcılar aslında daha genel bir ağ yapısı olan geri dönüşümlü ağların özel bir halidir. Bununla birlikte bu çalışmada ele alınan geri dönüşümlü ağda, çok katmanlı algılayıcılar için önceki başlıklarda tartışılan teorik çıkarımsamalar geçerlidir. Çünkü, bu çalışmada (ve genelde zaman serisi çalışmalarında) kullanılan geri dönüşümlü ağların içerik elemanlarının geriye doğru olan bağlantı ağırlıkları eğitilen ağırlıklar değildir. Eğitim sürecinde sürekli olarak aynı değere sahiptirler. Diğer bir ifade ile eğitilmemektedirler. Bu nedenle, bu tip geri dönüşümlü ağların tanıtımında, teorik olarak bu çalışmada MLP'ler için yapılan teorik açıklamaların ötesinde yeni açılımlara gerek duyulmamaktadır<sup>4</sup>. Aşağıdaki şekilde kısmi geri dönüşümlü bir ağ sunulmuştur.



**Şekil 28 : Kısmi Geri Dönüşümlü Ağ**

<sup>4</sup> Bu çalışmada ele alınan ağlar kısmi geri dönüşümlü ağlar olarak da adlandırılmaktadır. Kısmi geri dönüşümlü ağların dışında yapay sinir ağları literatüründe kullanılan ve Tam geri dönüşümlü ağlar olarak da adlandırılan ağlar mevcut olup, geri dönüşüm bağlantıları eğitilmektedir. Ancak bu tip ağlar zaman serisi analizinde kullanılmadıklarından burada ele alınmamıştır

Çok katmalı algılayıcılardaki yapıya ek olarak, bu ağ yapısında *içerik elemanları* farklılık yaratmaktadır. Şekil 28’de sunulan ağ yapısı Elman ağı olarak da adlandırılmaktadır. Elman ağları en basit geri dönüşümlü ağ yapısı olarak bilinir ve ayrıca *gizli geri dönüşümlü* ağlar olarak da adlandırılırlar. Aslında bu ikinci isimleri çalışma prensiplerine dayanılarak verilmiştir. Çünkü bu ağlarda ara katmanda (gizli katmanda) bulunan nöronların çıktıları bir zaman gecikmeli değer olarak tekrar kendi girdilerinden bir bölümünü oluştururlar. Bu nedenle gizli katmanda bulunan nöronlar kendi performanslarının kayıtları ile birlikte yada bir başka ifade ile kendi kısa dönemli içsel hafızaları ile birlikte yeni verileri işleme alırlar (*Gençay / Selçuk / Witcher, 2003, s.293*). Bu durum ilgili yapının altında yatan dinamiklerin öğrenilmesi için bir avantaj sağlamaktadır. Çünkü filtrelenmiş olarak ele alınan bu bilgi, ek bir açıklama gücünü de beraberinde getirir. Böylece içerik duyarlılığı ve hafıza ile birlikte daha dinamik bir yapı arz eder (*Gençay / Liu, 1997, s.123*) İleri beslemeli bir ağ ve geri dönüşümlü bir başka ağ için, aralarındaki farklılığın ortaya konulması açısından, verilebilecek matematiksel açıklamalar sırasıyla,

$$y_t = s \left( \beta_0 + \sum_{i=1}^q \beta_i g \left( \alpha_{i0} + \sum_{j=1}^n \alpha_{ij} x_{j,t} \right) \right) + \varepsilon_t$$

ve,

$$y_t = s \left( \beta_0 + \sum_{i=1}^q \beta_i g \left( \alpha_{i0} + \sum_{j=1}^n \alpha_{ij} x_{j,t} + \sum_{l=1}^q \delta_{il} h_{l,t-1} \right) \right) + \varepsilon_t$$

şeklindedir. Öte yandan geri dönüşümlü ağlar literatürde doğrusal dışı otoregresif hareketli ortalama (NARMA) yada dışsal girdili doğrusal dışı otoregresif model (NARX) olarak ele alınmaktadır. Moshiri , Cameron ve Scuse (1999), özellikle doğrusal dışı zaman serisi analizinde geri dönüşümlü ağların kompleks veri setlerinde daha başarılı sonuçlar verdiğini ve bu nedenle de, geri dönüşümlü ağların daha zengin bir dinamik yapıya sahip olduğunu belirtmektedirler (*Moshiri / Cameron / Scuse , 1999, s.227*).

Son olarak, geri dönüşümlü ağı tahmin prosedürü, MLP'lerde kullanılan teknikle elde edilebilir. Ancak burada dikkat edilmesi gereken iki özel durum sırasıyla, ara katmanın net girdi değeri hesaplanırken içerik elemanlarından gelen kısmın unutulmaması ve hatanın geri yayılımında ise içerik elemanlarının geri bağlantılarının dikkate alınmamasıdır. Daha öncede belirtildiği gibi içerik elemanlarının bağlantıları eğitilmemektedir. Başlangıçta ara katmandaki aktivasyon değeri bilinmediğinden, içerik elemanlarının başlangıç değeri olarak, genel bir yaklaşım, bir ara katmanın alabileceği maksimum değer yarısının içerik elemanlarının başlangıç değeri olarak atanmasıdır (*Öztemel, 2005, s.167*).

## **VI. YAPAY SİNİR AĞLARI İLE ZAMAN SERİSİ ANALİZİNE YENİDEN BAKIŞ**

Bu çalışmada buraya kadar, yapay sinir ağlarının tanıtımının ve tarihsel gelişiminin tanıtılmasını amaçlayan ikinci bölüm bir kenara bırakılırsa, birinci bölüm ve bu bölümde ele alınan konular, uygulama bölümü için bir alt yapı hazırlama amacını taşımaktadır. Bu başlık altında amaç hem ağ seçimi başlığı altında bu başlığa bırakıldığı açıklanan girdi ve çıktı katmanlarındaki nöron sayısına karar verilme probleminin, hem de zaman serisi analizindeki anahtar özelliklerin yapay sinir ağları ile nasıl ele alınacağına ortaya konulmasıdır.

Basitçe özetlemek gerekirse, bir zaman serisi ile çalışıldığında, ilk olarak serinin birinci bölümde tartışılan anahtar özelliklerden hangisi yada hangilerine sahip olduğunun belirlenmesi gerekir. İkinci olarak uygun gecikme yapısına karar verilir. Bu iki aşama sonucunda veri yaratma süreci tanımlanmış olur. Bu noktadan itibaren açıklanması gereken serinin, ilgili veri yaratma süreci altında birim kök içerip içermediğidir. Eğer seri ( $y_t$ ) birim kök içermiyorsa uygun bulunan veri yaratma süreci ile, eğer birim kök içeriyorsa da fark serisi ( $\Delta y_t$ ) kullanılarak tek değişkenli modelleme ve buna bağlı olarak önraporlama elde edilir.

İkinci olarak yine basitçe özetlemek gerekirse, modelleme aracı olarak yapay sinir ağlarının kullanılması durumunda, başlangıçta cevap verilmesi gereken sorular; girdi katmanında kaç nöron kullanılacağı, ara katman sayısının ne olacağı, ara katmanlarda kaç nöron kullanılacağı , çıktı katmanının da kaç nöron olacağı ve bu nöronların hangi transfer fonksiyonlarını kullanacağı ile eğitimde hangi algoritmanın tercih edileceğidir.

Buradan da anlaşılacağı gibi, yapay sinir ağlarının iktisadi zaman serilerinde modelleme aracı olarak kullanımı, kendilerine özgü sorunlara sahip iki ayrı alanın ortaklaşa ele alınmasını gerektirmektedir. Özellikle geleneksel olarak niteleyebileceğimiz zaman serisi analiz yöntemlerinin işin en başında kabul ettikleri fonksiyonel formun yanlış olabilme durumu yada veri setindeki sıradışı gözlemlerin tüm parametre tahminlerini yanlış kılma tehlikesi yapay sinir ağı modellerini güvenilirlik konusunda bir adım ileri çıkarmaktadır (Hill, O'Connor, Remus, 1996, s.1082). Öteyandan, yapay sinir ağlarının parametrik olmayan doğası ve tüm prosesin bilgisayarda işlediğinin düşünülmesi nedeniyle, yapay sinir ağları ile zaman serisi önraporlamalarına kara kutu yakıştırması yapılması ise bir tehlikedir (Faraway, Chatfield, 1998, s.236). Bu tehlikeye ek olarak, her özel bir durum için ayrı bir kural ortaya koymak uygulamacılar açısından ciddi bir problem olmaktadır (Setino / Tong, 2004, s.240) Bu gibi nedenlerle ilgili karmaşaya adimsal bir prosedürle yaklaşılması önem taşımaktadır. Bu amaçla Kaastra ve Boyd (1996) 8 adımlık, Yao ve Tan (2002) 7 adımlık bir prosedür önermektedirler. Birbirlerinin kopyası olmasa da aralarında oldukça büyük benzerlikler bulunan bu iki prosedür, bu çalışmada uygulama bölümü için bir el haritası olmak üzere aşağıdaki şekilde 5 adımlık bir yaklaşım olarak tekrar düzenlenmiştir<sup>5</sup>. Adımların tek bir geçişlik olmadığı ve bazı durumlarda geriye dönülüp bazı düzenlemelerin tekrar yapılmasının gerekliliğini belirtmede yarar vardır (Kaastra / Boyd, 1996, s.219).

---

<sup>5</sup> Bu çalışmada tek değişkenli zaman serileri ile çalışıldığından, kullanılan adimsal prosedür çok değişkenli yaklaşımlara göre daha kısadır.

## **1.ADIM: VERİ ÖNİŞLEME**

Yapay sinir ağlarının veri işleme yeteneğinin ortaya konulması için veri setinin ağa sunum formatı son derece önemlidir. Ham veri seti arzulanan formatta olmayabilir. Burada arzulanan format kavramı hem zaman serisi literatüründen hem de yapay sinir ağları literatüründen beslenmektedir. Zaman serisi analizine dayanan kısım verinin trend ve mevsimsellik içermesi durumudur. Saharda ve Patil (1992), yapay sinir ağlarının mevsimsel düzeltme yapılmadığı durumda yapay sinir ağları ile etkin öngörüler yapılabildiğini belirtmiş olmalarına rağmen Farway ve Chetfield (1995), Nelson v.d (1999), Zhang ve Qi (2005) deterministik trendten arındırılmış ve mevsimsel düzeltmenin yapılmış olduğu serilerle çalışıldığında yapay sinir ağlarının çok daha yüksek önraporlama performansına sahip olduğunu belirtmektedirler. Hatta bazı durumlarda trend ve mevsimsellikte düzeltme yapılmadığında yapay sinir ağı modelleri, ARIMA modellerinden daha kötü performans göstermektedirler (Zhang / Qi, 2005, s.513). Öteyandan, Stock ve Watson (1998), Tkacz (2001), mevsimsel düzeltme yapıldığı durumda, yapay sinir ağının önraporlama performansının düşebileceğini öne sürmektedirler. Bunun nedeni olarak da X11 filitreleme yöntemini ile, seride arka plandaki doğrusal dışılığın ortadan kaldırılmasını göstermektedirler (Tkacz, 2001, s.58). İkinci olarak uygulama bölümünde ele aldığımız serilerde karşılaşmadığımız için bu çalışmada da tartışılmayan sıra dışı gözlemlerin (outlier), seride olması durumunda , Yao ve Tan (2002) bu gözlemlerin seriden dışlanması gerektiğini belirtmektedirler. Ancak Franses (2002), Yapay sinir ağlarının sıra dışı gözlemleri modellemede başarılı olduğunu belirtmekte ve ayrıca seride yapısal kırılma varsa yapay sinir ağları ile bu kırılmaların da yakalanabileceğini ileri sürmektedir (Franses, 2000, s.215).

Veri önışleme adımının yapay sinir ağı literatüründen beslenen kısmı ise, seride bir normalizasyon işlemine gerek duyulmasıdır. Normalizasyon işlemi ile verinin küçük bir aralıkta ele alınması yapay sinir ağı modelinin performansını arttırmaktadır. Örneğin veri setinin 0.1 - 0.9 gibi bir aralık içerisinde ele alınması, kullanılan transfer fonksiyonunun limitlerini zorlamadan veri setinin

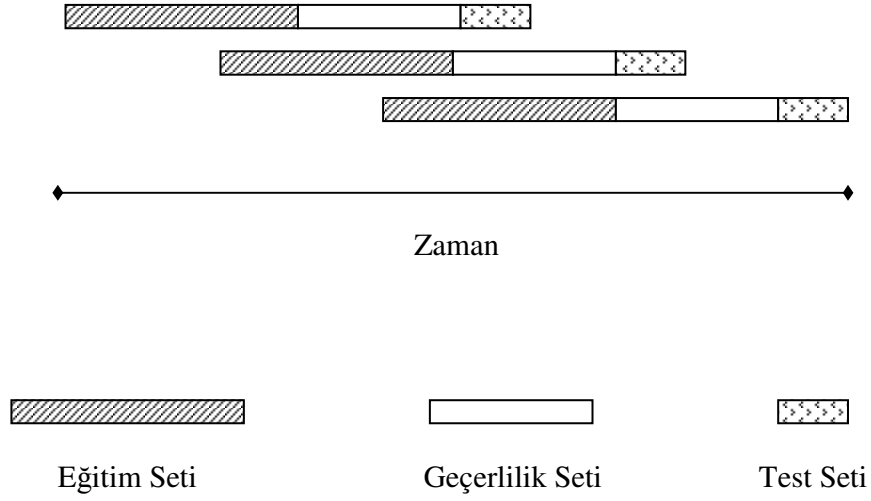
haritalanabilmesini sağlamaktadır (Cheng / Wagner / Lin, 1996, s.2). Eğer normalizasyon yapılmaz ise, girdi yada çıktı değerleri, yapay sinir ağı ile ele alınamayacak kadar büyük olabilir ve özellikle bu durum kullanılan yapay sinir ağı modelinde katman sayısı fazla ise daha büyük bir problem teşkil eder (Yao/ Tan, 2002, s.192) .

## **2. ADIM: VERİ ORGANİZASYONU**

Veri setinin ağa sunumunun organizasyonu, aslında bu çalışmanın ikinci bölümünde ağ seçimi başlığı altında çapraz geçerlilik yöntemi tartışılırken sunulmuştur. Burada buna ek olarak eğer ele alınan zaman serisi birim kök içeriyorsa nasıl bir yaklaşımda bulunulabileceği tartışılacaktır. Veri yaratma dinamiğinin altında yatan süreç eğer zaman içerisinde değişiyorsa, veri yaratma sürecine ilişkin tanımlanan parametrelerin tahminlerine de bu serbestliğin tanınması gereklidir. Eğer parametre tahminine bu serbestlik tanınmaz ise yapay sinir ağı parametreleri genelleştirme özelliklerini yitirirler (Gençay / Selçuk / Whitcher, 2002, s301).

Kayan pencere yöntemi, hareketli pencere yaklaşımı, ileri doğru yürüyen pencere yöntemi gibi adlarla adlandırılan yaklaşım, yapay sinir ağları ile durağan olmayan serilere ilişkin önraporlamaların elde edilmesinde bu amaçla kullanılmaktadır. Yöntemin mantığı aşağıda şekil 29'da gösterildiği gibi, eğitim, test, ve geçerlilik setlerinin zaman içerisinde ileriye doğru kaydırılmasına dayanır. Böylece lokal durağan pencerelerden (örneklemelerden), tekrar tekrar elde edilen yapay sinir ağı parametreleri veri yaratma sürecindeki zamana bağlı değişimi içerecek şekilde elde edilmiş olur.





**Şekil 29: Kayan Pencere Yöntemi**

Aslında bu yöntem gerçek zamanlı modellemelerde yeni gözlemler elde edildikçe eski gözlemlerin dışarıda bırakılması şeklinde kullanılmaktadır (Kaastra/ Boyd, 1996 s.224). Teorik olarak eğer seri durağan ise uzun veri seti modelin önraporlama performansını geliştirmekte ve aşırı uyum probleminden kaçınmaya yardımcı olmaktadır (Zhang, 2001, s.1119). Veri seti durağan değil ise, veri setinin kısa olması daha iyi sonuçlar vermekte ve ayrıca uzun veri setlerinin bu gibi durumlarda modelin önraporlama performansını bozduğu ileri sürülmektedir (Morants / Whalen / Zhang, 2004, s. 253-254).

Ele alınan zaman serisinin tüm gözlemlerinin veri yaratma süreci altında yatan dinamikler ve önraporlama üzerinde eşit derecede pay sahibi olmadığı ve son elde edilen gözlemlerin eski gözlemlerden daha önemli olduğu düşüncesinden hareketle Morants, Whalen ve Zhang (2004) hareketli pencere yöntemine benzeyen, yeni bir yöntemi ileri sürmüşlerdir. Ağırlıklandırılmış pencere yaklaşımı adlı bu yeni yöntem önraporlama hatalarına ağırlık verme (yeni hatalara daha yüksek, eski hatalara daha düşük) mantığıyla çalışmaktadır. Eğitim setinde “çekirdek” olarak adlandırılan en yeni gözlem grubuna %100 ağırlık verilirken, geri kalan gözlemler doğrusal olarak azalan şekilde ağırlıklandırılmaktadır.

Böylece t zaman endeksi, s eğitime alınan veri setinin başlangıç tarihine ait zaman endeksi, c ise çekirdek olarak adlandırılan gözlem grubunun ilk gözleminin zaman endeksi olmak üzere,

$$W_t = \begin{cases} 1 & , & t \geq c \text{ ise} \\ (t-c)/(t-s) & , & s \leq t < c \text{ ise} \\ 0 & , & t < s \text{ ise} \end{cases}$$

şeklinde ağırlıklar elde edilir. Sonuçta kalıntı kareler toplamı (SSE), ağırlıklandırılmış kalıntı kareler toplamı (WSSE) olur ve

$$WSSE = \sum_{t=s}^N W_t (y_t - \hat{y}_t)^2$$

şeklinde ifade edilir. Bu durumda eğitimde amaç WSSE'yi minimum kılacak yapay sinir ağı modelinin parametrelerini elde etmek olacaktır. Morantz, Whalen ve Zhang 7 ayrı veri seti üzerinde ağırlıklandırılmış pencere yaklaşımı ile elde edilen önraporlamaların performanslarının diğer yaklaşımlardan daha iyi sonuç verdiğini belirtmektedir.

### **3. ADIM: AĞ MİMARİSİNİN OLUŞTURULMASI**

Daha önce ikinci bölümde değinildiği gibi, oluşturulacak ağın kaç katman içermesi gerektiği konusu 1989-1993 yılları arasında Cybenko, Hornik, Funushasi vd. gibi araştırmacılar tarafından teorik olarak cevaplanmış ve tek katmanda dahi düzgün ayarlanmış parametreler ile yeterli sayıda nöron kullanılarak keyfi bir fonksiyona yakınsama sağlanabileceği gösterilmiştir. Ancak uygulamada bu teorik sonuçlar sadece bir rahatlık sağlamakta, ele alınan problemin tam çözümü için kesin ve net bir bilgi sunmamaktadır. Bundan dolayı ağ mimarisi dene ve gör yaklaşımıyla, problemin çözümünü veren benzer ağ sonuçlarının kıyaslanmasıyla elde edilmektedir. Bu kıyaslama için ise yine ikinci bölümde verilen kriterler kullanılmaktadır. Bu kıyaslamamanın önemi şu şekilde

belirtilebilir. Daha öncede tartışılan aşırı uyum olgusu ile karşılaşma durumunun yapay sinir ağlarında çok sık olarak karşımıza çıkma nedeni, eklenecek her bir nöronun, regresyonda eklenen her bir yeni parametreden çok daha fazla sayıda, serbestlik derecesi üzerinde baskıya yol açmasıdır. Çünkü her bir yeni eklenen nöron kendinden önceki ve sonraki katmanda bulunan tüm nöronlar ile bağlantı ve bunun sonucunda da ağırlık taşıyacağından, serbestlik derecesi üzerinde, tüm bu tahmin edilmesi gereken ağırlık ve artı bir de yan değeri kadar azalmaya yol açar. Hem katman hem de nöron sayısının artışı serbestlik derecesini düşürüp aşırı uyum problemine neden olur (Kaastra ve Boyd, 1996, s.225). Gerekenden az sayıda nöron kullanmak ise, ağırlık yakınsama yeteneğinin ortaya konulamaması anlamına gelecektir.

Öte yandan ikinci bölümde girdi ve çıktı katmanlarındaki nöron sayısının ne olacağının cevabı hatırlanacağı gibi bu kısma bırakılmıştı. Yapay sinir ağları ile çalışıldığında, ilgili ağda girdi katmanında kaç nöron kullanılacağına karar verme, modellemeye ilişkin en önemli basamaklardan biridir (Zhang, 2004, 11). Zaman serisi önraporlamaları için yapay sinir ağlarının kullanılması durumunda girdi katmanındaki nöron sayısı, kaç gecikme kullanılması gerektiğiyle ilgili bir konudur. Bu nedenle girdi katmanındaki nöron sayısı gecikme uzunluğuna bağlı olarak tespit edilir. Kullanılacak her bir gecikme için girdi katmanına bir nöron eklenmelidir (Zhang, 2001, s.1186). Kullanılacak gecikme uzunluğunun belirlenmesi için ise başlıca iki yol vardır<sup>6</sup>. Bunlardan birincisi sırası ile gecikme uzunluklarını artırıp modelin önraporlama performansını kontrol ederek, önraporlama performansında gelişmenin durduğu durumda gecikme sayısının ne olacağına karar vermektir. Bu yaklaşım geleneksel yapay sinir ağı yaklaşımı olarak adlandırılır. İkinci yol ise daha önce tanıtılan bilgi kriterlerini bu amaçla

---

<sup>6</sup> Üçüncü bir yol olarak, eğitimde kullanılacak algoritmanın en çok olabilirlik yöntemine dayandırıldığı durumda, Mederios, Terasvirta ve Rech (2002) tarafından geliştirilen bir yol söz konusu olup, en çok olabilirlik yönteminde olabilirlik fonksiyonuna cezalandırıcı bir terim eklenmesindeki problemlerden dolayı ağırlık yakınsama sorunu yaşamasının sıklıkla karşılaşıldığı bu yöntem, tahmin konusunda ele alınmadığından uzantısı olan yolda burada kullanılmamıştır. İlgili yöntem, Mederios, Terasvirta ve Rech (2002)'de bulunabilir.

kullanmaktır, bu yaklaşım ise Box Jenkins metodolojisinin yaklaşımıdır. BuHamra, Smaoui ve Gabr (2003), çalışmalarında, bilgi kriterlerinin kullanıldığı durumda ortalama hatanın %24 daha düşük elde edildiğini belirtmektedirler (BuHamra / Smaoui / Gabr, 2003, s.814).

Girdi katmanındaki nöron sayısı özellikle önraporlama performansı için ara katmanlardaki nöron sayısından çok daha hayati bir konudur (Zhang, 2001, 1199). Çıktı katmanındaki nöron sayısı ise, önraporlama da seçilecek yaklaşıma göre şekillenir. Daha öncede belirtildiği gibi tek adımlı yada çok adımlı önraporlamalar elde edilebilir. Eğer tek adımlı önraporlama yapılacaksa, tek bir nöron , eğer çok adımlı önraporlama yapılacak ise, adım sayısı kadar nöron çıktı katmanında yer almalıdır (Zhang, 2001, 1187).

Son olarak, hangi tip transfer fonksiyonu kullanılacağına karar vermede çok keskin kurallar olmasa da genelde, ağın ortalama davranışı öğrenmesi isteniyorsa lojistik, ortalamadan sapmanın öğrenilmesi isteniyorsa hiperbolik tanjant fonksiyonu, transfer fonksiyonu olarak kullanılmaktadır.

#### **4.ADIM: EĞİTİM**

Ağın eğitilmesi başlığı bu çalışmanın ikinci bölümünde detaylı olarak tartışılmıştır. Literatürdeki uygulamalı çalışmalar göstermiştir ki doğrusal dışı minimizasyon yapıldığında algoritmanın başlangıcı için atanan değerler sonucu etkilemektedir. Burada bir ekleme olarak, tüm bu algoritmaların minimize etmeye çalıştıkları ve daha önce hata enerjisi olarak tanımlanan ortalama kare hata yerine önraporlama konusunda sunulan diğer hata ölçütlerinin de bazı çalışmalarda kullanıldığını belirtmede yarar vardır.

## **5.ADIM: DEĞERLENDİRME**

Bu adım, yukarıdaki adımları sadece bir ağ için değil bir grup ağ için uygulanmasının önemine işaret eder. Özellikle önraporlama için, kullanılacak ağ tek bir ağ olmayıp, bir grup ağ olmalıdır (*Yao / Tan, 2002, s.194*). Bu grup içerisinde performansı en yüksek ağ değerlendirme adımında seçilecektir. Bu seçimde önraporlama performans ölçütlerinin ve bilgi kriterlerinin kullanılacağı açıktır. Gonzalez (2000), klasik istatistik ve ekonometrik modellemede, önraporlama adımının parametre tahminlerinden tamamen ayrı bir adım olduğunu, yapay sinir ağları modellerinin önraporlama performanslarının daha yüksek olmasının ana sebebinin ise bu değerlendirme adımında bu tür bir bilgi kullanmalarından kaynaklandığını belirtmektedir (*Gonzalez, 2000, s.11*). Öte yandan daha önce birinci bölümde belirtilen önraporlama yöntemlerinden *bağımsız yöntemde* ağdaki parametrelerin sayısı önraporlama yapılacak dönemin sayısına göre değişebilmektedir. Klein, (2004), bağımsız yöntemin iteratif yöntemle göre daha üstün olduğunu belirtirken (*Klein, 2004, s.230-231*), Heravi v.d (2004), o kadar kesin olmasa da, en azından, iteratif yöntemin en doğru model spesifikasyonu seçilse dahi gecikmeli değerlerin özellikle doğrusal dışı modeller ve Yapay sinir ağları modellerinde, doğrusal dışı bir şekilde modele girmesinden dolayı, küçük önraporlama hataları üretmediğini söylemektedirler (*Heravi, Osborn, Birchenhall, 2004, s.439*).

## DÖRDÜNCÜ BÖLÜM

# TÜRKİYE EKONOMİSİNE AİT DEĞİŞİK ZAMAN SERİSİ ÖZELİKLERİ GÖSTEREN MAKRO EKONOMİK SERİLER ÜZERİNDE YAPAY SİNİR AĞI MODELLERİNİN PERFORMANSLARININ ÖLÇÜLMESİ

## I. YÖNTEM

Bu bölüm, birinci bölümde sunulan ve zaman serilerinin sahip olduğu anahtar özellikler olarak adlandırılan, trend, mevsimsellik, yapısal kırılma, volatilité ve doğrusal dışı özellikler durumlarında alışıla gelmiş olarak kullanılan zaman serisi modelleriyle, yapay sinir ağı modellerinin performanslarının karşılaştırılması üzerine uygulamalardan oluşmaktadır. Böylece, zaman serisi analizinde yapay sinir ağı modellerini kullanmanın uygun olabileceği durumların belirlenmesi hedeflenmektedir. Bu amaçla, Türkiye ekonomisine ait ve yukarıda sayılan anahtar özelliklere sahip beş ayrı zaman serisi ele alınmıştır. Uygulamalarda MATLAB, R, NeuroSolutions, E-Views ve Nettle programları ve bu programlar altında çalışan rutinler kullanılmıştır.

İlk olarak ele alınan seri özel nihai tüketim serisidir. Çeyrek yıllık olarak ele alınan (1987:Q1- 2005:Q4) ve mevsimsel etkinin gözlemlendiği bu seri üzerinde, periyodik otoregresif model ve yapay sinir ağı modeli ayrı ayrı kurularak, bu iki modelin performansları değerlendirilmiştir. İkinci olarak ele alınan seri IMKB-100 endeksidir. Haftalık olarak ele alınan (31.7.1987 – 27. 4. 2007) bu seri üzerinde ise, Volatilité modellemesi başlığı altındaki ARCH - GARCH modellerinin performansı yapay sinir ağı modeliyle karşılaştırılmaktadır. Üçüncü olarak sanayi üretim endeksi serisi kullanılmıştır. Çeyrek yıllık olarak ele alınan (1980:Q1 – 2006:Q3) bu seri üzerinde yapısal kırılma test edildikten sonra uygun ARMA modeli ve yapay sinir ağı modelinin performansı karşılaştırılmıştır.

Dördüncü olarak, kapasite kullanım oranları serisi kullanılmıştır. Aylık olarak (Şubat 1991 – Ocak 2007) ele alınan bu seri üzerinde doğrusal dışı modellerin (TAR, LSTAR, ESTAR ve RNN) uygunluğu araştırılmıştır. Son olarak, ham petrol fiyatları serisi aylık olarak (Ocak1991- Aralık 2006) ele alınmıştır. Birim kök içeren bu seri üzerinde ağırlıklandırılmış pencere yöntemi kullanılarak yapay sinir ağı modelinin performansı basit bir AR modeli karşısında değerlendirilmiştir.

Uygulama bölümünün beş alt başlığındaki kullanılan yöntem şu şekildedir: İlk olarak ele alınan zaman serisinin hangi anahtar özelliklere sahip olduğu uygun testlerle test edilerek belirlenmiştir. Bu adım *tanımlama* adımı olarak da adlandırılabilir. İkinci olarak belirlenmiş özelliklere sahip zaman serisine uygun modeller seçilerek, bu modellerin parametreleri *tahmin* edilmiştir. Bu tahmin adımında yalnızca bir model tahmin edilmeyip uygun olabileceği düşünülen bir grup modelin tahmin edilmiş olması üçüncü adım olarak adlandırabilecek *diagnostik kontrolü* adımında, model seçim kriterlerinin de yardımıyla en iyi modelin belirlenmesine olanak sağlamıştır. Dördüncü adımda en iyi olduğuna karar verilmiş modelden *önraporlama* değerleri elde edilmiş ve bu önraporlama değerlerine ilişkin önraporlama performans ölçütleri hesaplanmıştır. Bu dört adım sonunda ele alınan seri, zaman serisi özellikleri belirlenmiş, bu özelliklere göre en uygun modelle modellemesi ve önraporlaması yapılmış ve daha sonra da önraporlama performans ölçütleri hesaplanmış olarak analiz edilmiş haldedir. Daha önce de belirtildiği gibi bu tez çalışmasının problemiği olarak ortaya konulan “yapay sinir ağları zaman serisi analizinde bir analiz aracı olarak kullanılabilir mi?” sorusunun cevabının alınması için farklı zaman serisi özelliklerine sahip bu seriler yapay sinir ağı modelleri kullanılarak da modellenmiştir. Yapay sinir ağı modellemesinde üçüncü bölümde “yapay sinir ağları ile yeniden bakış” başlığı altında sunulan; veri önışleme, veri organizasyonu, ağ mimarisinin oluşturulması, eğitim ve değerlendirme adımlarından oluşan adimsal süreç kullanılmıştır. Elde edilen önraporlama

değerlerine ilişkin performans ölçütleriyle ile yukarıda dördüncü adımda zaman serisi modelleriyle elde edilmiş olan performans ölçütleri karşılaştırılmıştır. Böylece yapay sinir ağı modellerinin örneklem içi önraporlamalar için klasik teknikler karşısındaki başarısı ya da başarısızlığı ortaya konulmuştur. Öte yandan üçüncü bölümde de ayrıntılı olarak tartışılmış olan aşırı uyum problemi göz önüne alınarak, kullanılan yapay sinir ağı modellerinin örneklem içi önraporlamalarda aşırı uyum probleminden dolayı eşit ya da daha iyi performans ortaya koymuş olabileceği gerçeğinden hareketle son bir adım olarak örneklem dışı önraporlama karşılaştırmaları da her bir başlık altında ayrıca ele alınmıştır.

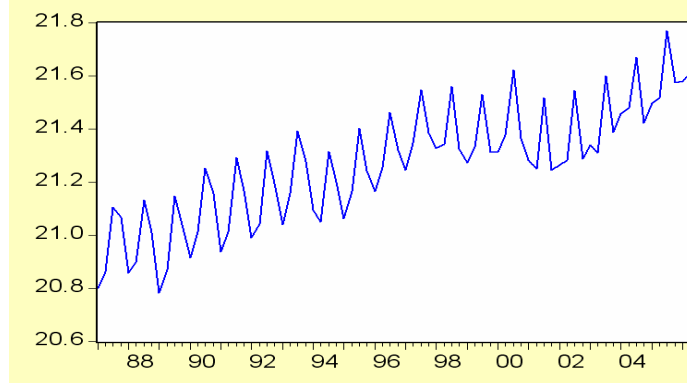
## **II. MEVSİMSELLİK İÇEREN BİR SERİ OLARAK ÖZEL NİHAİ TÜKETİM SERİSİNİN MODELLENMESİ**

Özel Nihai Tüketim, harcamalar yöntemiyle hesaplanan gayri safi milli hasılanın en önemli kalemidir. Özel sektör ve özel kişilerin nihai tüketim harcamalarından oluşur. Nihai tüketim harcaması olarak, dayanıksız tüketim mallarına (yiyecek, içecek vb.), yarı dayanıklı tüketim mallarına (elbise, ayakkabı vb.), dayanıklı tüketim mallarına (beyaz eşya vb.), hizmetlere (ulaşım, enerji vb.) ve son olarak konut sahipliğine yapılan harcamalar ele alınır. Türk ekonomisinin son on yılda ortalama %65'lik bölümü özel nihai tüketim harcamalarından oluşmaktadır. AB ülkeleri ile kıyaslandığında ise kişi başına düşen gelirin en az olduğu ülke olan Türkiye, gelire oranla hane halklarının en fazla harcama yaptığı ülkedir. Öte yandan, Dayanıklı tüketim mallarının harcamalar içindeki payı 2001 yılından 2005 yılına kadar yüzde 89 oranında büyüyerek yüzde 22'ye yükselmiştir. Bu durum, gelişmekte olan her ülkede gözlenebileceği gibi, tüketimin, GSMH'daki artış hızı kadar olmasa da artan bir şekilde devam ettiği ve GSMH'daki artış devam ettiği sürece tüketim kompozisyonunun dayanıksız tüketim mallarından yarı dayanıklı ve dayanıklı tüketim mallarına doğru dönüşeceği ve ancak gelir artışının belirli bir değerin üzerine ulaştığı durumda tasarrufların artmaya başlayabileceği şeklinde yorumlanmaktadır. Son olarak, hane halklarının tüketimlerinin ekonomik krizlerin hemen öncesinde en yüksek seviyeye ulaştığını ve Türkiyede 1994 ve 2001 yıllarında hane halkı tüketiminin



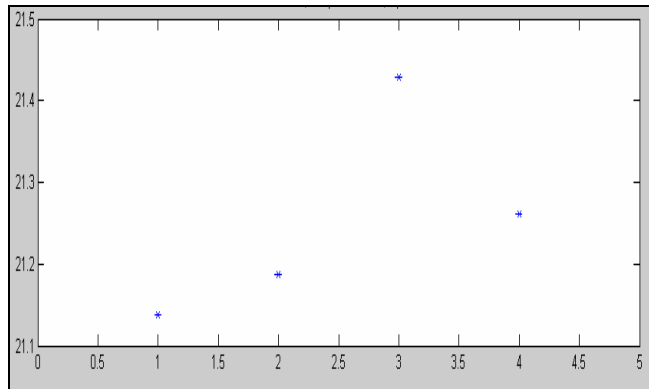
GSYİH'daki paylarının sırasıyla %70 ve %72 olduğunun belirtilmesinde fayda vardır.

Aşağıdaki şekil 30'da Türkiye için 1987:Q1- 2005:Q4 dönemi logaritmik özel nihai tüketim (ONT) serisinin zaman yolu grafiği görülmektedir.



**Şekil 30: Özel Nihai Tüketim (1987Q1 - 2005:Q4)**

Özel nihai tüketim serisindeki mevsimsel etki, belirli çeyreklerdeki maksimum ve minimumlar, şekil 30'da rahatlıkla görülebilmektedir. Mevsimsel etkinin ortaya konulması açısından çeyrekler bazındaki ortalamaları gösteren şekil 31 sunulmuştur.



**Şekil 31: Özel Nihai Tüketim: Çeyrekler Bazında Ortalamalar**

Çeyrekler bazında ortalamalara bakıldığında 3. çeyrek ortalamasının diğer ortalamalardan oldukça yüksek seyrettiği açıktır. Öte yandan 1. çeyrek ortalaması

ise en düşük ortalamayı oluşturmaktadır.

***İlk olarak***, serinin birim kök içerip içermediği araştırılmıştır. Üçüncü bölümde tartışıldığı üzere bu amaçla kullanılacak, Boswijk ve Franses (1996)'nın önerdiği en çok olabilirlik oranı testi,

$$LR = n \log (RSS_0 / RSS_1)$$

şeklinde Ki-Kare dağılımına sahip testtir. Burada,  $RSS_0$  ve  $RSS_1$  sırasıyla, kısıtlı ve kısıtsız modellerin kalıntı kareler toplamlarını göstermektedir. Kısıtlı model,

$$H_0 : \alpha_1 \alpha_2 \alpha_3 \alpha_4 = 1$$

hipotezinin getirdiği kısıt altında tahmin edilen,

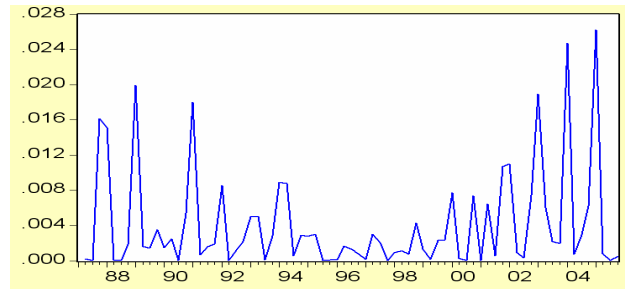
$$Y_t = \alpha_1 D_{1,t} Y_{t-1} + \alpha_2 D_{2,t} Y_{t-1} + \alpha_3 D_{3,t} Y_{t-3} + (\alpha_1 \alpha_2 \alpha_3)^{-1} D_{4,t} + \varepsilon_t$$

şeklindeki modeldir. Bu durumda kısıtsız model ise,

$$Y_t = \sum_{i=1}^4 \alpha_i D_{i,t} y_{t-1} + \sum_{s=1}^4 \mu_s D_s + \sum_{s=1}^4 \tau_s D_s t + \varepsilon_t$$

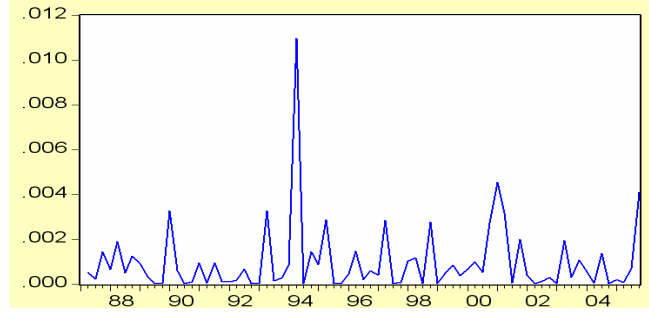
olacaktır. Aşağıda sırası ile, kısıtlı ve kısıtsız modellerin tahminleri ve kalıntı kareler toplamları sunulmuştur.

$$Y_t = 0.955 D_{1,t} Y_{t-1} + 1.002 D_{2,t} Y_{t-2} + 1.011 D_{3,t} Y_{t-3} + 0.992 D_{4,t}$$



**Şekil 32: Özel Nihai Tüketim: Kısıtlı Model Kalıntı Kareler**

$$Y_t = 0.997D_1y_{t-1} + 1.004D_2y_{t-1} + 1.011D_3y_{t-1} + 0.996D_4y_{t-1} \\ + 0.004D_1t - 0.001D_2t - 0.001D_3t - 0.002D_4t$$



**Şekil 33: Özel Nihai Tüketim: Kısıtsız Model Kalıntı Kareler**

Sonuçta, olabilirlik oranı test istatistiği  $LR = 76 \cdot \log(0.073 / 0.309) = 47.49$  olarak elde edilmiştir. Bu değer %1 anlamlılık düzeyinde 4 serbestlik dereceli  $\chi^2$  tablo değeri 13.28'den büyük olduğu için  $H_0$  birim kök hipotezi red edilir.

Bu aşamadan sonra uygun gecikme uzunluğuna karar verip modelin tahmin edilmesi aşamasına geçilebilir. Uygun otoregresif yapının belirlenmesi için üçüncü bölümde belirtildiği gibi AIC ve SIC enformasyon kriterleri kullanılmıştır.

**Tablo 6: Uygun PEAR (p) Modelinin Seçimi**

<i>Model</i>	<i>AIC</i>	<i>SIC</i>
PEAR(1)	-3,17	-2,92
PEAR(2)	-2,96	-2,90
PEAR(3)	-3,02	-2,82
PEAR(4)	-2,99	-2,64

Tablo 6'da de görüldüğü şekilde kullanılan iki bilgi kriteri de PEAR (1) modelini işaret etmektedir. ONT serisi için tahmin edilen PEAR(1) modeli,

$$\begin{aligned}
Y_{1,t} &= 0,125 + 0,12 y_{t-1} + \varepsilon_{1,t} \\
Y_{2,t} &= 0,212 + 0,74 y_{t-1} + \varepsilon_{2,t} \\
Y_{3,t} &= 0,147 + 0,92 y_{t-1} + \varepsilon_{3,t} \\
Y_{4,t} &= 0,095 + 0,61 y_{t-1} + \varepsilon_{4,t}
\end{aligned}$$

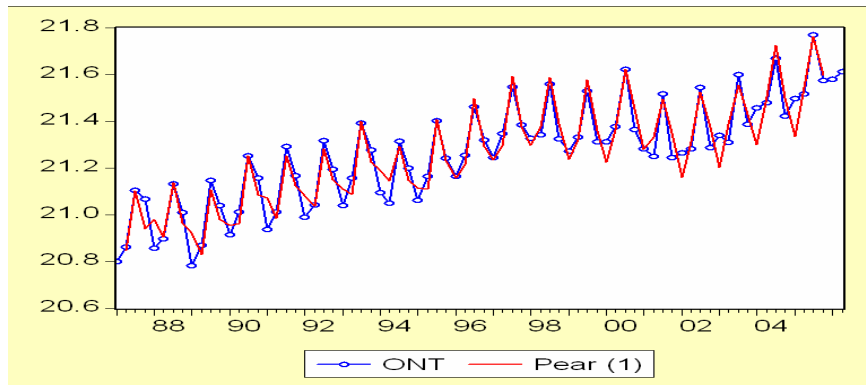
şeklindedir. Böylece PEAR(1) modelinin doğru model olduğu belirlendikten sonraki aşama olan otoregresif katsayılar da periyodik varyasyonun testine geçilebilir. Bu amaçla, yine kısıtlı ve kısıtsız model üzerinden elde edilen kalıntı kareler toplamlarının kullanıldığı F testi,  $F_{(3p,n-(s+sp))}$  dağılımına uyar ve

$$F = [(n-l)/s] / [(RSS_1 - RSS_0) / RSS_1]$$

şeklindedir. ONT serisi üzerinden elde edilen değerler yerine konulduğunda,

$$F = [(76-4)/4] / [0,310-0,073] / 0,310 = 23,60$$

değeri elde edilir. %1 anlamlılık düzeyinde  $F_{12,56}$  değeri ise 2,56 olduğundan, otoregresif parametrelerde periyodik varyasyonun varlığı kabul edilir. Böylece durağanlık kontrolü yapılmış, parametrelerinde periyodik varyasyonun varlığı ortaya konulmuş geçerli bir PEAR (1) modeli elde edilmiştir. Bu modele ilişkin örneklem içi önraporlama aşağıda şekil 34’te sunulmuştur.



**Şekil 34: Özel Nihai Tüketim – PEAR(1) Modeli Önraporlama Performansı**

Bu önraporlamaya ilişkin önraporlama performans ölçütleri ise aşağıda tablo 7’deki gibidir.

**Tablo 7: PEAR (1) Modeli Örneklem içi Önraporlama Ölçütleri**

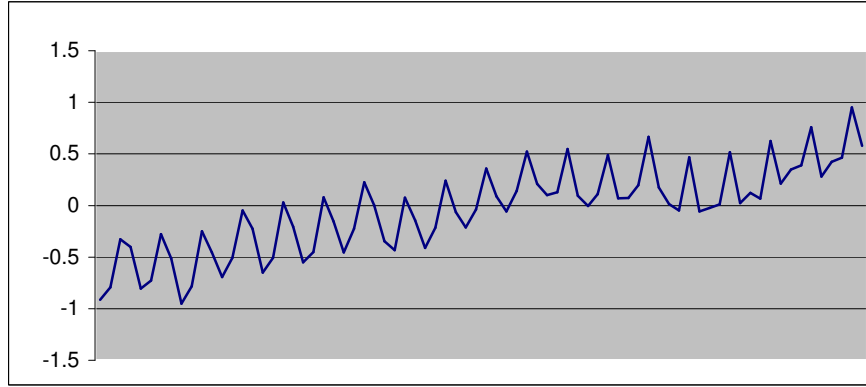
MAE	MAPE	MSE	RMSE
0.049964	0.002354	0.004075	0.556497

***İkinci olarak***, çok katmanlı algılayıcı ile mevsimsellik içeren ONT serisi modellenmiştir. Bu amaçla, üçüncü bölümde sunulan beş adımlık prosedür uygulanmıştır. Uygulamanın bu başlığı altındaki amaç, mevsimsellik içeren serilerin yapay sinir ağı modelleriyle ele alınması durumunda önraporlama performansında daha iyi sonuçlar elde edilip edilemeyeceğinin ölçülmesidir. Üçüncü bölümde yapay sinir ağları ile yeniden bakış – Veri önışleme adımı başlığı altında da değinildiği gibi Shadra ve Patil (1992), mevsimsel düzeltme yapılmassa da, iyi önraporlamalar elde edilebileceğini, Stock ve Watson (1998) ve Tkacz (2001), mevsimsel düzeltme yapıldığında yapay sinir ağı modellerinin önraporlama değerlerinin bozulacağını ileri sürmektedirler. Öte yandan Farway ve Chetfield (1995), Nelson v.d (1999), Zhang ve Qi (2005) mevsimsel düzeltmenin yapılmış olduğu serilerle çalışıldığında yapay sinir ağlarının çok daha yüksek önraporlama performansına sahip olduğunu belirtmektedirler. Hatta bu araştırmacılar mevsimsellikte düzeltme yapılmadığı durumda yapay sinir ağı modellerinin, ARIMA modellerinden daha kötü performans gösterdiklerini ileri sürmektedir. Bu nedenle, bu başlığın son bölümünde yukarıdaki PEAR modelinin önraporlama performans ölçütleriyle, bu ikinci adımda mevsimsel düzeltme yapılmamış seriden elde edilecek MLP modelinin önraporlama performans ölçütleri karşılaştırılacaktır.

Bu amaçla ilk olarak ONT serisi kullanılan transfer fonksiyonun limitlerini zorlamamak için normalizasyon işlemine alınmıştır. Hiperbolik tanjant fonksiyonu transfer fonksiyonu olarak kullanıldığı için aşağıdaki formüle veri seti  $[-1,1]$  aralığı içine alınmıştır.

$$\left[ \frac{X_i - \min(X_i)}{[\max(X_i) - \min(X_i)]} - 0.5 \right] \times 1.9$$

Bu normalizasyon işlemi sonrası ONT (NONT) serisi aşağıda şekil 35’de görülmektedir.



**Şekil 35: NONT Serisi**

İkinci olarak seride birim kök bulunmadığı PEAR modeli için test edildiğinden, veri yaratma dinamiğinin altında yatan sürecin zaman içerisinde değişmediği bilinmektedir. Bu nedenle bu çalışmada veri organizasyonu adımı altında anlatılan kısım atlanacaktır.

Daha önce de belirtildiği gibi uygulanan bu adımsal prosedürün adımları tek bir geçişlik değildir. Ara sonuçlarda geriye dönülüp bazı düzenlemelerin yapılması gerekir. Bu nedenle, ağ mimarisinin oluşturulması, tahmin ve değerlendirme adımları uygulamada eşanlı olarak ele alınacaktır. Aşağıda tablo 8’de NONT serisi üzerinde denen çeşitli ağların SIC ve bu ağ grupları içinde en iyi performans gösteren gecikme uzunluğuna sahip grubun, MSE ve MAPE ölçütleri verilmiştir. Öte yandan ağların eğitiminde ilk olarak Newton Rapson metodu denenmiş ancak H matrisinin tersi alınamamıştır, bu nedenle ağ Levenberg – Marquardt algoritması kullanılarak tahmin edilmiştir.

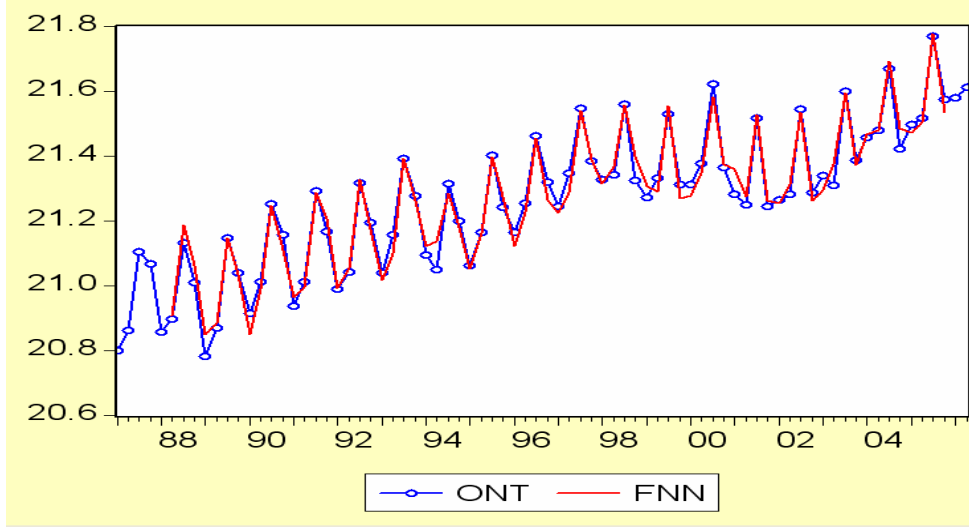
**Tablo 8: NONT Serisi İçin Ağ Seçimi**

AĞ ADI	SIC	MSE	MAPE	AĞ ADI	SIC	MSE	MAPE	AĞ ADI	SIC	MSE	MAPE
2-1-1	-182	---	---	3-1-1	-222	---	---	4-1-1	-303	---	---
2-2-1	-194	---	---	3-2-1	-248	---	---	4-2-1	-307	---	---
<b>2-3-1</b>	<b>-233</b>	<b>0.0307</b>	<b>1.72</b>	3-3-1	-249	---	---	4-3-1	-331	---	---
2-4-1	-227	---	---	3-4-1	-301	---	---	<b>4-4-1</b>	<b>-339</b>	<b>0.0059</b>	<b>2.088</b>
2-5-1	-219	---	---	3-5-1	-304	---	---	4-5-1	-326	---	---
2-6-1	-212	---	---	<b>3-6-1</b>	<b>-307.</b>	<b>0.0058</b>	<b>1.702</b>	4-6-1	-331	---	---
2-7-1	-205	---	---	3-7-1	-297	---	---	4-7-1	-324	---	---
2-8-1	-197	---	---	3-8-1	-298	---	---	4-8-1	-305	---	---
2-9-1	-193	---	---	3-9-1	-287	---	---	4-9-1	-338	---	---

**Tablo 8(Devam): NONT Serisi İçin Ağ Seçimi**

AĞ ADI	SIC	MSE	MAPE	AĞ ADI	SIC	MSE	MAPE
5-1-1	-331.	0.0075	1.087	6-1-1	-333	---	---
5-2-1	-340.	0.0072	0.993	6-2-1	-330	---	---
5-3-1	-330	0.0060	1.014	6-3-1	-334	---	---
<b>5-4-1</b>	<b>-361</b>	<b>0.0039</b>	<b>1.717</b>	<b>6-4-1</b>	<b>-340</b>	<b>0.0041</b>	<b>1.791</b>
5-5-1	-354	0.0041	1.151	6-5-1	-331	---	---
5-6-1	-345	0.0044	0.841	6-6-1	-333	---	---
5-7-1	-353	0.0040	1.202	6-7-1	-331	---	---
5-8-1	-322	0.0048	0.808	6-8-1	-298	---	---
5-9-1	-293	0.0046	0.811	6-9-1	-298	---	---

Tablo 8'e göre 5-4-1 ileri beslemeli ağı en iyi performansı gösteren ağıdır. 5 gecikme uzunluğu SIC'ye göre belirlenmiş, daha sonra 5 gecikme uzunluğuna sahip ağlar arasından ara katmandaki nöron sayısını belirlemek için MSE ölçütü kullanılmıştır. Burada bir ilginç durum her ağ grubu için son sütunda verilen MAPE ölçütünün en yüksek değerinin MSE ölçütünün en küçük olduğu ağ yapısında karşımıza çıkmasıdır. Ancak daha öncede değinildiği gibi Kaastra ve Boynd (1996), yapay sinir ağlarının MAPE'yi minimize edemediğini belirtmektedirler. Böylece bu çalışmada da MAPE ölçütünün kullanılmaması gerektiği savı desteklenmektedir. 5-4-1 ağının kullanılmasıyla elde edilen örneklem içi önraporlama aşağıda şekil 36'da görülmektedir.



**Şekil 36: Özel Nihai Tüketim 5-4-1 FNN Modeli Önraporlama Performansı**

Yukarıda şekil 36’de elde edilen önraporlama değerleri, ağıdan elde edilen değerlerin, çalışmanın başında yapılan normalizasyon işleminin tersinin gerçekleştirilmesi sonucu elde edilmiştir. Önraporlamaya ilişkin önraporlama performans ölçütleri de aşağıda tablo 9’da sunulmuştur.

**Tablo 9: 5-4-1 FNN Modeli Örneklem içi Önraporlama Ölçütleri**

MAE	MAPE	MSE	RMSE
0.025082864	0.001181	0.001109	0.29037

Tablo 4’teki ölçütler PEAR (1) modelinin örneklem içi önraporlamaları için hesaplanmış ölçütlerin verildiği tablo 7’deki değerlerle karşılaştırıldığında 5-4-1 ileri beslemeli ağının önraporlama performansının her ölçüt içinde daha iyi olduğu görülmektedir.

Bulunan bu ampirik sonuçlar, literatürdeki “mevsimsel serilerde yapay

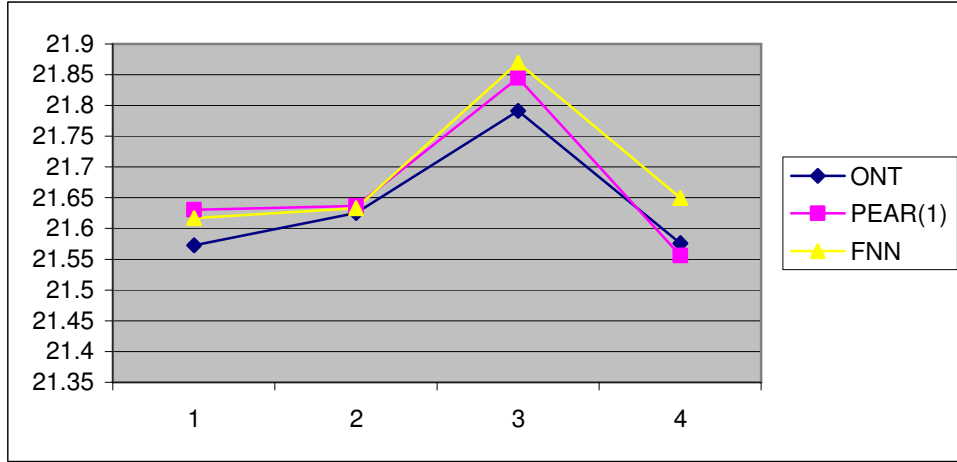


sinir ağı modelleri kullanılabilir mi ?” tartışmasında, kullanılabilir diyenlerin tarafını desteklemektedir. Ancak her ne kadar AIC ve MSE ölçütleriyle aşırı uyum problemi kontrol altında tutulmaya çalışılmış olsa da uygulamanın bu başlığı altında kullandığımız veri setinin küçük oluşu çapraz geçerlilik yöntemini kullanmamızı engellemiştir. Bu durum, ağın aşırı uyum probleminden dolayı örneklem içi önraporlamalarda yüksek doğruluğa sahip olduğu şeklinde yorumlanabilir. Bu nedenle ex post dönem önraporlaması önem arz etmektedir. Böylece modelin kurulmasında kullanılmamış, ancak gerçekleşmiş verilerle iki modelinde ex post dönem önraporlama performansları karşılaştırılacaktır. Bu dönem önraporlamalarının da 5-4-1 FNN modelinin lehine sonuç vermesi durumu, hem yukarıdaki 5-4-1 FNN modelinin örneklem içi önraporlamada aşırı uyum göstermediği hem de yapay sinir ağlarının mevsimsellik gösteren serilerde de kullanılabileceği yönünde Shadra ve Patil (1992), Stock ve Watson (1998) ve Tkacz (2001)’in görüşlerini destekler nitelikte bir sonuç olacaktır.

Aşağıdaki şekil 37 ve tablo 10 model kurulması için kullanılan örneklem kesildiği 2005 yılının son çeyreğinden sonrası dönemde gerçekleşmiş 4 özel nihai tüketim değerini ve PEAR (1) ile 5-4-1 FNN modellerinin ex post önraporlama değerlerini göstermektedir.

**Tablo 10: Özel Nihai Tüketim İçin Gerçekleşen Değerler ve Ex post Önraporlama Performansları**

Model	2006-Q1	2006-Q2	2006-Q3	2006-Q4	TMS
<i>5-4-1 FNN</i>	21.61316	21.62622	21.85108	21.6502	0.17545
<i>PEAR(1)</i>	21.63008	21.64162	21.84902	21.5564	0.11024
<i>GERÇEK</i>	21.57316	21.62517	21.79108	21.5658	-----



**Şekil 37: Özel Nihai Tüketim İçin Expost Önraporlama Performansları**

İlk olarak tablo 10'daki değerler üzerinden hesaplanan toplam mutlak sapma (TMS) değeri PEAR(1) modeli için 0.11024 iken 5-4-1 FNN için ise 0.17545 şeklinde elde edilmiştir. İkinci olarak, Tablo 10'daki değerler üzerinden Diebold Mariano (1995) testi ile FNN modelinin expost dönem önraporlama performansı ile PEAR(1) modelinin performansı arasında bir fark olup olmadığı test edilmiştir. Bu amaçla kullanılan Diebold ve Mariano'nun önerdikleri test aşağıdaki şekilde olup, boş hipotezi iki modelin önraporlamaları arasında istatistiksel açıdan bir farkın olmadığı yönündedir ve test standart normal dağılıma sahiptir.

$$S = \left(\frac{m}{4}\right)^{-\frac{1}{2}} \left[ \sum_{i=1}^m d_i - \frac{m}{2} \right]$$

Burada, m, önraporlama yapılan dönem sayısıdır.  $d_i$ , FNN modelinin önraporlama hatası karesinin, PEAR(1) modeli önraporlama hata karesinden büyük olduğunda 1, aksi durumda 0 değerini alır.

Yukarıdaki iki model için S istatistiği,

$$S = (4/4)^{-1/2} [0+0+1+1 - (4/2)] = 0$$

şeklinde değer alır. Bunun anlamı  $H_0$  hiotezinin red edilemeyeceğidir. Yani iki

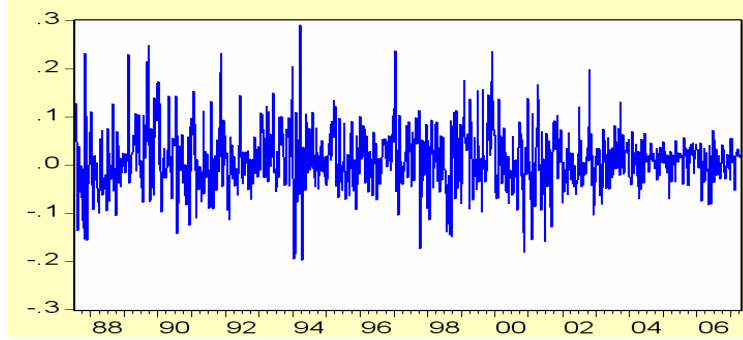
modelin önraporlamaları arasında istatistiksel açıdan bir fark yoktur.

Elde edilen bu sonuçlar, literatürde süregelen mevsimsellik içeren serilerde yapay sinir ağı modellerinin kullanılıp kullanılmayacağı tartışması için Türkiye ekonomisine ait gerçek bir serinin analizine dayalı cevap taşımaktadır. Ancak diğer taraftan bu cevabın sadece ONT serisi temelinde değerlendirilmesi gereğide açıktır. Kanımızca ele alınan ONT serisinde mevsimsel etki ile birlikte doğrusal dışı bir yapı mevcuttur. 5-4-1 FNN modelinin örneklem içi ve örneklem dışı önraporlama başarısında bizce bu doğrusal dışı yapının varlığı rol oynamaktadır. Literatürde mevsimselliğin ve doğrusal dışılığın birlikte ele alınabileceği bir test ve bir modelleme aracının olmaması bu tartışmanın dahada derinleştirilebilmesini engellemektedir. En azından bulunan bu sonuçlar *bazı durumlarda* yapay sinir ağı modellerinin mevsimsel etkiye sahip serilerde de iyi önraporlama değerleri veren bir modelleme aracı olarak kullanılabileceğini göstermektedir.

### III. İMKB-100 HAFTALIK GETİRİLERİ İÇİN VOLATİLİTE MODELLEMESİ

Rassal bir değişkenin zaman içerisindeki dalgalanması o değişkene ilişkin volatilitiyi ortaya çıkarır. Hisse senedi fiyatı ve özellikle de bu fiyatın oynaklığı iktisadi olarak birkaç farklı açıdan önem taşımaktadır. İlk olarak, hisse senedi fiyatındaki volatilitenin yüksekliği, hisse senedinin fiyatının doğru olmadığını bir göstergesi olarak düşünülebilir. Buna bağlı olarak sermaye piyasasının fonksiyonel rolü tartışmaya açık hale gelir. İkinci olarak, ilgili volatiliteler, opsiyon ve türev piyasalarının sergileyecekleri performans için de önemli bir göstergedir. Yatırımcılar ve şirketlerin özellikle orta vadeli projeksiyonlarına ve bu yolla da tavırlarına yön veren bu piyasaların, hisse senedi volatilitesine duyarlılığı son derece yüksektir. Üçüncü olarak, hisse senedi piyasasındaki volatiliteler toplam talep üzerinde de etkiye sahip olabilmektedir. Özellikle, finansal piyasalardaki dalgalanmaların ve negatif yönlü hareketlerin, hane halkı refahında ve buna bağlı olarak da tüketim harcamaları üzerinde kısıtlayıcı etkiye yol açabileceği ileri sürülmektedir. Ancak bu görüş, hisse senedi sahiplerinin daha çok yüksek gelir grubuna dahil kişilerden oluştuğu teziyle eleştirilirse, hisse senedi piyasasındaki volatilitenin artışı, riskin arttığı anlamına geleceğinden, fonların başka finansal araçlara kaymasını ve bu yollarda şirketlerin fonlama maliyetlerinde bir artışı ortaya çıkarabilmektedir. Bu durum da en azından yatırım kararlarının ertelenmesini beraberinde getirmektedir.

Yukarıda sayılan nedenlerden dolayı volatiliteler modellemesi iktisadi açıdan önem arz etmektedir. Aşağıda şekil 38’de İMKB 100 endeksi için 20.07.1987 – 9.04.2007 tarihleri arasında haftalık getiri serisinin zaman yolu grafiği sunulmuştur.



**Şekil 38 : İMKB 100 Haftalık Getiri (20.07.1987-09.04.2007)**

İlk olarak ADF testi ile serinin birim kök içerip içermediği test edilmiş ve sonuç aşağıda Tablo 11’de sunulmuştur.

**Tablo 11: İMKB 100 Serisi Getirileri (R) İçin AD Testi**

Düzey	Kesmesiz ve Trendsiz	Kesmeli	Kesme ve Trendli
İMKB	-23.45	-23.92	-23.92
%95 Güv. Düz. Kritik Değerler	-1.94	-2.89	-3.45
%99 Güv. Düz. Kritik Değerler	-2.58	-3.49	-4.05

Yukarıda tablo 11’de de görüldüğü gibi İMKB 100 haftalık getiri serisi durağan bir seridir. Seride birim kök varlığını ileri süren boş hipotez %1 anlamlılık düzeyinde red edilmektedir.

İkinci olarak, seride ARCH ve GARCH etkilerinin varlığı test edilmiştir. Bu amaçla, birinci aşamada ortalama modeli olarak en iyi ARMA modeli tespit edilmiştir. Model seçimi için değerlendirilen ölçütler sırasıyla, modelin parametrelerinin anlamlılığı, düzeltilmiş  $R^2$  değerinin yüksekliği, AIC ve SIC değerlerinin düşüklüğü, kalıntı kareler toplamı (SSE) değerinin küçüklüğü, Ençok olabilirlik oranının (OLB) yüksekliği, F istatistiğinin anlamlılığıdır. En iyi olarak

belirlenen üç model için sonuçlar aşağıda tablo12’de sunulmuştur.

**Tablo 12: İMKB 100 Serisi Getirileri İçin En İyi Ortalama Modeli**

	ARMA(3,3)	ARMA(2,3)	ARMA( 3,2 )
C	0.010 (0.003)	0.010 (0.002)	0.010 (0.003)
$\phi_1$	0.86 (0.073)	1.40 (0.014)	0.81 (0.05)
$\phi_2$	-1.16 (0.039)	-0.96 (0.013)	-1.07 (0.04)
$\phi_3$	0.33 (0.072)	-----	0.27 (0.03)
$\varphi_1$	-0.58 (0.074)	-1.15 (0.033)	-0.52 (0.036)
$\varphi_2$	1.02 (0.040)	0.61 (0.061)	0.92 (0.031)
$\varphi_3$	-0.04 (0.072)	0.25 (0.030)	-----
Adj.R <sup>2</sup>	0.09	0.09	0.09
AIC	-2.85	-2.86	-2.85
SIC	-2.82	-2.83	-2.82
SSE	3.41	3.44	3.43
OLB	1472.92	1472.75	1470.09
F	17.63	20.91	19.93
Prob.(F)	0.000	0.000	0.000

Yukarıdaki tablo 12’den hareketle ARMA(2,3) modeli en iyi model olarak belirlenmiştir.

İkinci aşamada, yukarıdaki ARMA(2,3) modelinin kalıntılarında hareketle ARCH LM testi ile koşullu değişen varyansın varlığı test edilmiş ve test sonuçları tablo 13’de verilmiştir. Buna göre ARMA (2,3) modelinin kalıntısına dayalı,

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \dots + \alpha_q \hat{\varepsilon}_{t-q}^2 + v_t$$

şeklindeki regresyon yardımıyla,  $LM = (T-p-q) R^2$  şeklinde hesaplanan test istatistiği ile,

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_q = \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \alpha_1 \neq \alpha_2 \neq \dots \neq \alpha_q \neq \beta_1 \neq \beta_2 \neq \dots \neq \beta_p \neq 0$$

hipotezi test edilmektedir.

**Tablo 13: İMKB 100 Serisi Getirileri İçin ARCH LM Testi**

<i>ARCH LM Testi</i>	<i>İMKB 100 Getirileri</i>
LM (1)	31.67 Prob.(0.0000)
LM (5)	60.46 Prob.(0.0000)
LM (10)	68.74 Prob.(0.0000)
LM (20)	76.51 Prob.(0.0000)

Tablo 13'den de görüldüğü şekilde  $H_0$  hipotezi sırasıyla 1, 5, 10 ve 20 gecikme uzunluğu için red edilmektedir. Bunun anlamı İMKB 100 getiri serisinde GARCH etkisinin olduğudur. Böylece GARCH modeli tahmin edilebilir. Yine en iyi model tahminini belirlemek için düzeltilmiş  $R^2$  değerinin yüksekliği, AIC ve SIC değerlerinin düşüklüğü, kalıntı kareler toplamı (SSE) değerinin küçüklüğü, En çok olabilirlik oranının (OLB) yüksekliği, F istatistiğinin anlamlılığı, ve RMSE, MAE, MAPE ve Theil'in eşitsizlik katsayısı (U) ölçütlerinin küçüklüğü ele alınacaktır. Ayrıca modeller değerlendirilirken GARCH modelinin

$\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0$  ve  $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$  şeklindeki kısıtları da göz önüne

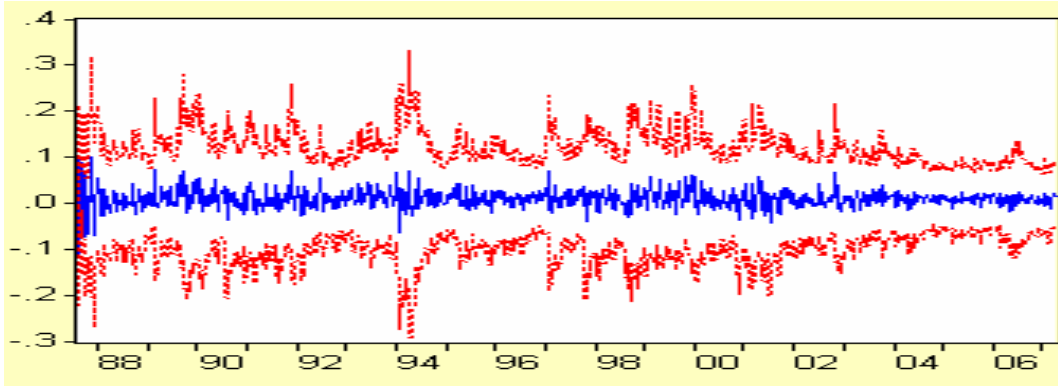
alınmalıdır. Aşağıda tablo14'de elde edilen en iyi dört modelin tahminleri sunulmuştur.

**Tablo 14: İMKB 100 Haftalık Getiri Serisi İçin Model Tahminleri**

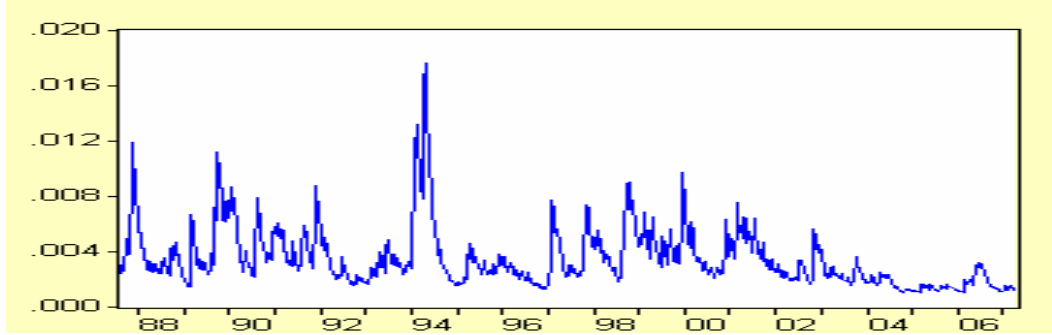
	GARCH(1,1)	GARCH(1,2)	GARCH(2,1)	GARCH(2,2)
C	0.009 (0.002)	0.010 (0.002)	0.009 (0.002)	0.009 (0.002)
$\phi_1$	1.28 (0.034)	-0.36 (0.35)	1.27 (0.034)	-0.37 (0.34)
$\phi_2$	-0.90 (0.032)	0.26 (0.28)	-0.91 (0.035)	0.30 (0.19)
$\varphi_1$	-1.018 (0.044)	0.63 (0.36)	-1.00 (0.045)	0.65 (0.35)
$\varphi_2$	0.61 (0.043)	0.09 (0.31)	0.62 (0.046)	-0.14 (0.24)
$\varphi_3$	0.22 (0.033)	0.007 (0.083)	0.22 (0.035)	-0.021 (0.064)
$\alpha_0$	0.0001 (0.000)	0.00005 (0.000)	0.000006 (0.000)	-0.0000008 (0.000)
$\alpha_1$	0.11 (0.015)	0.10 (0.019)	0.20 (0.039)	0.17 (0.004)
$\alpha_2$	-----	-----	-0.17	-0.17 (0.004)
$\beta_1$	0.86 (0.020)	0.48 (0.24)	0.96 (0.006)	1.60 (0.003)
$\beta_2$	-----	0.40 (0.23)	-----	-0.60 (0.001)
Adj.R <sup>2</sup>	0.081	0.080	0.088	0.08
AIC	-2.96	-2.95	-2.97	-2.98
SIC	-2.92	-2.90	-2.92	-2.92
SSE	3.45	3.47	3.44	3.47
OLB	1531.075	1526.923	1537.684	1539.863
F	12.34	10.01	10.94	8.93
Prob.(F)	0.000	0.000	0.000	0.000
RMSE	0.058	0.058	0.058	0.058
MAE	0.043	0.043	0.043	0.043
MAPE	216.23	233.68	213.58	231.85
U	0.70	0.72	0.71	0.72



Böylece durağanlık kontrolü yapılmış seri için GARCH kısıtlarını yerine getiren en iyi model GARCH (1,1) modeli olarak elde edilmiştir. Önraporlama değerlerinin örneklem içi önraporlama performans ölçütleri yukarıdaki tablo 14’de mevcuttur, önraporlama grafikleri ise aşağıda sırasıyla şekil 39 ve şekil 40’da verilmiştir.



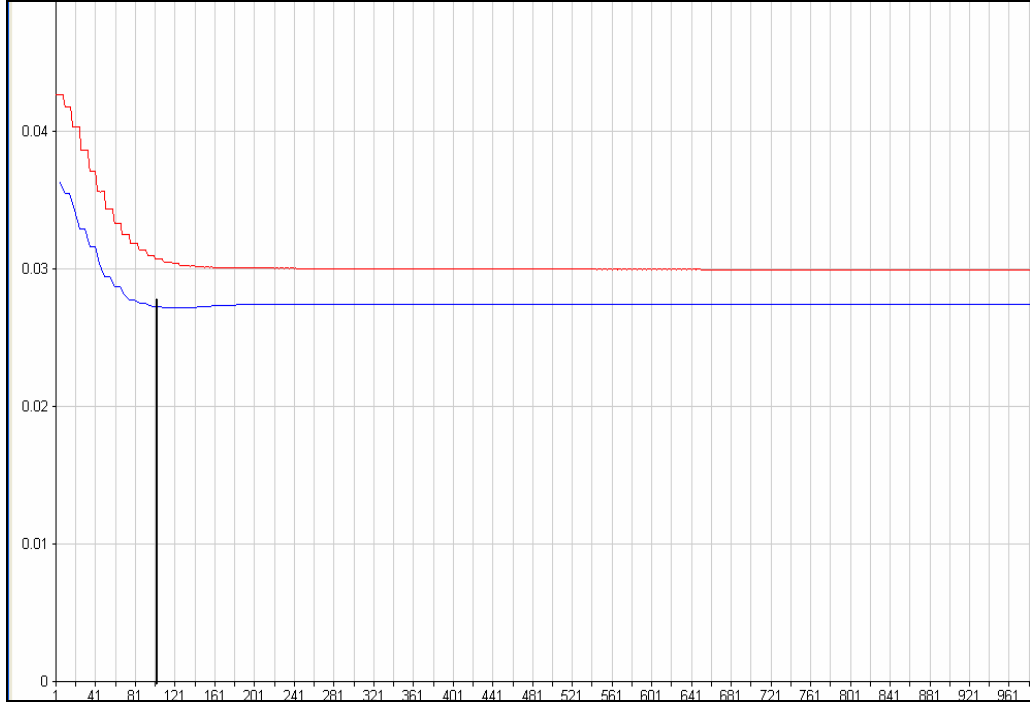
*Şekil 39: İMKB 100 Haftalık Getirileri İçin Ortalama Modeli Önraporlaması*



*Şekil 40: İMKB 100 Haftalık Getirileri İçin Varyans Modeli Önraporlaması*

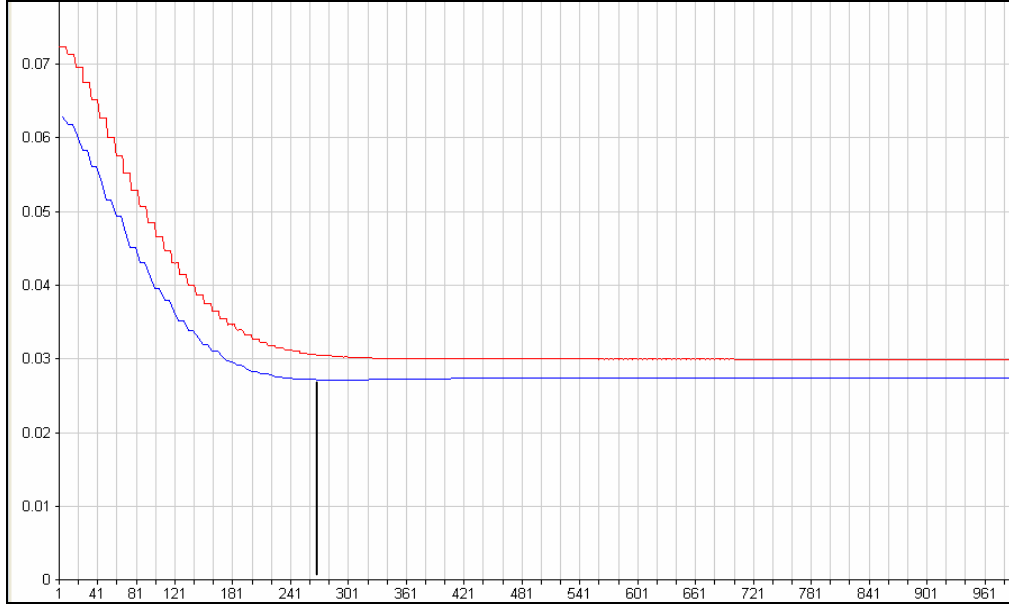
**İkinci adımda** İMKB 100 haftalık getiriler serisi yapay sinir ağı kullanılarak modellenecektir. Uygulamanın bu başlığı altında amaç, her türlü doğrusal dışı ilişkiyi yakalayabilecek modeller olarak tanıtılan yapay sinir ağı modellerinin varyanstaki doğrusal dışılığı yakalamada başarılı olup olmayacağını ortaya konulmasıdır. Bu nedenle yukarıda GARCH(1,1) modelinden elde edilen önraporlama performansı aşağıda yapay sinir ağı modelinden elde edilecek önraporlama performansı ile karşılaştırılacaktır.

Getirilerin orijinal olarak  $[-1,1]$  aralığında yer alması ve getiri serisinin durağan oluşu nedeniyle 1. ve 2. adımlar otomatik olarak atlanmıştır. Ağda kullanılan transfer fonksiyonu yine veri setinin orijinal değerlerinden dolayı hiperbolik tanjant fonksiyonu olarak belirlenmiştir. Ayrıca veri setinin yeterince geniş oluşu nedeniyle çapraz geçerlilik yönteminden yararlanılabilmektedir. Veri seti, literatür takip edilerek, %70 eğitim, %20 geçerlilik ve %10 test olmak üzere parçaya ayrılmıştır. Böylece eğitim için toplam 1030 gözlemin 721 tanesi, geçerlilik için 206 tanesi ve test için 103 tanesi kullanılmıştır. İkinci bölümde tanıtılan geri yayılım algoritması ve yine ikinci bölümde tanıtılan erken durdurma metodu kullanılarak ağın eğitimi, geçerlilik setindeki MSE yükselmeye başlayınca dek devam ettirilmiştir. Yükselmenin başladığı anda eğitim kesilerek ağın aşırı uyum problemine girmesi engellenmiştir. Bu işlem girdi katmanında farklı sayıda ve ara katmanında farklı sayıda nöron bulunan çeşitli ağ tiplerine uygulanmıştır. Böylece aşırı uyum probleminden kaçınılırken aynı zamanda en uygun ağ yapısı da, en düşük MSE değerine ulaşan ağ olarak belirlenmeye çalışılmıştır. Aşağıda şekil 41-44 en iyi olarak nitelendirilebilecek 4 ağa ilişkin geçerlilik ve eğitim setlerinin iterasyonlar boyunca MSE değerlerinin değişimini göstermektedir.



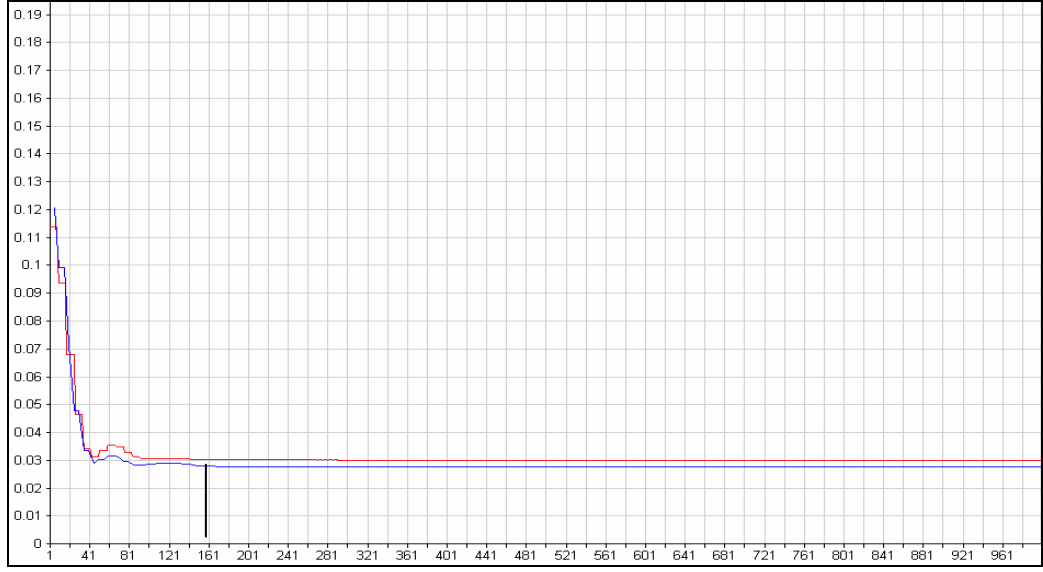
**Şekil 41: 2-3-1 FNN Cross Validation Eğitim Performansı**

Şekil 41’de kullanılan 2-3-1 ileri beslemeli ağ geçerlilik setinde 111. iterasyonda 0.02881 MSE değerine ulaşmaktadır. Bu noktadan sonra geçerlilik setindeki hata artmaya başlamıştır. Bu durumda kesilen ağ eğitiminde eğitim setindeki MSE değeri ise, 0.03181 dir



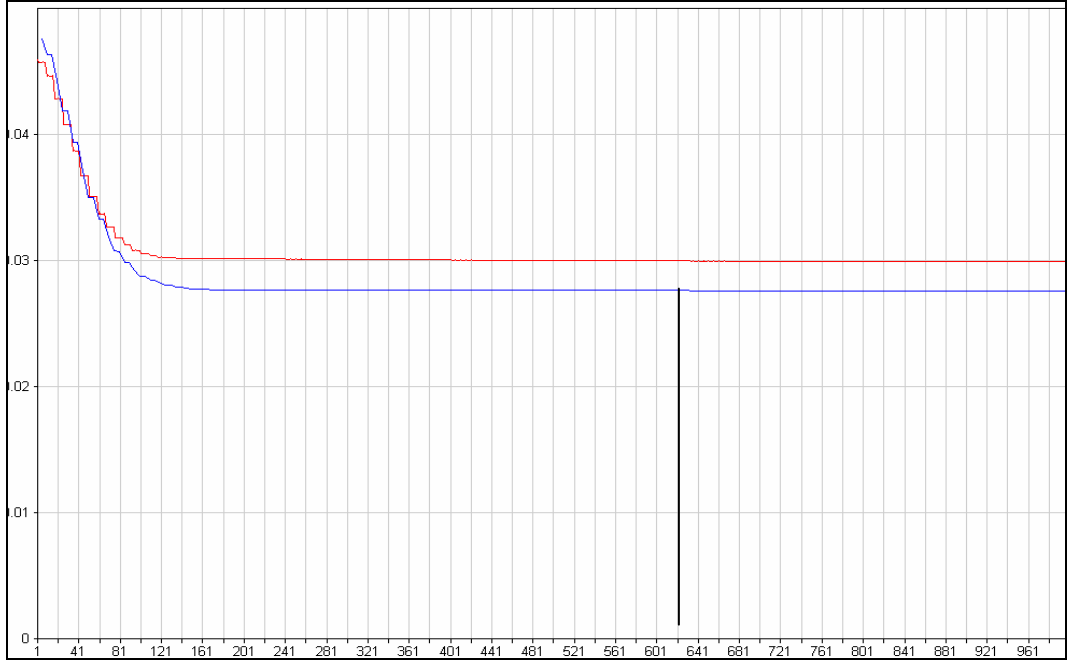
***Şekil 42: 3-4-1 FNN Cross Validation Eğitim Performansı***

Şekil 42’de 3-4-1 ileri beslemeli ağı ile geçerlilik setinde 273. iterasyonda 0.02896 MSE değerine ulaşılmış ve bu noktadan sonra geçerlilik setinde MSE artmaya başlamıştır. Bu noktada kesilen ağ eğitiminde, eğitim setindeki MSE değeri 0.03069’dur.



***Şekil 43: 2-5-1 FNN Cross Validation Eğitim Performansı***

Şekil 43’de 2-5-1 ileri beslemeli ağı ile geçerlilik setinde 168. iterasyonda 0.02886 MSE değerine ulaşılmış ve bu noktadan sonra geçerlilik setinde MSE artmaya başlamıştır. Bu noktada kesilen ağ eğitiminde, eğitim setindeki MSE değeri 0.03085’dir.

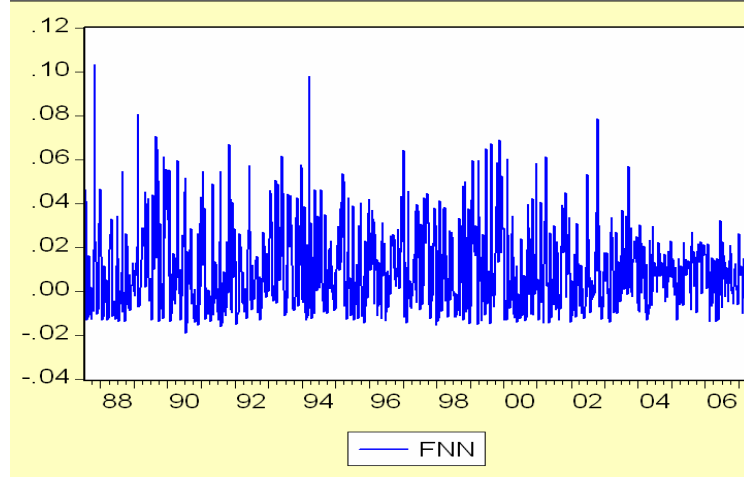


**Şekil 44: 3-5-1 FNN Cross Validation Eğitim Performansı**

Şekil 44'de 3-5-1 ileri beslemeli ağı ile geçerlilik setinde 636. iterasyonda 0.02897 MSE değerine ulaşılmış ve bu noktadan sonra geçerlilik setinde MSE artmaya başlamıştır. Bu noktada kesilen ağ eğitiminde, eğitim setindeki MSE değeri 0.03077'dir.

Yukarıdaki şekillerde sunulan tüm bu ağların eğitim ve geçerlilik setlerindeki hata yapılarını elde etmek için kullanılan geri yayılım algoritmasında momentum katsayısı ( $\alpha$ ) ve öğrenme oranı ( $\eta$ ) değerleri sırasıyla 0.5 ve 0.1 olarak kullanılmıştır. İkinci bölümde belirtildiği gibi Shadra ve Patil (1992) ve Haykin (1999)'un deneysel çalışmalarda en iyi sonucu verdiğini belirttikleri ikili değerler bunlardır.

Bu dört ağ içerisinde bu durumda en iyi eğitilmiş ağ 3-4-1 ağıdır. Aşağıda bu ağa ilişkin örneklem içi önraporlama şekil 45'de görülebilir.



**Şekil 45: 3-4-1 FNN Ağı ile İMKB-100 Haftalık Getirilerinin Önraporlaması**

Bu önraporlamaya ilişkin performans ölçütleri de aşağıda tablo 15’de sunulmuştur.

**Tablo 15: İMKB 100 Haftalık Getiri Serisi İçin 3-4-1 FNN Önraporlama Performans Ölçütleri**

MAE	MAPE	RMSE
0.067	235.18	0.075

Hesaplanan üç ölçüt içinde 3-4-1 ileri beslemeli ağı GARCH (1,1) modelinden daha kötü performans göstermiştir. Üstelik ağ eğitiminde gözlem sayısı yeterince büyük olduğunda kullanılabilen ayrıca, Nakamura (2005) ve Gonzalez (2000) tarafından aşırı uyum problemine takılmadan ağın hem eğitimini hemde katmanlardaki nöron sayısını belirlemede en iyi yaklaşım olarak belirtilen çapraz geçerlilik ve erken durdurma metotlarının beraber kullanılmış olmasına rağmen, varyanstaki doğrusal dışılığı da bir bilgi olarak kullanabilen GARCH modeli, volatilité modellemesinde yapay sinir ağı modellerinden daha iyi

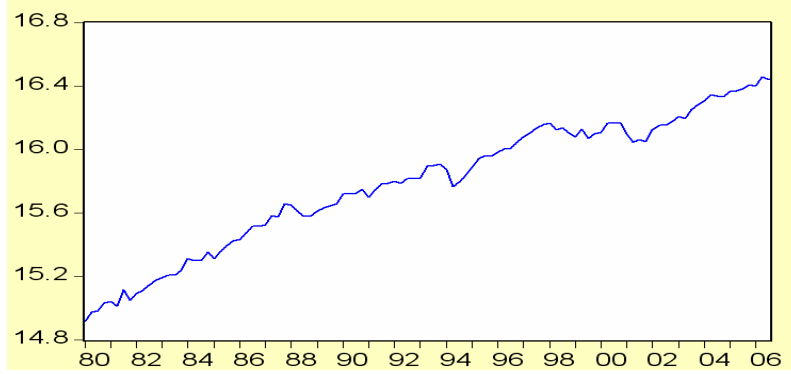
alıřmaktadır. Kanımızca yapay sinir ađı modelinin nraporlama performansı, yapay sinir ađı modellerinin tahmininde kullanılan algoritmaya varyanstaki bilginin de bir girdi olarak kullanılabileceđi řekilde bir takım dzenlemelerin yapılmasıyla arttırılabilir. Mevcut durumda GARCH modelinin bařarısı yapay sinir ađı modeliyle yakalanamamaktadır.



#### IV. YAPISAL KIRILMA GÖSTEREN BİR SERİ OLARAK SANAYİ ÜRETİM ENDEKSİNİN MODELLENMESİ

Sanayi üretim indeksi, madencilik sanayi, imalat sanayi, elektrik gaz ve su sektörlerindeki sanayi işyerlerinin ürettiği fiziksel çıktıdaki değişimi ölçmeyi amaçlamaktadır. Sanayi sektörünün zaman içindeki seyrini görmek, dönemler arasındaki değişimleri izlemek, ekonomideki konjonktürel gelişimi takip etmek için kullanılan Sanayi üretim indeksi, üretim yolu ile hesaplanan Sanayi Sektörü GSMH'sine yaklaşık bir tahmindir. İktisadi hayatın en önemli kısa dönem göstergelerinden biri olan sanayi üretim indeksi, ekonomi için dönüm noktalarının erken dönemde ortaya konulmasında oldukça kullanışlı bir endekstir. Gelişmekte olan ülkeler için sanayi üretiminin izlediği yol ve bu endeksteeki artış ve azalışların bir diğer önemli yanı ise, ekonominin büyümesi ve gelişmesi ile birlikte tarım sektörünün toplam üretim içerisindeki payının azalmasından ileri gelmektedir. Sanayi ve ona bağlı olarak hizmet sektörünün toplam üretim içerisindeki payının artması ile yaratılan yeni istihdam sadece mevcut işsizlere ve çalışma çağına girmekte olan yaş grubundakilere değil, ayrıca tarım sektöründen açığa çıkan yeni işsizlere de pay edilmektedir. Sanayi üretiminin artması ve buna bağlı olarak hizmet sektörünün büyümesi ve bu gelişiminde *öngörülebilmesi* bu nedenle de önemlidir. Teknik açıdan ise hem doğrusallığından hem de durağanlığından emin olduğumuz bir seri üzerinde doğrusal bir model ile yapay sinir ağı modelinin performansları karşılaştırılmış olacaktır.

Aşağıda şekil 46'da 1980 birinci çeyreği ile 2006 üçüncü çeyreği arasında Türkiye ekonomisi için sanayi üretim indeksi zaman yolu grafiği görülmektedir.



**Şekil 46 : Sanayi Üretim Endeksi (1980Q1-2006Q3)**

İlk olarak serinin birim kök içerip içermediğinin ortaya konulması için ADF testi uygulanmıştır. Aşağıdaki tablo16’da sanayi üretim endeksi (SÜE) için ADF test sonuçları sunulmuştur.

**Tablo 16: SÜE Serisi İçin ADF Testi**

<i>Düzyey – Birinci fark</i>	<i>Kesmesiz ve Trendsiz</i>	<i>Kesmeli</i>	<i>Kesme ve Trendli</i>
SÜE	4.23	-1.59	-2.82
$\Delta SÜE$	-9.90	-11.55	-11.59
%95 Güv. Düz. Kritik Değerler	-1.94	-2.89	-3.45
%99 Güv. Düz. Kritik Değerler	-2.58	-3.49	-4.05

ADF test sonuçlarına göre SÜE serisi düzey olarak ele alındığında bir birim kök içermektedir. Serinin birinci farkı ise durağandır.

İkinci olarak, serinin düzey olarak ele alındığında, bulunan birim kökün aslında ele alınan dönem içerisinde bir yapısal kırılmadan etkilenerek ADF test sonuçlarını birim kök bulma yönünde yanıltıcı kılmasından kaynaklanıp kaynaklanmadığını ortaya koymak için, üçüncü bölümde yapısal kırılma başlığı

altında sunulan Peron 1997 yöntemi ile birim kök testi tekrarlanmıştır. Düzeyde bir kırılma içeren birinci model ve hem kesme hem de eğimde bir kayma içeren ikinci model uygun olmadığından, eğimde kayma içeren üçüncü model baz alınarak elde edilen test sonuçları aşağıda tablo17’de sunulmuştur.

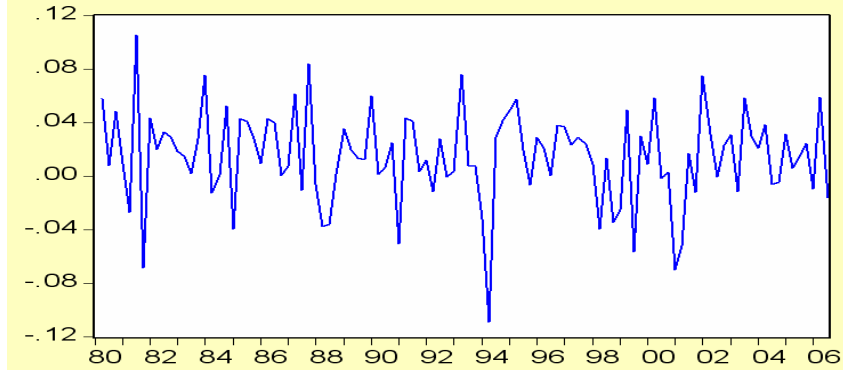
**Tablo 17: SÜE Serisi İçin Peron 1997 testi**

<i>Seri</i>	<i>T</i>	<i>T<sub>b</sub></i>	<i>Gecikme</i>	<i>t<sub>γ</sub></i>
SÜE	104	1994:2	9	-2.88

Parantez içerisindeki değerler t istatistikleri olup, 100 gözlem için 1% ve 5% kritik değerler sırasıyla -5.38 ve -4.67’dir

Öncelikle kırılma dönemi bilinmediği için,  $\lambda = \frac{T_B}{T}$  olmak üzere, ve toplam gözlem sayısının ilk ve son %15’lik kısmı dışlanarak elde kalan gözlemler üzerinden,  $T_B=1,2,3,\dots,T$  şeklinde ayrı ayrı regresyon tahminleri yapılmış ve  $\gamma < 0$  olduğundan, minimum  $|t_\gamma|$ ’yi veren dönem olan 1994 yılının ikinci çeyreği kırılma dönemi olarak tespit edilmiştir. İkinci olarak 9 gecikme uzunluğu AIC kriteri göz önüne alınarak uygun bulunmuştur. Sonuçta elde edilen model üzerinden hesaplanan  $t_\gamma$  istatistiği yukarıdaki tablo 17 ‘de sunulmuştur. Buna göre,  $\alpha=1$  boş hipotezini test eden t istatistiği, -2.88 olup %1 ve %5 anlamlılık düzeylerinde red edilememiştir.

Perron 1997 prosedürü ile de birim kökün kırılmadan kaynaklanmadığı açıklığa kavuştuktan sonra yapılması gereken serinin birinci farkını alarak önraporlama yapılacak modeli kurmaktır. Aşağıda şekil 47’de SÜE serisinin birinci farkı ( $\Delta SÜE$ ) serisinin zaman yolu grafiği görülmektedir.



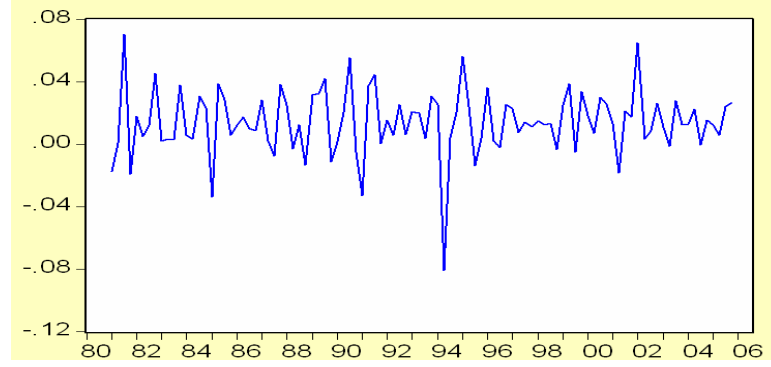
**Şekil 47 : Sanayi Üretim Endeksi Birinci Farkı (1980Q2-2006Q3)**

Önceki ADF testinde de açıklığa kavuşturulduğu gibi birinci fark serisi durağan bir seridir. Böylece bu seri için önraporlama modeli elde edilebilecektir. Aşağıda tablo18’de bu amaçla elde edilmiş en iyi üç modelin tahminleri ve modeller arasında en iyi modelin seçimi için kullanılacak kriterler sunulmuştur. Model seçimi için değerlendirilen ölçütler yine sırasıyla, modelin parametrelerinin anlamlılığı, düzeltilmiş  $R^2$  değerinin yüksekliği, AIC ve SIC değerlerinin düşüklüğü, kalıntı kareler toplamı (SSE) değerinin küçüklüğü, Ençok olabilirlik oranının (OLB) yüksekliği, F istatistiğinin anlamlılığı, ve RMSE, MAE, MAPE ve Theil’in eşitsizlik katsayısının (U) ölçütlerinin küçüklüğüdür. En iyi örneklem içi önraporlamayı veren modeli bulmak için denenen modellerde, peron 1997 yöntemiyle kırılma tespit edilen 1994Q2 dönemi için kukla değişken (b) kullanılmıştır.

**Tablo 18:  $\Delta S\ddot{U}E$  Serisi İçin Model Tahminleri**

	ARMA(1,3)	ARMA(3,3)	ARMA(3,4)
C	0.014 (0.002)	0.01 (0.00)	0.01 (0.00)
$\phi_1$	0.52 (0.022)	0.27 (0.011)	-0.49 (0.018)
$\phi_2$	-----	-0.24 (0.009)	-0.34 (0.016)
$\phi_3$	-----	0.72 (0.008)	-0.39 (0.013)
$\varphi_1$	-0.76 (0.021)	-0.47 (0.010)	0.42 (0.017)
$\varphi_2$	0.45 (0.011)	0.35 (0.11)	0.49 (0.015)
$\varphi_3$	-0.40 (0.009)	-0.87 (0.05)	0.43 (0.017)
$\varphi_4$	-----	-----	-0.47 (0.13)
b	-0.12 (0.032)	-0.12 (0,030)	-0.64 (0.016)
Adj.R <sup>2</sup>	0.07	0.12	0.21
AIC	-3.88	-3,95	-4.04
SIC	-3.75	-3.77	-3.84
SSE	0.12	0.10	0.09
OLB	208.72	210.36	216.33
F	1.87	3.35	4.86
Prob.(F)	(0.12)	(0.00)	(0.00)
RMSE	0.033	0.031	0.029
MAE	0.026	0.024	0.022
MAPE	1662.40	618.36	1998.86
U	0.50	0.49	0.44

Tablo 18'e göre ARMA(3,4) modeli örneklem içi önraporlamalar için en iyi önraporlama performansını veren modeldir. Bu modele ilişkin örneklem içi önraporlama grafiği aşağıda şekil 48'de sunulmuştur.



**Şekil 48: ARMA (3,4) Örneklem İçi Önraporlama Değerleri**

Bu önraporlama değerlerine ilişkin önraporlama performans ölçütleri yukarıda tablo 18’de mevcuttur.

**ikinci olarak**, durağan olduğu bilinen birinci fark serisi yapay sinir ağı ile modellenmiştir. Uygulamanın bu başlığı altında amaç doğrusal olduğu bilinen ve kırılmaya sahip durağan bir serinin modellenmesinde yapay sinir ağı modellerinin uygunluğunun araştırılmasıdır. Zhang (2000), Zhang ve Patuwo (1998) doğrusal serilerde Franses (2000) doğrusal ve kırılmalı serilerde yapay sinir ağı modellerinin başarılı şekillerde kullanılabileceğini belirtirlerken, diğer taraftan Desai ve Bharati (1998) aşırı uyum probleminin bu tip serilerle çalışıldığında çok daha sık rastlanan bir problem olarak ortaya çıktığını ileri sürmektedir. Gorl (1994) ise doğrusal olmayan bir modelleme aracı olarak yapay sinir ağlarının doğrusal serilerde kullanılmasının sakıncalı olduğunu belirtmektedir. Bu nedenle Türkiye ekonomisine ait sanayi üretim endeksi birinci fark serisi yapay sinir ağı ile modellenip önraporlama performansı yukarıda ARMA(3,4) modelinden elde edilen önraporlama performansı ile karşılaştırılmıştır.

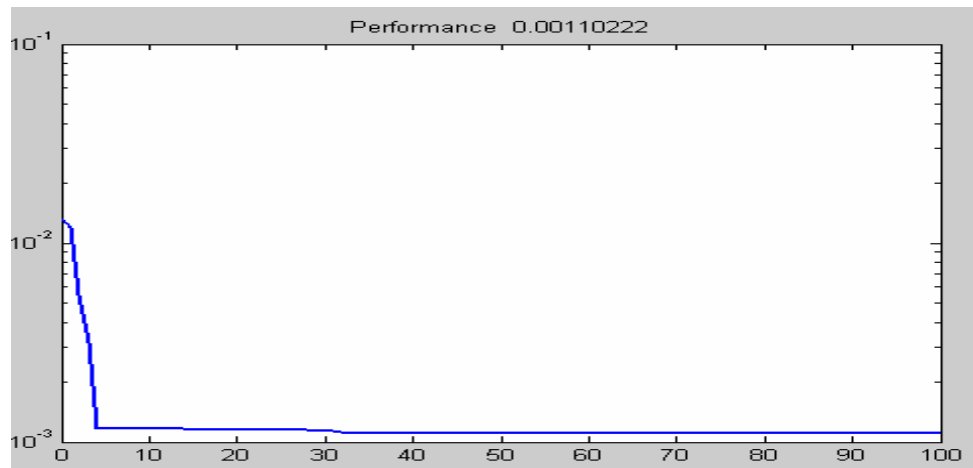
Serinin durağan olduğu ve orijinal değerlerinin  $[-1,1]$  sınırları içerisinde olduğu göz önüne alınarak, adımsal prosedürün ilk iki adımı atlanmış ve ayrıca ağda kullanılacak transfer fonksiyonunun hiperbolik tanjant fonksiyonu olmasına karar verilmiştir. Öte yandan 3.,4. ve 5. adımlar yine eşanlı olarak ele alınmıştır.

Tahmin algoritması olarak Newton Rapsion metodu kullanılmıştır. Eğitim MSE'deki değişim 0.01'e düşüne dek devam ettirilmiştir. Çapraz geçerlilik yöntemi ele alınan veri setinin yeterince büyük olması nedeniyle tercih edilmemiştir. Girdi ve ara katmanlarda nöron sayısına karar vermek için aşağıdaki tablo 19'da sunulan SIC ve MSE ölçütleri kullanılmıştır.

**Tablo 19:  $\Delta SUE$  Serisi İçin Ağ Seçimi**

AĞ ADI	SIC	MSE	AĞ ADI	SIC	MSE	AĞ ADI	SIC	MSE	AĞ ADI	SIC	MSE
2-3-1	-640	0.0012	3-3-1	-624	---	4-3-1	-618	---	5-3-1	-610	---
2-4-1	-607	---	3-4-1	-629	---	4-4-1	-631	---	5-4-1	-621	---
2-5-1	-624	---	3-5-1	-634	0.0014	4-5-1	-587	---	5-5-1	-635	0.0014
2-6-1	-593	---	3-6-1	-612	---	4-6-1	-640	0.0011	5-6-1	-593	---
2-7-1	-635	---	3-7-1	-618	---	4-7-1	-599	---	5-7-1	-590	---
2-8-1	-602	---	3-8-1	-577	---	4-8-1	-612	---	5-8-1	-570	---
2-9-1	-613	---	3-9-1	-591	---	4-9-1	-574	---	5-9-1	-562	---
2-10-1	-585	---	3-10-1	-570	---	4-10-1	-560	---	5-10-1	-540	---

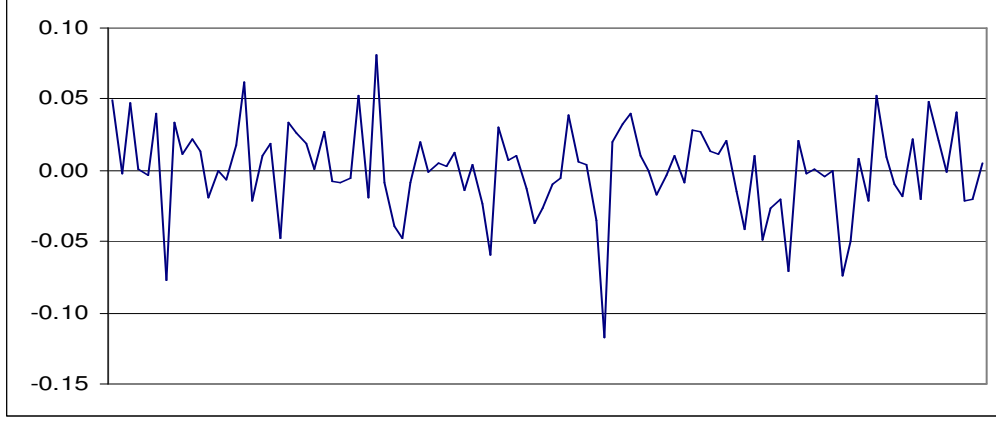
Tablo 19'da da görüldüğü gibi 4-6-1 ileri beslemeli ağ en küçük SIC ve en düşük MSE değerlerine ulaşan ağdır. Ağın eğitiminin iterasyonlar boyunca MSE değişimi (100. iterasyona kadar) aşağıda şekil 20'de görülebilir.



**Şekil 49 : Ağın Eğitiminde İterasyonlar Boyunca MSE değişimi**

Aşağıda bu ağın kullanılmasıyla elde edilen örneklem içi önraporlama

şekil 50’de sunulmuştur.



**Şekil 50 : 4-6-1 FNN Örneklem İçi Önraporlama Değerleri**

Yukarıda grafiği görülen 4-6-1 ileri beslemeli ağının örneklem içi önraporlama performansı ölçütleri ise aşağıda tablo 20’de görülebilir.

**Tablo 20:  $\Delta SUE$  İçin 4-6-1 FNN Önraporlama Performans Ölçütleri**

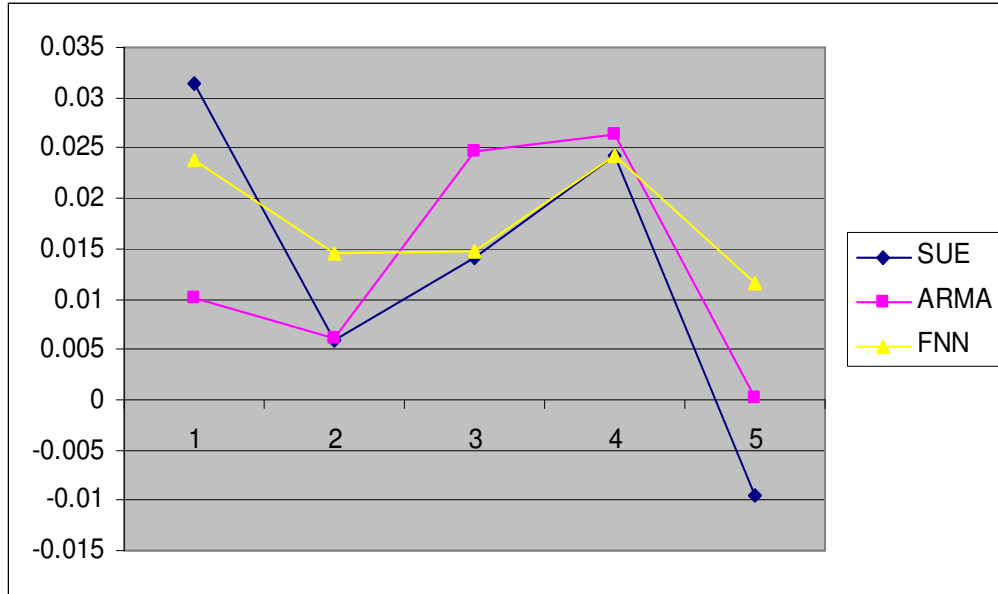
MAE	MAPE	RMSE
0.01433	1237.545	0.155831

Bu elde edilen önraporlama performans ölçütlerinin tablo 13 ‘deki ARMA (3,4) modeli için verilen ölçütlerle karşılaştırılması yapıldığında, MAE ve MAPE ölçütüne göre 4-6-1 Ağının, RMSE ölçütüne göre ise ARMA(3,4) modelinin daha iyi olduğu söylenebilir. Bu nedenle yine ex-post dönem önraporlaması yapılması önem arz etmektedir. Aşağıda tablo 21’de  $\Delta SUE$  için 2005 birinci çeyreğinden, 2006 birinci çeyreği dahil olmak üzere 5 gerçekleşen ve modellerde kullanılmamış gözlem ile ARMA(3,4) ve 4-6-1 FNN modellerinin bu dönem için ex-post önraporlamaları sunulmuştur. Bu değerler ayrıca şekil 51’de görülebilir.



**Tablo 21:  $\Delta SUE$  İçin Gerçekleşen Değerler ve Expost Önraporlama Performansları**

Model	2006-Q1	2006-Q2	2006-Q3	2005Q4	2006-Q1	TMS
4-6-1 FNN	0.023921	0.014613	0.01467	0.024248	0.011549	0.037947
ARMA(3,4)	0.01001	0.006184	0.024624	0.026341	0.000238	0.044165
GERÇEK	0.031464	0.00584	0.014165	0.024231	-0.00956	-----



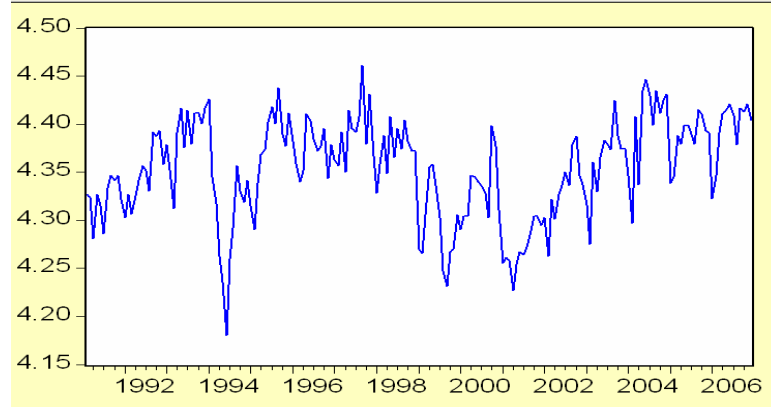
**Şekil 51:  $\Delta SUE$  Serisi İçin Gerçekleşen Değerler ve Expost Önraporlamalar**

Yukarıda tablo 21’de 5 dönem için hesaplanan toplam mutlak sapma değeri 4-6-1 FNN modeli için 0.037947 iken ARMA (3,4) modeli için 0.044165 olarak elde edilmiştir. Daha önce de tanıtılan Diebold Mariano (1995) S istatistiği, 4-6-1 FNN modeli temel alınarak hesaplandığında  $S = 2.236$  değeri, elde edilmektedir. Testin boş hipotezi iki modelin önraporlamaları arasında fark yoktur şeklinde olduğu bir kez daha hatırlatılacak olursa, 4-6-1 ağının expost önraporlama performansının istatistiksel olarak %5 anlamlılık düzeyinde (standart normal dağılım kritik değeri %5 için 1.96 olduğundan) daha iyi olduğu

söylenbilir. Bu elde edilen sonuçlar, doğrusal serilerde de yapay sinir ağlarının önraporlama amacıyla kullanılabilmesine dair bir bilgi olarak ileri sürülebilir niteliktedir. Kullanılan yapay sinir ağı modelinin sadece örneklem içi değil aynı zamanda örneklem dışı dönem önraporlamalarında da doğrusal modelden daha iyi sonuç veriyor olması ve istatistiksel olarak da iki modelin önraporlamalarının gerçekleşen değerlerden farkının anlamlı oluşu FNN modelinin kırılma gösteren seride kullanılabilmesi yönünde bir sonuçtur.

## V. KAPASİTE KULLANIM ORANLARI SERİSİNDE DOĞRUSAL DIŞILIĞIN MODELLENMESİ

Kapasite kullanım oranı bir fiil üretimde bulunan fabrika ve ekipmanların cari üretim oranıdır. Kapasite kullanım oranındaki düşüşler ekonominin durgunlaşmaya gittiğinin işaretini verirken artışlar ise ekonominin genişlemeye yöneldiği anlamına gelir. Kapasite Kullanım Oranının düzenli takibi, sanayideki üretim düzeyinin, talep ve yatırım bazında izlenmesini sağlar. Bir ekonomideki mevcut kapasitenin kullanımının talep düzeyine göre belirlendiğini kabul edilirse, yatırımların faiz oranlarının düşük düzeylerde olacağı ve/veya talebin yüksek düzeyde devam edeceği beklentileri ile artacağı açıktır. Üretim kapasitesinin bu yeni yatırımlar sayesinde artırılması ile kapasite kullanım oranı normal düzeylere düşecektir. Kapasite kullanımının yüksek olduğu dönemlerde, genellikle sanayi üretimi de artmaktadır. Ekonomik büyümenin hızlı olduğu dönemlerde, kapasite kullanım oranının yüksek olması enflasyonist baskıların ağırlaşmasını da beraberinde getirir. Kapasite kullanım oranının, devamlı ve uzun bir süre yüksek düzeylerde kalması bu durumda o kadar da olumlu bir işaret olarak algılanmamalıdır. Çünkü bu tip durumlar aslında sanayi yenileme yatırımlarının gecikmiş olduğunu işaret etmektedir. Aşağıda şekil 52’de Türkiye ekonomisi için 1991 Şubat – 2006 Aralık dönemi için logaritmik kapasite kullanım oranları serisinin (KKO) zaman yolu grafiği sunulmuştur.



**Şekil 52: Kapasite Kullanım Oranları (02.1991-12.2006)**

Bu seri, ARMA, kendine özgü doğrusal dışılık formu bulunan TAR ve STAR modelleri ve genel bir doğrusal dışı yaklaşıma sahip geri dönüşümlü bir ağ (RNN) kullanılarak modellenecek, örneklem içi ve örneklem dışı ex-post dönem için önraporlamalar karşılaştırılacaktır. Seri için öncelikle ADF birim kök testi uygulanmıştır. Sonuçlar aşağıda tablo 22’de sunulmuştur.

**Tablo 22: KKO Serisi İçin ADF testi**

Düzy	Kesmesiz ve Trendsiz	Kesmeli	Kesme ve Trendli
KKO	0.041	-5.08	-5.17
%95 Güv. Düz. Kritik Değerler	-1.94	-2.89	-3.45
%99 Güv. Düz. Kritik Değerler	-2.58	-3.49	-4.05

Kesmeli, kesme ve trend içeren modeller için birim kök hipotezi %1 anlamlılık düzeyinde red edilmektedir. KKO serisi birim kök içermemektedir.

İlk olarak ARMA modelleme yaklaşımıyla KKO seri modellenmiştir. Bilgi kriterleri kullanılarak en uygun modelin seçimi aşağıda tablo 23’de

sunulmuştur.

**Tablo 23: Uygun ARMA(p,q) Modelinin Seçimi**

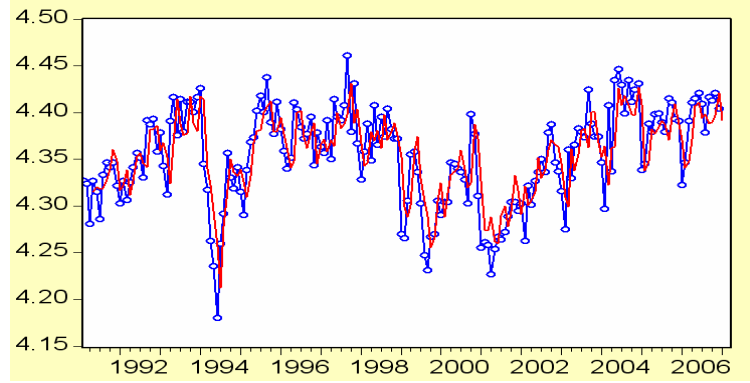
<i>Model</i>	<i>AIC</i>	<i>SIC</i>
AR(1)	-3.93	-3.89
AR(2)	3.93	-3.88
ARMA(2,2)	-3.94	-3.86
ARMA(3,2)	-4.00	-3.90

Tablo 23'e göre en uygun model olarak en düşük AIC ve SIC değerlerine sahip olan ARMA(3,2) modeli belirlenmiştir. Aşağıda tablo24'de bu modelin tahmini ve örneklem içi önraporlama performans ölçütleri sunulmuştur.

**Tablo 24: ARMA(3,2) Modeli**

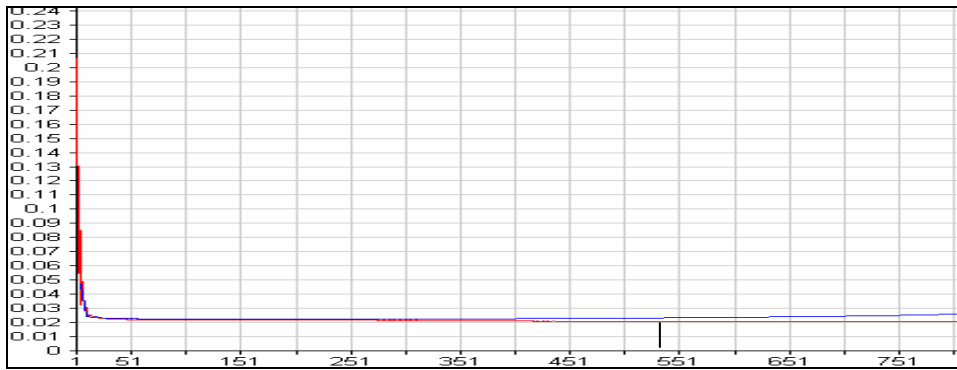
	ARMA(3,2)
C	4.36 (0.000)
$\phi_1$	0.16 (0.002)
$\phi_2$	-0.40 (0.000)
$\phi_3$	0.69 (0.000)
$\varphi_1$	0.54 (0.000)
$\varphi_2$	0.98 (0.000)
Adj.R	0.61
SSE	0.19
OLB	388
RMSE	0.032
MAE	0.035
MAPE	0.058

ARMA (3,2) modelinin örneklem içi önraporlamalarının grafiği ise aşağıda şekil 53'de görülebilir.



**Şekil 53: ARMA(3,2) Modeli Örneklem İçi Önraporlaması**

**İkinci olarak,** KKO serisi RNN ile modellenmiştir. İlk olarak veri seti, lojistik fonksiyon transfer fonksiyonu olarak kullanılacağı için [0 1] aralığına çekilmiştir. İkinci olarak veri seti çapraz geçerlilik yönteminden yararlanmak için %70 eğitim, %20 geçerlilik ve %10 test seti olmak üzere üç parçaya ayrılmıştır. Böylece eğitim seti için 134, geçerlilik seti için 38 ve test seti içinde 19 veri kullanılmıştır. Levenberg – Marquardt algoritması tahmin yöntemi olarak seçilmiş ve yine erken durdurma metodu kullanılarak geçerlilik setindeki hata minimum düzeye indiği durumda eğitim sona erdirilmiştir. Böylece bu yöntemle hem aşırı uyum problemine yakalanılmamış hem de ağ spesifikasyonu belirlenmiştir. Aşağıda şekil 54’de bu eğitim süreci, seçilen ağ yapısı 2-3-1 RNN modeli için gösterilmektedir.

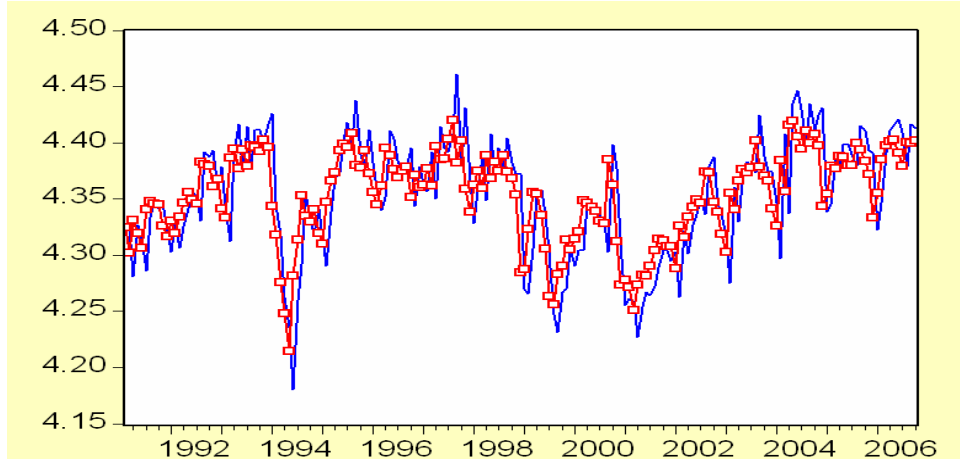


**Şekil 54: 2-3-1 RNN İçin Çapraz Geçerlilik Yönteminde Erken Durdurma**

539. iterasyonda geçerlilik setindeki MSE 0.01989 olarak minimum

değerine ulaşmıştır. Bu iterasyondaki eğitim setindeki MSE ise 0.01938'dir. bu noktada eğitim kesilerek ağın aşırı uyum problemine girmesi engellenmiştir.

Aşağıda, elde edilen ağırlık değerleriyle oluşan örneklem içi önraporlama şekil 55'de görülebilir.



*Şekil 55: 2-3-1 RNN ile KKO serisi Örneklem İçi Önraporlama*

Bu modele ilişkin önraporlama performans ölçütleri ise aşağıda tablo 25'de görülebilir.

*Tablo 25: KKO İçin 2-3-1 RNN Önraporlama Performans Ölçütleri*

MAE	MAPE	MSE	RMSE
0.0244	0.0005	0.0019	0.024

**Üçüncü olarak**, KKO serisi TAR modeli ile modellenmiştir. Üçüncü bölümde de belirtildiği gibi TAR modelinin kurulumunda yanıtlanması gereken sorular şu şekildedir. İlk olarak modelin, TAR modelimi yoksa M-TAR modeli mi olması gerektiğine karar verilmelidir. İkinci olarak eşik değer ne olacağına ve kaç gecikme ile belirleneceğine karar verilmelidir. Üçüncü olarak eşik değer altında ve üstünde çalışacak modellerin kaç gecikme kullanacağına karar

verilmesi gerekmektedir. Bu amaçla yapılan bir dizi model tahmini ve ulaşılan sonuçlar aşağıda tablo 26’de sunulmuştur.

**Tablo 26:Eşik Değer Modeli İçin En Uygun Modelin Belirlenmesi**

MODEL	Eşik Değer İçin Gecikme	Eşik Değer	Düşük Rejim İçin Gecikme	Yüksek Rejim İçin Gecikme	AIC
TAR	1	4.383588	2	2	-749.4664
TAR	0	4.409763	2	1	-747.3268
M-TAR	0	4.409763	2	2	-747.3195
TAR	0	4.409163	2	1	-746.5370
TAR	1	4.383588	1	2	-745.3388
TAR	1	4.383588	2	1	-745.3267
M-TAR	0	4.409763	2	2	-744.5867
TAR	0	4.335327	1	1	-744.5722
TAR	0	4.357980	1	1	-744.1777
M-TAR	1	4.383588	2	2	-74433397

Tablo 26’den da anlaşılacağı gibi eşik değer belirlemede bir gecikmeli değişkeni kullanan yada diğer bir ifade ile  $KKO_{t-1}$  ‘e göre eşik değer belirlendiği ve 4.383588 değerini eşik değer olarak kullanan, düşük rejimde 2 ve yüksek rejimde yine 2 gecikmeli otoregresif yapıya sahip TAR modeli en düşük AIC değeri olan -749.4664 değeri ile en uygun model olarak belirlenmiştir.

Model,

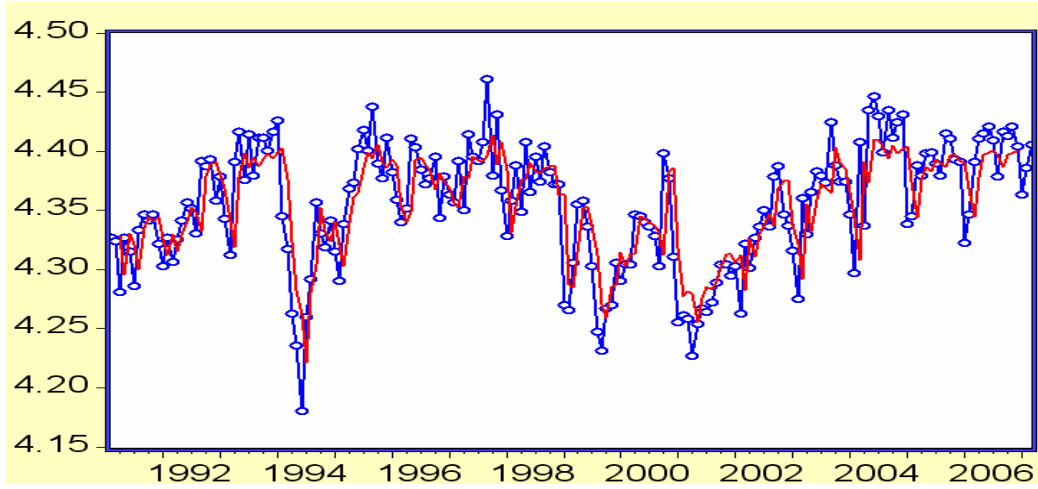
$$KKO_t = I_t(4.11 + 0.75KKO_{t-1} - 0.003KKO_{t-2}) + (1 - I_t)(3.85 + 0.32KKO_{t-1} + 0.03KKO_{t-2}) + \varepsilon_t$$

$$I_t = \begin{cases} 1 & KKO_{t-1} > 4.383588 \text{ ise} \\ 0 & KKO_{t-1} \leq 4.383588 \text{ ise} \end{cases}$$

şeklinindedir. Burada dikkat çekici bir durum yüksek rejim için geçerli modelin ilk otoregresif gecikmesinin 0.75 gibi yüksek bir değer olmasıdır. Bu durumun



anlamı, genelde Türkiye ekonomisinde yenileme yatırımlarının geciktiği ve özellikle ekonominin genişleme dönemlerinde enflasyonist baskıların oluştuğudur. Çünkü modelden, yüksek rejimde pozitif şokların kalıcılığının yüksek olduğunu, düşük rejimde ise ortalamaya dönme isteğinin kuvvetli olduğu anlaşılmaktadır. Modele ilişkin örneklem içi önraporlama grafiğide aşağıda şekil 56’da sunulmuştur.



*Şekil 56: TAR ile KKO Serisi Örneklem İçi Önraporlama*

Bu önraporlamaya ilişkin performans ölçütleri ise tablo 27’de sunulmuştur.

*Tablo 27: KKO İçin TAR Modeli Önraporlama Performans Ölçütleri*

MAE	MAPE	MSE	RMSE
0.0247	0.0005	0.0020	0.025

**Dördüncü olarak,** KKO serisi STAR modeli ile modellenmiştir. Bu amaçla üçüncü bölümde de belirtildiği gibi öncelikle STAR lojistik yada üstel fonksiyonlardan hangisini kullanacağına karar vermek gerekmektedir. Diğer bir ifadeyle modelin LSTAR yada ESTAR modellerinden hangisinin olması

gerektiğine karar verilmelidir. Bu nedenle ilk olarak daha önce tahmin edilmiş ARMA(3,2) modelinin kalıntıları üzerinden aşağıdaki regresyon hesaplanmıştır.

$$\begin{aligned}
 e_t = & a_0 + a_1 y_{t-1} + \dots + a_p y_{t-p} \\
 & + b_1 y_{t-1} y_{t-d} + \dots + b_p y_{t-1} y_{t-p} \\
 & + c_1 y_{t-1} y_{t-d}^2 + \dots + c_p y_{t-1} y_{t-d}^2 \\
 & + d_1 y_{t-1} y_{t-d}^3 + \dots + d_p y_{t-1} y_{t-d}^3 + \varepsilon_t
 \end{aligned}$$

Burada d gecikme uzunluğu 3 gecikmeye kadar ele alınmış ve elde edilen regresyonun T.R<sup>2</sup> istatistiği 20.16 olarak elde edilmiştir. Bu değer %5 anlamlılık düzeyinde 9 serbestlik dereceli kritik  $\chi^2$  tablo değeri olan 16.92'dan büyük olduğu için doğrusallık boş hipotezi red edilmiştir. İkinci adımda yine yukarıdaki regresyona dayanarak ve yine üçüncü bölümde belirtildiği gibi,

$$H_0 : d_1 = d_2 = \dots = d_p = 0$$

boş hipotezini test edilir. Eğer bu hipotez red edilirse, geçerli modelin LSTAR modeli olduğuna karar verilir. Eğer bu hipotez red edilemiyorsa geçerli model ESTAR modeli olacaktır. Bu amaçla hesaplanan F testi (3,176)s.d. ile 4.655 olarak hesaplanmıştır ve test %5 istatistiksel anlamlılık düzeyinde anlamlı bulunmuştur. Böylece hem eşik değeri etkisi anlamlı bulunmuş hemde LSTAR modelinin kullanılması gereken model olduğu sonucuna varılmıştır.

Aşağıdaki tablo 28'de LSTAR modelinin spesifikasyonuna karar vermek için denenen en iyi beş modelin yapısı sunulmaktadır.

**Tablo 28: LSTAR Model Spesifikasyonunun Belirlenmesi**

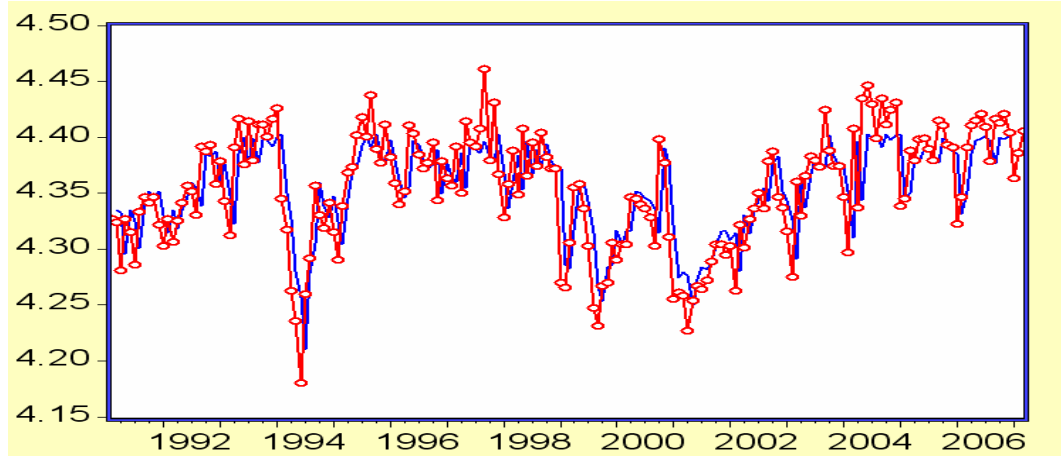
MODEL	Eşik Değer için Gecikme	Düşük Rejim için Gecikme	Yüksek Rejim için Gecikme	AIC
LSTAR	1	1	1	360.72
LSTAR	1	1	2	362.76
LSTAR	1	2	1	362.94
LSTAR	1	2	2	364.64
LSTAR	0	1	2	369.13

Tablo 28'ye göre eşik değeri için  $KKO_{t-1}$  değişkenini kullanan düşük ve yüksek rejimde bir otoregresif gecikme yapısıyla çalıştırılan ve 360.72 AIC değerine sahip model en uygun model olarak belirlenmiştir.

Tahmin edilen model,

$$KKO_t = (0.62 + 0.85KKO_{t-1} + ((3.93 - 3.69KKO_{t-1}) / (1 + \exp(6.06(KKO_{t-1} - 4.39))))$$

şeklinindedir. Bu modele ilişkin örneklem içi önraporlama aşağıda Şekil 57'de sunulmuştur.



**Şekil 57: LSTAR ile KKO Serisi Örneklem İçi Önraporlama**

Modele ilişkin önraporlama performans ölçütleri ise aşağıdaki Tablo 29'daki gibidir.

**Tablo 29: KKO İçin LSTAR Önraporlama Performans Ölçütleri**

MAE	MAPE	MSE	RMSE
0.0244	0.0006	0.0010	0.024

Buraya kadar kullanılan modellerin örneklem içi performans ölçütleri değerlendirildiğinde en iyi uyum gösteren modelin belirlenmesi güçtür. MSE ölçütüne göre LSTAR modeli, MAPE ölçütüne göre RNN ve TAR modelleri, MAE ölçütüne göre RNN ve TAR modelleri ve RMSE ölçütüne göre LSTAR ve

RNN modelleri örneklem içi önraporlamalarda daha iyi performans göstermişlerdir.

Bu nedenle bir kez daha ex-post dönem önraporlamaya başvurulmuştur. Buraya kadar ele alınan modellerin tümünün 2007 Ocak-Nisan dönemi için ex-post dönem önraporlamaları aşağıda tablo 30'da gerçekleşen değerler ile karşılaştırılmalı olarak sunulmuştur.

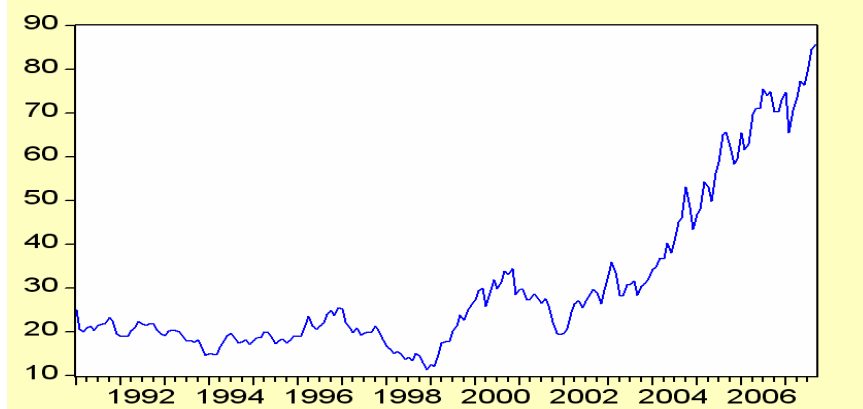
**Tablo 30: KKO İçin Ex-post Önraporlamalar**

	<b>2007/1</b>	<b>2007/2</b>	<b>2007/3</b>	<b>2007/4</b>	<b>TMS</b>
<b>Gerçekleşen</b>	4.3630	4.3857	4.4054	4.4042	----
<b>ARMA (3,2)</b>	4.3915	4.3514	4.3821	4.4258	0.1077
<b>TAR</b>	4.4040	4.3956	4.3924	4.3908	0.0776
<b>LSTAR</b>	4.4099	4.4011	4.3922	4.3855	0.0635
<b>RNN</b>	4.3910	4.4064	4.3921	4.4057	0.0635

Tablodaki son sütun ex-post önraporlama yapılan 4 dönem için, toplam mutlak sapmayı göstermektedir. İlginç bir şekilde LSTAR ve RNN modellerinin toplam mutlak sapmaları da ex-post dönem için eşit çıkmıştır. TAR modelinin bu iki modelden farkı ancak (dört dönem toplamı için) 0.0141 olarak elde edilmiştir. sonuçta RNN ve LSTAR başta olmak üzere üç modelde başarılıdır denilebilir.

## VI. AĞIRLIKLANDIRILMIŞ PENCERE YÖNTEMİNİN HAM PETROL FİYATLARI ÜZERİNE UYGULAMASI

Ham petrol fiyatlarındaki yükseliş ve düşüşlerin makroekonomi için yüksek derecede öneme sahip olduğu tüm iktisatçılar tarafından kabul edilmektedir. Petrol fiyatlarındaki yükselişler enflasyon, döviz kuru ve ödemeler dengesi üzerinde baskılara yol açarak GSMH üzerinde negatif etkiye bulunmaktadır. Petrol fiyatlarındaki düşüş ise, artışların yarattığı negatif etkiye oranla daha sınırlı pozitif gelişmeyi beraberinde getirir. Bu nedenler, petrol fiyatlarının önraporlamaları özellikle petrol fiyatlarındaki artış dönemlerinde gelişmekte olan petrol ithalatçısı ülkeler için önem arz etmektedir. Aşağıda şekil 58’de 1991 Ocak – 2007 Mayıs dönemi için ham petrol fiyatları serisinin zaman yolu grafiği sunulmuştur.



**Şekil 58: Ham Petrol Fiyatları 1991Ocak – 2007Mayıs**

Şekilde 2001 yılı Kasım ayı sonrası ortaya çıkan trendli dönem göze çarpmaktadır. Ham petrol varil fiyatlarının önraporlamalarının elde edilmesinde bu dönemin dinamiklerinin etkili olacağı açıktır. Bu dönem zarfında ham petrol varil fiyatları 19.67 dolardan 77.23 dolara çıkmıştır. Aşağıda tablo 31 logaritmik dönüşümü yapılmış ham petrol varil fiyatları serisi (COP) için ADF birim kök testi sonuçlarını göstermektedir.

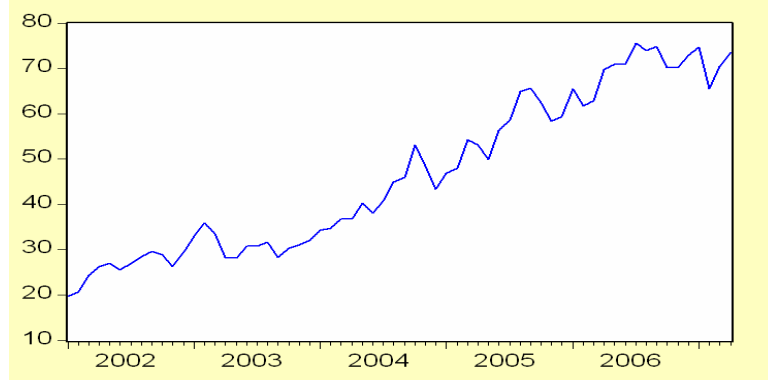
**Tablo 31: COP için ADF testi**

<i>Düzey / Birinci Fark</i>	<i>Kesmesiz ve Trendsiz</i>	<i>Kesmeli</i>	<i>Kesme ve Trendli</i>
COP	1.27	0.37	-1.94
$\Delta$ COP	-12.76	-12.85	-12.95
%95 Güv. Düz. Kritik Değerler	-1.94	-2.89	-3.45
%99 Güv. Düz. Kritik Değerler	-2.58	-3.49	-4.05

Tablo 31’den de anlaşılacağı gibi COP serisi için birim kök boş hipotezi red edilememekte iken serinin birinci farkı durağandır. Diğer taraftan yukarıda da belirtildiği gibi, şekil 58’de 2001 Kasım ayından itibaren seri pozitif bir trende sahiptir. Bu nedenle üçüncü bölümde veri organizasyonu başlığı altında da tartışılan ağırlıklandırılmış pencere yaklaşımının seriye uygulanmasına karar verilmiştir. Hatırlanacağı gibi, ele alınan zaman serisinin tüm gözlemlerinin veri yaratma süreci altında yatan dinamikler ve önraporlama üzerinde eşit derecede pay sahibi olmadığı ve son elde edilen gözlemlerin eski gözlemlerden daha önemli olduğu düşüncesinden hareketle Morants, Whalen ve Zhang (2004) hareketli pencere yöntemine benzeyen, yeni bir yöntemi ileri sürmüşlerdir. Ağırlıklandırılmış pencere yaklaşımı adlı bu yeni yöntem *önraporlama hatalarına ağırlık verme* (yeni hatalara daha yüksek, eski hatalara daha düşük) mantığıyla çalışmaktadır. Eğitim setinde “çekirdek” olarak adlandırılan en yeni gözlem grubuna %100 ağırlık verilirken, geri kalan gözlemler doğrusal olarak azalan şekilde ağırlıklandırılmaktadır.

Ele aldığımız ham petrol varil fiyatları serisinin çekirdek gözlem grubu yukarıda da belirtildiği gibi 2001 yılı Kasım ayından sonraki dönem olarak belirlenmiştir. Ham petrol varil fiyatlarında bu dönemden sonra izlenen sürekli artışın göz önünde tutularak modellenmesin ve buna bağlı olarak önraporlanmasında bu gözlem grubunu oluşturan veri yaratma sürecinin temel

etkiye sahip olduğu açıktır. Aşağıda şekil 59’da bu uygulamada ele alınan ham petrol varil fiyatları serisinin çekirdek olarak adlandırılan 11 / 2001 – 04 / 2007 döneminin zaman yolu grafiği sunulmuştur.



**Şekil 59: Ham Petrol Varil Fiyatları 2001Kasım – 2007Mayıs**

Diğer taraftan, yine üçüncü bölümde veri organizasyonu başlığı altında tartışıldığı şekilde, çekirdek dışında kalan ancak önraporlama hataları ağırlıklandırılmış bir şekilde modele dahil edilecek grup ise bu uygulamada Aralık 1998 – Ekim 2001 dönemi olarak belirlenmiştir. Şekil 58’den de rahatlıkla görülebileceği gibi Ocak 1991’den başlayarak süre gelen yapının ilk olarak bozulduğu ancak Kasım 2001 sonrası dönemden de farklı bir yapıya sahip bu döneme ait bilgi önraporlama hatası ağırlıklandırılmış biçimde modele alınacaktır.

Böylece t zaman endeksi, s (96. gözlem, Aralık 1998) eğitime alınan veri setinin başlangıç tarihine ait zaman endeksi, c (131. gözlem Kasım 2001) ise çekirdek olarak adlandırılan örneklem grubunun ilk gözleminin zaman endeksi olmak üzere,

$$W_t = \begin{cases} 1 & , & t \geq c \text{ ise} \\ (t - c) / (t - s) & , & s \leq t < c \text{ ise} \\ 0 & , & t < s \text{ ise} \end{cases}$$

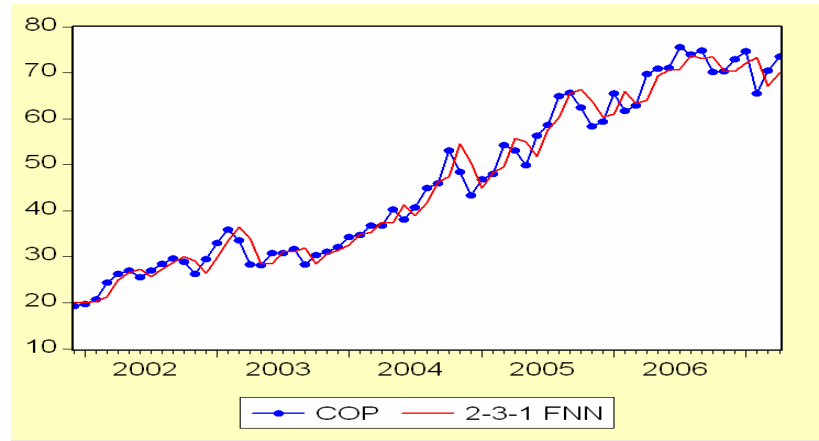
şeklinde elde edilecek ağırlıklar için, bu uygulamadaki önraporlama hataları ağırlık ayarlaması,

$$W_t = \begin{cases} 1 & , & t \geq 131 \text{ ise} \\ [(t-131) / (t-96)] & , & 1 \leq t < 131 \text{ ise} \\ 0 & , & t < 96 \text{ ise} \end{cases}$$

şeklinde yapılmıştır.

Uygulamanın bu başlığı altında model seçimi için herhangi bir yöntem izlenmemiş basit bir ileri beslemeli ağ yapısı 2-3-1 ileri beslemeli ağı kullanılmıştır. Ağ mimarisi açısından, basit bir ağ yapısı ile ara katmandaki yapay sinir hücrelerinin sayısı üç olduğu için ve tek katman kullanıldığı için aşırı uyum problemine yakalanma riski yoktur. Eğitimde iterasyon sayısının fazla olmasından kaynaklanabilecek aşırı uyum probleminin üstesinden gelmek için ise, eğitim 100. iterasyonda kesilmiştir.

Modelden elde edilen örneklem içi önraporlama grafiği aşağıda şekil 60'da sunulmuştur.



**Şekil 60: 2-3-1 FNN ile COP Serisi Örneklem İçi Önraporlama**

Modele ilişkin bazı değerlendirme ölçütleri de aşağıda tablo 32'de sunulmuştur.



**Tablo 32: 2-3-1 İleri Beslemeli Ağı için Bazı Değerlendirme Ölçütleri**

MAE	MAPE	MSE	RMSE
1.94905	0.04476	3.352273	24.28992

*İkinci olarak*, yukarıda ağırlıklandırılmış pencere yönteminde “çekirdek” olarak adlandırılan Kasım 2001 – Nisan 2007 dönemine ait COP serisi ARIMA modelleme yöntemiyle ele alınmıştır. İlk olarak bu döneme ait seriye ADF birim kök testi uygulanmıştır.

**Tablo 33: Çekirdek COP için ADF testi**

Düzy / Birinci Fark	Kesmesiz ve Trendsiz	Kesmeli	Kesme ve Trendli
COP	2.18	- 0.21	-4.14
%95 Güv. Düz. Kritik Değerler	-1.94	-2.89	-3.45
%99 Güv. Düz. Kritik Değerler	-3.17	-3.47	-4.09

Tablo 33’den de görülebileceği gibi çekirdek COP serisi kesme ve trendli modelde durağandır. Bunun anlamı serinin deterministik trend ve durağan bileşenden oluştuğudur.

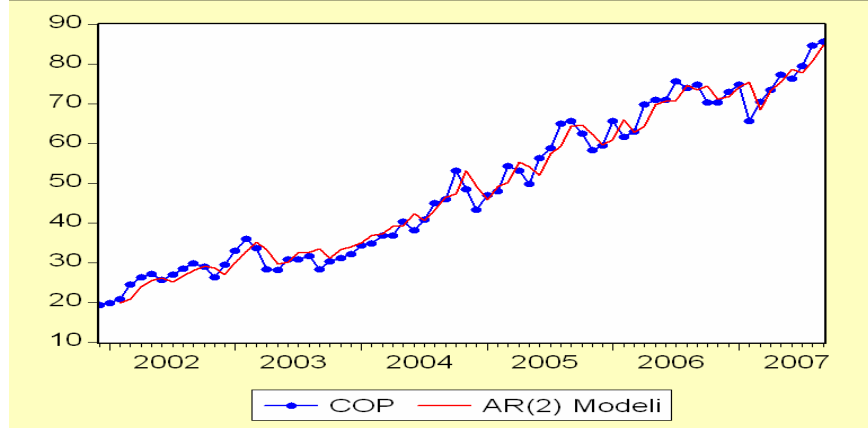
Bu serinin modellenmesinde t zaman kuklası birbirini izleyen dönemlerde birer birer arttırılarak ele alınmış ve deterministik trend için bir kukla değişken kullanılmıştır. Bilgi kriterlerinin kullanımıyla elde edilmiş en iyi model aşağıda sunulmuştur<sup>7</sup>.

$$COP_t = 15.72 + 0.95 t + 0.78COP_{t-1} - 0.06COP_{t-2} + \varepsilon_t$$

(5.43)    (13.96)    (6.23)    (2.66)

<sup>7</sup> Parantez içindeki rakamlar t istatistiklerini göstermektedir.

Yukarıdaki modelin kullanılmasıyla elde edilen örneklem içi önraporlama grafiği şekil 61’de sunulmuştur.



**Şekil 61: AR(2) Modeli ile COP Serisi Örneklem İçi Önraporlama**

Bu önraporlamalara ilişkin önraporlama performans ölçütleri ise aşağıda tablo 34’de sunulmuştur.

**Tablo 34: AR(2) Modeli için Bazı Değerlendirme Ölçütleri**

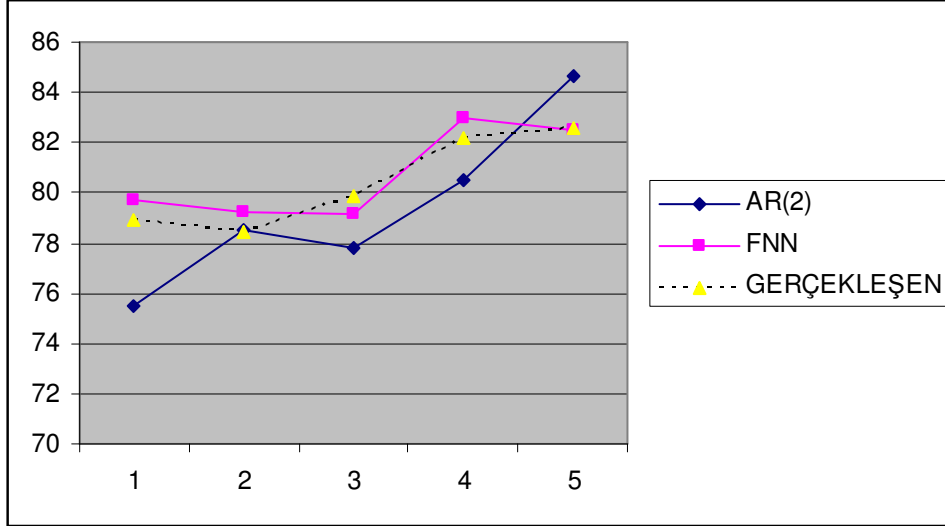
MAE	MAPE	MSE	RMSE
1.980742	0.045735	3.377458	24.38099

Tablo 34’deki değerler, yapay sinir ağı modelinin önraporlama ölçütlerinin sunulduğu tablo 32 ile karşılaştırıldığında ağırlıklandırılmış pencere yöntemi kullanılarak elde edilmiş yapay sinir ağı modelinin önraporlamaları, yaklaşık olarak, MAE ölçütünde %4, MAPE ölçütünde binde 1, MSE ölçütünde %2.5 ve RMSE ölçütünde %10 farklarla daha iyi sonuçlar vermiştir. Son olarak her iki modelin expost dönem önraporlamaları da karşılaştırılmıştır. Aşağıda sırasıyla tablo 35 ve şekil 62 gerçekleşen değerler ile yapay sinir ağı modeli ve AR(2) modeli expost dönem önraporlamalarını 2007 Mayıs – Eylül dönemi için göstermektedir.

**Tablo 35: COP Serisi için Gerçekleşen Değerler ve Expost Önraporlamalar**

AR(2)	FNN	GERÇEKLEŞEN
75.47844	79.71761	78.89245
78.50385	79.20564	78.46268
77.82451	79.17590	79.87104
80.51655	83.01010	82.18780
84.68436	82.50125	82.57473

**Şekil 62: COP Serisi Gerçekleşen Değerler ve Expost Önraporlamalar**



Diebold Mariano (1995) testi ile FNN modelinin expost dönem önraporlama performansı ile AR(2) modelinin performansı arasında bir fark olup olmadığı test edilmiştir. Bu amaçla Diebold ve Mariano'nun önerdikleri test istatistiği bir kez daha kullanılmıştır. Yukarıdaki iki model için S istatistiği,

$$S=(5/4)^{-1/2}(1+0+1+1+1-(5/2)) = 1.3416$$

şeklinde hesaplanmıştır. Bulunan değer istatistiksel olarak ancak %8 anlamlılık düzeyinde FNN modelinin daha iyi önraporlamalar verdiği anlamına gelmektedir.

## SONUÇ

Ekonomik zaman serilerinin sahip olduđu başlıca beş anahtar özellik Trend, Mevsimsellik, Kırılma, Volatilité ve Doğrusal dışılıktır. Zaman serileri analizinin tarihsel hikayesi zaman serilerinin gösterdiği ana özelliklerin modelleme konusunda ortaya çıkardıkları problemlerin çözümlenmesi paralelinde şekillenmiştir. Bu nedenle de tek bir model çerçevesinde konuyu ele alabilme yaklaşımı yetersiz kalmaktadır.

İlk olarak ekonomik zaman serileri için en sık rastlanan ortak özellik pozitif bir trende sahip olmaktır. Elbette bunun en büyük nedeni, iktisadi büyüklüklerin nüfus artışı, teknolojik gelişmeler vb. nedenlerden dolayı artma eğiliminde olmasıdır. Diğer taraftan ekonomide bazı kriz ve benzeri dönemler serilerin ortalamaları üzerinde kalıcı etkiler bırakabilmektedir. Bu durumda problem serinin durağanlığının ADF tipi testlerle test edilememesidir. Çünkü böyle bir durumda seride kırılmanın ortaya çıktığı dönem öncesinde küçük değerler(büyük değerler) yine küçük değerler(büyük değerler) tarafından izlenirken kırılma sonrasında göreceli büyük değerler (küçük değerler) yine göreceli büyük değerler (küçük değerler) tarafından izlenir. Böylece  $y_{t-1}$ 'in otoregresif katsayısının 1'e eşitliğini test eden birim kök testleri birim kökü red etmeme yönünde yanlı olurlar.

İkinci olarak, devreler bir yıldan uzun olmamak şartıyla bir zaman serisindeki tekrarlanan döngüsel hareketlerin tümüne mevsimsel dalgalanma adı verilir. Periyodik otoregresif model (PEAR), AR modelinin otoregresif parametresinin zaman içerisinde mevsimler boyunca dinamikleştirilmesine dayanır. Böylece, örneğin çeyrek yıllık veri seti ile çalışıldığı durumda, her bir çeyreğin diğerlerine göre göreceli etkisi ayrıştırılabilir. Aslında bu yöntem ekonomik birimlerin farklı mevsimlerde farklı davranış kalıpları göstermelerinden dolayı oldukça kullanışlı bir yoldur.

Üçüncü olarak, ele alınan zaman serisi sabit varyansa sahip olmadığı durumda bildik yöntemler modellemede ve öngörü yapmada başarılı olamamaktadır. Otoregresif koşullu değişen varyans (ARCH) modelleri ise varyansa öngörü hatalarının karelerinin bir fonksiyonu olarak değişim serbestliği getirmektedir. GARCH modelleri ise, koşullu varyansın ARMA yapısında da olduğu durumda tahminine olanak sağlamaktadır.

Dördüncü olarak, İktisatta kullanılan istatistiksel ve ekonometrik yöntemler, doğrusallık varsayımı ile aslında iktisat teorisinin bazen arka planda zimni olarak kabul ettiği bazen de direkt vurguladığı doğrusal dışı ilişkileri göz ardı etmektedir. Bu durum gerçek hayata ilişkin fenomenlerin modellemesinde kimi zaman ciddi sınırlamalar yaratmaktadır. Örneğin, ücretlerin aşağıya doğru gösterdiği yapışkanlık, ekonominin genişleme dönemlerinde işsizlik oranının azalmasındaki yavaşlığa rağmen, daralma dönemlerinde işsizlik oranındaki artışın hızlı olması ya da kısa ve uzun dönemli faiz oranlarındaki farkın bir eşik değere göre farklı hareket etmesi ve benzeri iktisat teorisinden gelen açılımlar, doğrusal modelleme araçlarıyla modelleme yapıldığında göz ardı edilmiş olmaktadır. Bu durumda veri yaratma sürecinin TAR modeli olarak ele alınması, otoregresif katsayıya, örneğin  $y_{t-1} > 0$  durumunda  $\phi_1$ ,  $y_{t-1} \leq 0$  durumunda ise  $\phi_2$  gibi ayrı değerler alma şansı tanıyan bir sürecin işletilmesini sağlar. Diğer taraftan, TAR modelinin iktisat teorisince vurgulanan rejim değişikliklerini yakalamadaki başarısına karşın, bu rejim değişikliklerini hızlı yada başka bir ifade ile keskin bir şekilde olduğunu varsayması eleştirilebilir. Bu nedenle bazı durumlarda, istatistiksel modellemenin sunduğu STAR modelleri ailesi kullanışlı bir araçtır. STAR modelleri, rejim geçişlerinin daha yumuşak bir şekilde gerçekleşmesine izin veren modellerdir. Bu nedenle daha sofistike bir modelleme aracı olarak görülürler.

Diğer taraftan son on yılda, zaman serisi analizinde yeni ve kullanışlı bir modelleme aracı olarak Yapay Sinir Ağı modellerinden sıklıkla ve olumlu bir şekilde söz edilmektedir. Yüksek örüntü tanıma kabiliyeti, evrensel yakınsayıcı özelliği ve bu nedenle de her tür karmaşık ilişkiyi yakalayabilme yeteneği,

regresyon yaklaşımının tersine veri setinin dağılımına ve tahmin edilecek parametrelere doğrusallığa ilişkin herhangi bir varsayım gerektirmemesi gibi çekici özelliklerinin yanında, kendine özgü farklı bir terminolojisinin olması, tüm sürecin bilgisayarlarda işlediğine dair kara kutu yakıştırması, parametrelerinin yorumlanamaması gibi dezavantajları olan bu modelleme ailesinin daha yakından incelenmesi gerekmektedir.

Yapay sinir ağları insan beyninin işleyişinden esinlenerek ortaya konulmuş modellerdir. İnsan beyni sinir hücrelerinin ve bu hücreler arasındaki bağlantıların işleyişiyle çalışmasını gerçekleştirir. İnsan beyninin avantajı, yüksek derecede paralel hesaplama yapabilen bir yapıya sahip olma ve bununla ortaya çıkan bilgi işleme yeteneğidir 1940'lardan önce bazı yapay sinir ağı çalışmaları olsa da, bu alanda ilk ciddi çalışmanın 1943 yılındaki Warren McCulloch ve Walter Pitts tarafından yapıldığı kabul edilir. Bu çalışmayı, Rosenblatt tarafından geliştirilen Basit algılayıcı modeli ve Windrow ve Hoff tarafından geliştirilen ADALİNE izlemiştir. Yapay sinir ağlarının zaman serisi analizinde ciddi bir şekilde kullanıldığı çalışmalardan ilki 1988 yılında White tarafından yapılmıştır. İlgili yapay sinir ağı modelinde kaç katman kullanılacağı ve katmanlardaki sinir hücrelerinin sayısının ne olacağı gibi araştırmacının bilgi tecrübe ve becerisine bırakılmış yada diğer bir ifade ile subjektivitenin hakim olduğu ve bir takım sorulara tam bir yanıtın olmayışı, bu yıllarda araştırmacıları yapay sinir ağı modellerine biraz soğuk bakmasını beraberinde getirse de, 1989-93 yılları arasında Cybenko (1989), Funahashi (1989), Hornik (1989, 1991, 1993) ve Hornik, Stinchcombe and White (1989, 1990), tarafından yapılan teorik çalışmalar bu anlamdaki şüpheleri gidermiştir. Bu tarihten sonra da finansal ve ekonomik temelli çalışmalar hızla artmıştır.

Zaman serisi analizi konu başlığının içerisinde değerlendirilebilecek ve yapay sinir ağı modellerinin kullanıldığı bilimsel çalışmalarda kullanılan ağlar, çok katmanlı algılayıcılar ve geri dönüşümlü ağlardır. Bu ağlar öğretmenli öğrenme stratejisi ile çalışmaktadır. Öğretmenli öğrenme stratejisi ise İstatistiksel Öğrenme Teorisine dayanmaktadır. İstatistiksel Öğrenme Teorisinin amacı kullanılacak öğrenme algoritmasının özellikleri üzerine istatistiksel bir çatı

oluşturmaktır.

Çok katmanlı algılayıcılar yapay sinir ağları modelleri ailesi içerisinde en çok kullanılan ağ tipidir. Aslında tek katmanlı algılayıcının genelleştirilmiş bir hali olarak da düşünülebilir. Bu ağlarda girdi sinyali katmanlar boyunca ileriye doğru işlenerek çıktı üretilmektedir. Aslında ileri beslemeli bir ağ transfer fonksiyonlarının katmanlar olarak düzenlendiği bir yapıdır. İleri beslemeli ağlar için, transfer fonksiyonları üzerinden tanımlanan çok önemli bir konu da bu ağların evrensel yakınsayıcı özellikleridir. *Evrensel yakınsama teoremi*, ileri beslemeli bir ağın yeterli sayıda gizli katman ve bu katmanlarda yeterli sayıda nöron ile doğru ayarlanmış parametreler ve bir transfer fonksiyonu yardımıyla seçilen herhangi bir fonksiyona yakınsayabileceğini belirtmektedir.

Verili bir ağ yapısı ve belirlenmiş transfer fonksiyonları için en önemli konu ağırlık tahminlerinin elde edilmesidir. Bu anlamda en sık kullanılan algoritma geriye yayılım algoritmasıdır. Bununla beraber doğrusal olmayan en küçük kareler yöntemi de geniş kullanım alanı bulunan ve bazı üstünlüklere sahip olan bir yöntemdir.

Öte yandan ağın ne tip bir yapıda olacağı konusunda da çok kesin kurallar bulunmamaktadır. Burada ağın yapısı ile kast edilen, ağın kaç katmandan oluşacağı ve katmanlardaki nöron sayısının ne olacağı hususudur. Fonksiyon uzayındaki karmaşıklığın ağ ile yakalanması yanında aşırı uyum probleminden kaçınılması gerekmektedir. Bu nedenle dene gör yöntemi gibi subjektifliğe açık bir yaklaşım yerine belli başlı kriterlerle bu konunun açıklığa kavuşturulması önem arz etmektedir. Böylece farklı araştırmacıların farklı ağ seçimi yada farklı ağa güvenme karmaşası engellenebilir. Bu amaçla Teorik Bilgi Kriterleri, Çapraz Geçerlilik Yöntemi, Erken Durdurma Metodu ve bazı Hata Kriterleri kullanılmaktadır.

Uygulama bölümünün değerlendirilmesine geçmeden basitçe özetlemek gerekirse, bir zaman serisi ile çalışıldığında, ilk olarak serinin yukarıda tartışılan anahtar özelliklerden hangisi yada hangilerine sahip olduğunun belirlenmesi gerekir. İkinci olarak uygun gecikme yapısına karar verilir. Bu iki aşama

sonucunda veri yaratma süreci tanımlanmış olur. Bu noktadan itibaren açıklanması gereken serinin, ilgili veri yaratma süreci altında birim kök içerip içermediğidir. Eğer seri ( $y_t$ ) birim kök içermiyorsa uygun bulunan veri yaratma süreci ile, eğer birim kök içeriyorsa da fark sers ( $\Delta y_t$ ) kullanılarak tek değişkenli modelleme ve buna bağlı olarak önraporlama elde edilir.

Modelleme aracı olarak yapay sinir ağlarının kullanılması durumunda, başlangıçta cevap verilmesi gereken sorular; girdi katmanında kaç nöron kullanılacağı, ara katman sayısının ne olacağı, ara katmanlarda kaç nöron kullanılacağı, çıktı katmanının da kaç nöron olacağı ve bu nöronların hangi transfer fonksiyonlarını kullanacağı ile eğitimde hangi algoritmanın tercih edileceğidir.

Buradan da anlaşılacağı gibi, yapay sinir ağlarının iktisadi zaman serilerinde modelleme aracı olarak kullanımı, kendilerine özgü sorunlara sahip iki ayrı alanın ortaklaşa ele alınmasını gerektirmektedir. Özellikle geleneksel olarak niteleyebileceğimiz zaman serisi analiz yöntemlerinin işin en başında kabul ettikleri fonksiyonel formun yanlış olabilme durumu yada veri setindeki sıradışı gözlemlerin tüm parametre tahminlerini yanlış kılma tehlikesi yapay sinir ağı modellerini güvenilirlik konusunda bir adım ileri çıkarmaktadır. Bu çalışmada uygulama bölümü için bir el haritası oluşturmak amacıyla 5 adımlık bir prosedür düzenlenmiştir. Bu adımsal prosedür; Veri ön işleme, Veri organizasyonu, Ağ mimarisinin Oluşturulması, Eğitim ve Değerlendirme adımlarından oluşmaktadır.

Uygumla bölümü, birinci bölümde sunulan ve zaman serilerinin sahip olduğu anahtar özellikler olarak adlandırılan, trend, mevsimsellik, yapısal kırılma, volatilité ve doğrusal dışı özellikler durumlarında alışıla gelmiş olarak kullanılan zaman serisi modelleriyle, yapay sinir ağı modellerinin performanslarının karşılaştırılması üzerine uygulamalardan oluşmuştur. Böylece, zaman serisi analizinde yapay sinir ağı modellerini kullanmanın avantajlı olduğu durumların belirlenmesi ve literatürdeki bu alandaki tartışmalara gerçek seriler üzerinden kanıt sunulması hedeflenmektedir. Bu amaçla, Türkiye ekonomisine ait ve yukarıda sayılan anahtar özelliklere sahip beş ayrı makroekonomik zaman serisi



kullanılmıştır.

Uygulamanın ilk alt başlığı altındaki amaç, mevsimsellik içeren serilerin yapay sinir ağı modelleriyle ele alınması durumunda önraporlama performansında daha iyi sonuçlar elde edilip edilemeyeceğinin ölçülmesidir. Bu nedenle Türkiye ekonomisine ait Özel Nihai Tüketim serisi kullanılmıştır. Üçüncü bölümde yapay sinir ağları ile yeniden bakış – Veri önışleme adımı başlığı altında da değinildiği gibi Shadra ve Patil (1992), mevsimsel düzeltme yapılmasa da, iyi önraporlamalar elde edilebileceğini, Stock ve Watson (1998) ve Tkacz (2001), mevsimsel düzeltme yapıldığında yapay sinir ağı modellerinin önraporlama değerlerinin bozulacağını ileri sürmektedirler. Öte yandan Farway ve Chetfield (1995), Nelson v.d (1999), Zhang ve Qi (2005) mevsimsel düzeltmenin yapılmış olduğu serilerle çalışıldığında yapay sinir ağlarının çok daha yüksek önraporlama performansına sahip olduğunu belirtmektedirler. Hatta bu arařtırmacılar mevsimsellikte düzeltme yapılmadığı durumda yapay sinir ağı modellerinin, ARIMA modellerinden daha kötü performans gösterdiklerini ileri sürmektedir. Bu nedenle, PEAR modelinin önraporlama performans ölçütleriyle, bu ikinci adımda mevsimsel düzeltme yapılmamış seriye uygulanan MLP modelinin önraporlama performans ölçütleri karşılaştırılmıştır. Bulunan bu ampirik sonuçlar, literatürdeki “mevsimsel serilerde yapay sinir ağı modelleri kullanılabilir mi ?” tartışmasında, kullanılabilir diyenlerin tarafını desteklemektedir.

Uygulamanın ikinci alt başlığındaki amaç Volatilite modellemesinde yapay sinir ağı modellerinin kullanılıp kullanılamayacağını arařtırmaktır. Bu amaçla İMKB 100 haftalık getiri serisi kullanılmıştır. Bu başlık altında tahmin edilen 3-4-1 ileri beslemeli ağı GARCH (1,1) modelinden daha kötü performans göstermiştir. Üstelik ağı eğitiminde gözlem sayısı yeterince büyük olduğunda kullanılabilen ayrıca, Nakamura (2005) ve Gonzalez (2000) tarafından aşırı uyum problemine takılmadan ağıın hem eğitimini hem de katmanlardaki nöron sayısını belirlemeye yarayan çapraz geçerlilik ve erken durdurma metotlarının beraber kullanılmış olmasına rağmen, varyanstaki doğrusal dışılığı da bir bilgi olarak kullanabilen GARCH modeli, volatilite modellemesinde yapay sinir ağı

modellerinden daha iyi çalışmaktadır.

Uygulamanın üçüncü alt başlığındaki amaç yapısal kırılmalı bir serinin modellemesinde yapay sinir ağı modellerinin kullanılıp kullanılmayacağını araştırmaktır. Bu amaçla Türkiye ekonomisine ait Sanayi Üretim Endeksi kullanılmıştır. Tahmin edilen 4-6-1 ileri beslemeli ağının örneklem içi ve ex-post önraporlama performansının daha iyi olduğu sonucuna varılmıştır. Bu sonuç, yapısal kırılmalı doğrusal serilerde de yapay sinir ağlarının önraporlama amacıyla kullanılabileceğini desteklemektedir.

Uygulamanın dördüncü alt başlığındaki doğrusal dışı bir serinin modellemesinde yapay sinir ağı modellerinin kullanılıp kullanılmayacağını araştırmaktır. Bu amaçla Türkiye ekonomisine ait Kapasite Kullanım Oranları serisi kullanılmıştır. Serinin doğrusal dışı bir yapıya sahip olup olmadığı test edildikten sonra seri uygun bulunan TAR, LSTAR ve RNN modelleriyle modellenip örneklem içi ve ex-post periyot önraporlamaları karşılaştırılmıştır. İlginç bir şekilde LSTAR ve RNN modellerinin toplam mutlak sapmaları ex-post dönem için eşit çıkmıştır. Ex-post periyot için TAR modelinin bu iki modelden farkı ancak (dört dönem toplamı için) 0.0141 olarak elde edilmiştir. Uygulama sonucunda RNN ve LSTAR başta olmak üzere üç modelde başarılıdır denilebilir.

Uygulamanın son alt başlığındaki amaç birim köke sahip bir seride yapay sinir ağı modellerinin kullanılıp kullanılmayacağını ağırlıklandırılmış pencere yöntemiyle ele alınmasıdır. Bu amaçla ham petrol varil fiyatları serisi yapay sinir ağı modeliyle modellenmiştir. İkinci olarak aynı serinin birim kök sahip olmayan en geniş son penceresi ise ARIMA modelleme yaklaşımıyla modellenmiştir. Örneklem içi ve örneklem dışı önraporlamalarında yapay sinir ağı modeli daha doğru önraporlamalar üretmiştir.

Sonuçta yapay sinir ağları modellerinin ilgili seride mevsimsellik, yapısal kırılma, doğrusal dışılık gibi anahtar özellikler bulunması durumunda başarı ile kullanılabileceği, ancak volatiliteye sahip serilerde ise varyanstaki doğrusal dışı yapının yapay sinir ağları ile yakalanamadığı ortaya konulmuş ve

ayrıca birim köke sahip serilerde ağırlıklandırılmış pencere yönteminin önraporlama performansını arttırabilecek şekilde kullanılabileceği belirlenmiştir. Uygulamada gerçek seriler kullanıldığı için, çalışma sonuçlarının bu seriler özelinde değerlendirilmesi gereği de açıktır.

## **Terimler Sözlüğü**

***Adaptive*** : Uyarlayıcı

***Adaptive Linear Element (ADELINE)*** : Uyarlayıcı Doğrusal Eleman

***Associative Memory*** : Çağrışimli Bellek

***Backpropagation Algorithm*** : Geriye Yayılım Algoritması

***Bias*** : Yan

***Feedforward Neural Network*** : İleri Beslemeli Ağ

***Gradient Descent*** : Dereceli Azaltma

***Learning*** : Öğrenme

***Linear Vector Quantization Model*** : Doğrusal Vektör Parçalama Modeli

***Pattern Recognition*** : Örüntü Tanıma

***Perceptron*** : Basit Algılayıcı

***Reinforcement Learning*** : Destekleyici Öğrenme

***Self Organization Features Map*** : Kendi kendini organize edebilen Özellik Haritası

***Structural Risk Minimization*** : Yapısal Risk Minimizasyonu

***Supervised Learning*** : Öğretmenli Öğrenme

***Training*** : Eğitim

***Unsupervised Learning*** : Öğretmensiz Öğrenme

## **KAYNAKÇA**

Abraham, A. (2005); “Artificial Neural Networks”, **Handbook of Measuring System Desing**, John Willey and Sons, New York.

Abu-Mostafa, Y.S. (1986); “Neural Networks for Computing”, In Proceedings of the American Institute of Physics Meeting, pp. 1-6.

Amari, S. (1994); “[Neural Networks: A Review from Statical Perspective]: Comment”, **Statistical Science**, Vol. 9, No. 1, pp. 31-32.

Anthony, M. and Bartlett, P.L. (1999); **Neural Network Learning: Theoretical Foundations**, Cambridge University Pres, Cambridge.

Arifovic, J. and Gencay R. (2001); “Using Genetic Algorithms to Select Architecture of a Feedforward Artificial Neural Network”, **Physica (A)**, 289, pp. 574-594.

Arrold, M. (2003); “Reasoning About Nonlinear AR Models Using Expectation Maximization”, **Journal of Forecasting**, 22, pp. 479-490.

Ashworth, J. and Thomas, B. (1999); “Patterns of Seasonality in Employment in Tourism in the UK”, **Applied Economic Letters**, 6, pp. 735-739.

Azoff, E.M. (1994); **Neural Network Time Series Forecasting of Financial Markets**, John Willey and Sons, West Sunsex.

Barron A. R. and Sheu C. (1991). Approximation of density functions by sequences of exponential families. **Annals of Statistics**, Vol.19, pp.1347-1369.

Barron, A.R. (1994); “[Neural Networks: A Review from Statistical Perspective]: Comment”, **Statistical Science**, Vol. 9, No. 1, (February), pp. 33-35.

Beinenstock, E. and Geman, S. (1994); “[Neural Networks: A Review from Statistical Perspective]: Comment”, **Statistical Science**, Vol. 9, No. 1, (February), pp. 36-38.

Bollerslev, T. (1986); “Generalized Autoregressive Conditional Heteroskedasticity”, **Journal of Econometrics**, 31, pp. 307-327.

Boswijk H. P., Franses P.H., Haldrup N. (1997): “Multiple Unit Roots in Periodic Autoregression”, *Journal of Econometrics*, vol:80, 1, 167-193.

Bousquet, O., Bouchenen, S. and Lugosi, G. (2003); “Introduction to Statistical Learning Theory”, <http://www.kyb.mpg.de/publications/pdfs/pdf2819.pdf> Erişim Tarihi (23.07.2007).

Box J. (1996): “Discussion on the Papers on 'Statistics and Mathematics'”, *Journal of the Royal Statistical Society: Series D (The Statistician)*, Vol. 47 Issue 2 Page 273 July 1998

Breiman, L. (1994); “[Neural Networks: A Review from Statistical Perspective]: Comment”, **Statistical Science**, Vol. 9, No. 1, (February), pp. 38-42.

Branson, R. (1991); **Matrix Methods**, Academic Press, INC, San Diego.

Bu Hamra, S., Smaoui, N. and Gabr, M. (2003); “The Box-Jenkins Analysis and Neural Networks: Prediction and Time Series Modelling”, **Applied Mathematical Modelling**, 27, pp. 805-815.

Caner, M. ve Hansen, B. E 1997. "Threshold Autoregressions with a Unit Root," **Boston College Working Papers in Economics 381**, Boston College Department of Economics.

Cao, C.Q. and Tsay, R.S. (1992); “Nonlinear Time Series Analysis of Stock Volatilities”, **Journal of Applied Econometrics**, 7, pp. 165-185.

Castro, T.B. and Osborn, D. (2005); “Cointegration for Periodically Integrated Processes”, *The School of Economics Discussion Paper Series 0522*, School of Economics, The University of Manchester.

Chan, M.C, Wong, C.C. and Lam C.C. (2000); “**Financial Time Series Forecasting by Neural Network Using Conjugate Gradient Learning Algorithm and Multiple Linear Regression Weight Initialization**”, *Computing in Economics and Finance*

Chatfield, C. (1997). **The analysis of time series: An introduction**. Chapman and Hall, New York.

Cheng, B. and Titterington, D.M (1994a); "Neural Networks: A Review From a Statistical Perspective", **Statistical Science**, Vol. 9, No. 1, (pp. 2-30).

Cheng, B. and Titterington, D.M (1994b); "Neural Networks: A Review From a Statistical Perspective", Rejoinder, **Statistical Science**, Vol. 9, No. 1, (pp. 49-54).

Clements, M.P., Franses, P.H. and Swanson, N.R. (2003); "Forecasting Economic and Financial Time Series with Nonlinear Models", **International Journal of Forecasting**, Vol,20, pp.169,183

Clements, M.P. and Smith, J. (2001); "Evaluating Forecast from SETAR Models of Exchange Rates", **Journal of International Money and Finance**, 20, pp. 133-149.

Csaji, B.C. (2004); Approximation with Artificial Neural Networks, Msc Thesis, Faculty of Mathematics and Computing Science, Eindhoven University of Technology, Netherland [http://www.sztaki.hu/~csaji/CsBCs\\_MSc.pdf](http://www.sztaki.hu/~csaji/CsBCs_MSc.pdf) Erişim Tarihi:(18.05.2005)

Curram S. P., ve Minger J. (1994); "Neural Networks, Decision Tree Induction and Discriminant Analysis: Emprical Comprasion", **Operational Research Society**, Vol:45, No:4, pp. 440-450

DeLugrio A.S., (1998); **Forecasting Principles and Applications**, McGrew Hill, NewYork.

Dickey, P.A. (2005); "Stationarity Issues in Time Series Model", Proceedings of the 13. Annual SAS Users Group, International Conferance, SAS Institute Inc. Gray, New York.

Enders, W. (2004); **Applied Econometric Time Series**, 2. Edition, John Willey and Sons, New York.

Engel A. ve Van denBroeck C. (2001); **Statistical Mechanics of Learning**, Cambridge University Press, Cambridge

Engle, R.F. (1982); “Autoregressive Conditional Heteroskedasticity with Estimates of Variance of UK Inflation”, **Econometrica**, 50, pp. 987-1007.

Evgeniou, T., Pontil, M. and Poggio, T. (1999); A unified framework for regularization networks and support vector machines. **CBCL Paper 171/AI Memo 1654**, Massachusetts Institute of Technology, Cambridge, MA.

Evgeniou, T., Pontil, M. and Poggio, T. (2000); “Statistical Learning Theory: A Primer”, **International Journal of Computer Vision**, 31-1, pp. 9-13.

Faraway, J. and Chatfield, C. (1998); “Time Series Forecasting with Neural Networks: A Comparative Study Using the Airline Data”, **Applied Statistics**, Vol. 47, No. 2, pp. 231-250.

Feng, H. and Liu, J. (2002); “A SETAR Model for Canadian GDP: Non Linearities and Forecast Comparisons”, **Working Paper EWP 0206**, University of Victoria.

Francis, L. (2001); “The Basics of Neural Networks Demystified”, **Casualty Actuarial Society**, Contingencies, November, (pp. 56-61).

Franses, P.H. (1998); **Time Series Models for Business and Economic Models**, Cambridge University Pres, Cambridge.

Franses, P.H. (2000); **Non Linear Time Series Models in Empirical Finance**, Cambridge University Pres, Cambridge.

Fu, Q., Fu, R.L., and Sun, Y. (2004); “Self-Exciting Thershold Autoregressive Model (SETAR) to Forecast the Well Irrigation Rice Water Requirement”, **Nature and Science**, 2 (1), pp. 36-43.

Gaynor, P.E, Kirkpatrick, R.C. (1994); **Introduction to Time Series Modelling and Forecasting in Business and Economics**, McGraw-Hill, New York.

Gencay, R. and Liu, T. (1997); “Nonlinear Modeling and Prediction with Feedforward and Reccurent Networks”, **Physica (D)**, 108, pp. 119-134.



Gencay, R., Selçuk F.ve Whitcher B. (2002); *An Introduction to Wavelets and Other Filtering. Methods in Finance and Economics*, Academic Press, New York

Ghahramani, Z. (2004); “Advanced Lectures in Machine Learning”, **Unsupervised Learning**, Ed. O. Basquet, U. Van Luxburg and G Raetsch, Springer Verlag, pp. 71-112.

Gonzales, S. (2000); “Neural Networks for Macroeconomic Forecasting A Complementary Approach to Linear Regression Models”, **Finance Canada Working Paper**, No. 2000-07.

Gonzales, D. S. and Costo, A.V. (2001); “Comparing Neural Networks and Efficiency Techniques in Nonlinear Production Functions”, **7. European Workshop on Efficiency and Productivity Analysis**, at University of Oviedo, September, pp. 25-27.

Gonzalo, J. and Pitarakis, J.Y. (2005); “Lag Length Estimation in Large Dimensional Systems”, **Journal of Time Series Analysis**, Vol. 23, Iss. 4, pp. 401-423.

Gren, W.H. (1997); **Econometric Analysis**, 3. Edition, Prentice-Hall, New Jersey.

Hagan, M.T. and Demuth, H.B. (1995); **Neural Network Design**, International Thomson Publishing, London.

Hammer and Villmann (2003); **Mathematical Aspects of Neural Networks**, ESANN, Bruges, Belgium.

Hand, D.T. (1998); “Breaking Misconceptions – Statistics and Its Relationship to Mathematics”, **The Statistician**, Vol. 47, No. 2, pp. 245-250.

Hansen, B.E. 1997. "Threshold effects in non-dynamic panels: Estimation, testing and inference," Boston College Working Papers in Economics 365, Boston College Department of Economics.

Hatanaka, M. (1996); *Advanced Text in Econometrics: Time Series Based Econometrics*, Ed: C.W.J. Granger, G.E. Mizan, Oxford University Pres, New

York.

Haug, A.A. and Siklos, P.L. (2006); “The Behaviour of Short Term Interest Rates: International Evidence of Nonlinear Adjustment”, *Studies in Nonlinear Dynamics and Econometrics*, Vol. 10, Iss. 4, Article 6.

Haykin, S. (1999); **Neural Networks**, Prentice – Hall, London.

Heravi, S., Osborn, D.R. and Birchenhall, C.R. (2004); “Linear Versus Neural Network Forecasts for European Industrial Production Series”, **International Journal of Forecasting**, 20, pp. 435-446.

Herbrich, R. et al. (1999); “Neural Networks in Economics, Background, Applications and New Developments”, <http://www.smartquant.com/references/NeuralNetworks/neural7.pdf> Erişim Tarihi:(20.08.2004).

Hertz J. A., Palmer R.G., Krogh A. (1991); **Introduction to the Theory of Neural Computations**, Westview Press, USA

Hill, T., O’Conner, M. and Remus, W. (1996); “Neural Network Models for Time Series Forecasts”, **Management Science**, Vol. 42, No. 7, pp. 1082-1092.

Hylleberg, S. (1992); **Modelling Seasonality**, Ed. by S. Hylleberg, Oxford University Press, New York.

Irwin, G., Lightbody, G. And McLoone, S. (1994); “Comparison of Gradient Based Training Algorithms for Multilayer Perceptrons”, <http://www.smartquant.com/references/NeuralNetworks/neural7.pdf>, Erişim Tarihi: (26.11.2005).

İşığışok, E. (1999); Türkiye’de Enflasyon’un Varyansının ARCH ve GARCH Modelleri ile Tahmini, **Uludağ Üniversitesi İ.İ.B.F. Dergisi**, Cilt:17, Sayı:3 pp.1-17.

Jordan, M. and Bishop C. (1996); “Neural Networks”, **CRC Handbook of Computer Science**, Ed. A. Truceker, CRC Press, Boca Raton, pp.1-29

Kaashock, J.F. and van Dijk, H.K. (2001); **Neural Networks as Econometric Tool**, Econometric Institute Report EI 2001-05.

Kaelbling, L.P., Litman, M.L. and Moore, A.W. (1996); “Reinforcement Learning: A Survey”, **Journal of Artificial Intelligence Research**, 4, (pp. 237-285).

Kantz, H. and Schreiber, T. (2003); **Nonlinear Time Series Analysis**, Cambridge University Press, Cambridge.

Kaasra, I. And Boyd, M. (1996); “Designing a Neural Network for Forecasting Financial and Economic Time Series”, **Neurocomputing**, 10, pp. 215-236.

Kenny, B. and Durbin, J. (1982); “Local Trend Estimation and Seasonal Adjustment of Economic and Social Time Series”, **Journal of the Royal Statistical Society Series A**, Vol. 145, No. 1, pp. 74-85.

Klimasauskas, C. (1991); Neural Nets Tell Why, **Dr. Dobb’s Journal**, April 1991, pp. 16-24

Kline, D.M. (2004); “Methods for Multi-Step Time Series Forecasting with Neural Network”, **Neural Networks in Business Forecasting**, Ed. by G.P. Zhang, Idea Groups, Hershey.

Koskivaara E. (2000); “Artificial neural network models for predicting patterns in auditing monthly balances”. **Journal of Operational Research Society**, Volume 51, Number 9, September 2000, pp. 1060-1069.

Kuan, C.M. (2006); “Artificial Neural Networks”, **IEAS Working Paper**, No. 06-A010, September, Institute Economics Academia Sinica.

Kuan, C.M. and Liu T. (1995); “Forecasting Exchange Rates Using Feedforward and Recurrent Neural Networks”, **Journal of Applied Econometrics**, Vol. 10, No. 4, pp. 347-364.

Kurozumi, E. (2002); “Testing for Periodic Stationarity”, **Econometric Reviews**, Vol. 21, No. 2, pp. 243-270.

Lawrence, A.D. Back, Tsoi A.C. and Giles C. L. , “On the distribution of performance from multiple neural network trials”, *IEEE Trans. on Neural Networks*, Vol. 8, No. 6, pp. 1507-1517, 1998.

Liew, S.K.V., Baharumshah, Z.A. and Lau, S.H. (2002); “Forecasting Reformance of LSTAR Exchange Rate Model: The Original and Reparametrised Versions”, **MPRA Paper**, No. 511.

Linderberg B.C. (1982); “International Comprasion of Growth in Demand for A new Durable Consumer Product”, **Journal of Marketing Research**, Vol.19, No:3, pp.364-371.

Loh, A.P. and Fong, K.F. (1993); “**Backpropagation Using Generalized Least Squares**”, IEEE International Conference on Neural Networks, San Francisco CA (1993), pp.592-597.

Madsen, K., Nielsen, H.B. and Tingleff, O. (2004); **Methods for Nonlinear Least Squares Problems**, 2. Edition, Informatics and Mathematical Modelling Technical Universtiy of Denmark.

Mc Cabe, M. and Tremayne, A. R. (1995); “Testing a Time Series for Difference Stationarity”, **The Annals of Statistics**, Vol. 23, No. 3, pp. 1015-1028.

Mc Clelland, J. (1994);”[Neural Networks: A Review from a Statistical Perspective]: Comment: Neural Networks and Cognitive Science: Notivations and Applications”, **Statistical Science**, Vol. 9, No. 1, (pp. 42-45).

Mikhail, O., Eberwein, C.J. and Handa J. (2003); “On the Evidence of Nonlinear Structure in Canadian Unemployment”, [www.bus.ucf.edu/omikhail](http://www.bus.ucf.edu/omikhail), Erişim tarihi: 06.08.2005.

Mizrach, B. (2006); **Nonlinear Time Series Analysis**, Prepared for the New Palgrave Dictionary of Economics, 2. Edition [ftp://snde.rutgers.edu/Rutgers/wp/2006-04.pdf](http://snde.rutgers.edu/Rutgers/wp/2006-04.pdf) Erişim Tarihi: 15.01.2007

Morantz, B.H., Whalen, T. and Zhang, G.P. (2004); “A Weighted Window Approach to Neural Network. Time Series Forecasting”, **Neural Networks in Business Forecasting**, Ed. by. G.P. Zhang, Idea Group, Hershey, pp. 251-265.

Moshiri, S., Cameron, N.E. and Scuse, D. (1999); “Static, Dynamic and Hybrid Neural Networks in Forecasting Inflation”, **Computational Economics**, 14, pp. 219-235.

Nakamura, E. (2005); "Inflation Forecasting Using a Neural Network", **Economics Letters**, 86, pp. 373-378.

Ng, S.C., Leung, S.H. and Luke, A. (1999); "A Generalized Back Propagation Algorithm for Faster Convergence", **Neural Processing Letters**, Vol. 9, Number 1, pp.13-23.

Oja, E. (2002); "Unsupervised Learning in Neural Computation", **Theoretical Computer Science**, 287, (pp. 187-207).

Oren, S.S. (1973); "Quasi Newton Algorithms: Approaches and Motivations", **Proceedings of the IEEE Conference on Decision and Control**, San Diego, CA (December / 1973)

Öztemel, E. (2003); **Yapay Sinir Ağları**, Papatya Yayıncılık, İstanbul.

Qi, M. (2001); "Predicting US Recessions with Leading Indicators Via Neural Network Model", **International Journal of Forecasting**, 17, pp. 383-401.

Qi, M. and Zhang, G.P. (2001); "An Investigation of Model Selection Criteria for Neural Network Time Series Forecasting", **European Journal of Operational Research**, 132, pp. 666-680.

Patterson, K. (2005); **An Introduction – Applied Econometrics: A Time Series Approach**, St. Martin's Press, New York.

Pednault, E.P.D. (1998); "Statistical Learning Theory", MIT **Encyclopedia of the Cognitive Sciences**, Ed. M. Jordan and S. Russell, AI/Computer Science, MIT press, Cambridge.

Pena, D. And Rodriguez, J. (2006); "Detecting Non Linearity in Time Series by Model Selection Criteria", <http://halweb.uc3m.es/esp/Personal/personas/dpena/articles/ijfnonlinearityf.pdf> Erişim Tarihi: (12.02.2007)

Perron, P. (1989); "The Great Crash the Oil Price Shock, and the Unit Root Hypothesis", **Econometrica**, Vol:57, November 1989, pp. 1361-1401.

Perron, P. and Vogelsang, T.J. (1992); "NonStationary and Level Shifts with an Application to Purchasing Power Parity", **Journal of Business and Economic Statistics**, 10, pp.310-320.

Pindyck, R.S. and Rubinfeld, D.L. (1981); **Economic Models and Economic Forecasts**, McGraw-Hill, Kogakusha.

Poon S. H., Granger C. W. (2003); "Forecasting Volatility in Financial Markets: A Review", *Journal of Economic Literature*, Vol.41, No:2, pp.478-539.

Ripley, B.D. (1994); "Neural Networks: A Review from Statistical Perspective: Comment", **Statistical Science**, V. 9, No. 1, pp. 45-48.

Roberts, L. (1989); "Are Neural Nets Like the Human Brain", **Science New Series**, Vol. 243, No. 4890, (pp. 481-482).

Rodriguez, J.V.P., Tora, S. and Felix, J.A. (2004); "STAR and ANN Models: Forecasting Performance on the Spanish 'Ibex 35' Stock Index", **Journal of Empirical Finance**, Vol.12, 3, pp.490-509.

Rodriguez, R.J. (2004); "Oil Price Shocks: Testinger for Nonlinearity", **Working Paper** No. 115, Universita Delgi Studi Salerno.

Rothman, P. (1998); "Ferecasting Asymmetric Unemployment Rates", **The Review of Economics and Statistics**, Vol. 80, No. 1, pp. 164-168.

Rudebusch, G.d. (1992); "Trends and Random Walks in Macroeconomic Time Series: A Re-examination", **International Economic Review**, Vol. 33, No. 3, pp. 661-680.

Shachmurove, Y. (2002); "Applying Artificial Neural Networks to Business, Economics and Finance", **Penn CARESS Working Paper**, UCLA Department of Economics.

Salam A.M., Salam S., Feridun M. (2006); "Forecasting Inflation in Developing Nations: The Case of Pakistan", **International Research Journal of Finance and Economics**, No. 3, pp. 138-159

Schiffman, W. and Werner, M.J.R. (1994); "**Optimization of the Bacpropogation Algorithm for Training Multilayer Perceptens**", *IEEE Transactions on Neural Networks*, Proceedings of the IEEE Conference on Decision and Control, San Diego, CA (1994).

Schott, James R. (1997); **Matrix Analysis for Statistics**, John Willey and Sons, New York.

Serletis, A. (1992): "The Random Walk in Canadian Output", **The Canadian Journal of Economics**, Vol. 25, Iss. 2, pp. 392-406.

Serper, Ö. (2004); Uygulamalı İstatistik 2, Ezgi Kitabevi, Bursa.

Setino, R. and Thog, J.Y.L. (2004); "An Approach to Generate Rules from Neural Networks for Regression Problems", **European Journal of Operational Research**, 155, pp. 239-250.

Sevüktekin, M. ve Nargeleçekenler, M. (2005); **Zaman Serileri Analizi**, Nobel Yayın Dağıtım, Ankara.

Sevüktekin, M. ve Nargeleçekenler, M. (2006); "İMKB Getiri Modellemesi ve Önraporlaması" Ankara Üniversitesi Siyasi Bilgiler Fakültesi Dergisi, Cilt:61, No:4, s.243-265.

Shadra, R.,Patil R., "Neural Networks as Forecasting Experts: An Empirical Test, **Proc. 1990 IJCNN Meeting, 2 (1990), 491-494**

Stock, J.H. and Watson, M.W. (1986); "Does GNP Have Unit Root", **Economic Letters**, 22, pp. 147-151.

Su, W.C., Chang, H.L. and Chen, Y.S. (2007); "Stock Prices and Dividends in Taiwan's Stock Market: Evidence Based on Time Varying Present Value Model, **Economic Bulletin**, Vol. 7, No. 4, pp. 1-12.

Tal, B. and Nazareth L. (1995); "Artificial Intelligence and Economic Forecasting", **Canadian Business Economics**, Spring, (pp. 69-74).

Terasvirta, T. and Anderson, H.M. (1992); "Characterizing Nonlinearities in Business Cycles Using Smooth Transition Autoregressive Models, **Journal of Applied Econometrics**, Vol. 7, pp. 119-136.

Tibshirani, R. (1994); "[Neural Networks: A Review from Statistical Perspective]: Comment", **Statistical Science**, Vol. 9, No. 1, (February), pp. 48-49.

Tkacz, G. (2001); "Neural Network Forecasting of Canadian GDP Growth",

**International Journal of Forecasting**, 17, pp. 57-69.

Veelentruf, L. (1995); **Analysis and Applications of Artificial Neural Networks**, Prentice Hall, London.

Vogelsang, T.J. and Perron, P. (1998); “Additional Tests for a Unit Root Allowing the Possibility of Breaks in the Trend Function”, **International Economic Review**, 39, pp. 1073-1100.

Wang, J. and Zhang, W. (1997); “**A Dynamic Backpropagation Algorithm with Application to Gain Scheduled Aircraft Flight Control System Design**”, 1997 IASTED International Conference on Intelligent Information Systems (IIS '97) pp.(133-137)

Warner, B. and Misra, M. (1996); “Understanding Neural Networks as Statistical Tools”, **The American Statistician**, Vol. 50, No. 4, (pp. 284-293).

Wendemuth, A. (2006); “Neural Networks Revisited: A Statistical View on Optimization and Generalization”, <http://www.iesk.et.uni.magdeburg.de>, (Erişim Tarihi: 03.02.2007).

Wiggins, V.L. and Looper, L.T. (1991); “Neural Networks: A Primer”, **Air Force Systems Command Brook**, Air Force Base, Texas, Armstrong Laboratory.

Williamowski, B.M and Chen, Y. (1999); “**Efficient Algorithm for Training Neural Networks with one Hidden Layer**”, Proc. of the International Joint Conference on Neural Networks (IJCNN'99), Washington, D.C., USA, July 10-16, 1999, pp 1725-1728.

Williamowski, B.M et al. (2001); **An Algorithm for Fast Convergence in Training Neural Network**, Neural Networks, 2001. Proceedings. IJCNN '01. International Joint Conference on Neural Networks, Washington DC (2001).

Yao, J.T. and Tan, C.L. (2002); “Neural Networks for Technical Forecasting of Foreign Exchange Rates”, **Neural Networks in Business, Techniques and Applications**, Ed. by S. Kate, Idea Group, Hershey, pp. 189-204.



Yu, X., Efe, M.O. and Kaynak, O. (2002); “A General Backpropagation Algorithm for Feedforward Neural Networks Learning”, **IEEE Transactions on Neural Networks**, Vol. 13, No. 1, January, pp. 251-254.

Yurtođlu, H. (2005); “Yapay Sinir Ağları Metodolojisi ile Öngörü Modellemesi: Bazı Makroekonomik Deđişkenler için Türkiye Örneđi”, **DPT Uzmanlık Tezi**, Yayın No. DPT 2683, Ankara.

Zhang, G.P. (2001); “An Investigation of Neural Networks for Linear Time Series Forecasting”, **Computers and Operations Research**, 28, pp 1183-1202.

Zhang, G.P. (2004); “Business Forecasting with Artificial Neural Networks: An Overview”, **Neural Network in Business Forecasting**, Ed. by G.P. Zhang, Idea Group, Hershey.

Zhang, G.P., Patumo, B.E and Hu, M.Y. (1998); “Forecasting with Artificial Neural Networks, The State of the Art”, **International Journal of Forecasting**, 14, pp. 35-62.

Zhang, G.P. and Qi, M. (2005); “Neural Network Forecasting for Seasonal and Trend Time Series”, **European Journal of Operational Research**, 160, pp. 501-514.

## ÖZGEÇMİŞ

**Doğum Yeri ve Yılı** : Soma / MANİSA 1974

<b>Öğr.Gördüğü Kurumlar</b>	<b>Başlama Yılı</b>	<b>Bitirme Yılı</b>	<b>Kurum Adı</b>
Lise	1988	1991	Soma Linyit Lisesi
Lisans	1995	1999	Uludağ Üniv. İİBF İktisat Böl.
Yüksek Lisans	1999	2002	Uludağ Üniv. Sos. Bil. Enst. Ekonometri
Doktora	2002	2007	Uludağ Üniv. Sos. Bil. Enst. Ekonometri

**Medeni Durum** : Evli

**Bildiği Yabancı Diller ve Düzeyi** : İngilizce İyi

<b>Çalıştığı Kurum (lar)</b>	<b>Başlama ve Ayrılma Tarihleri</b>	<b>Çalışılan Kurumun Adı</b>
1.	2001-	Uludağ Üniv. İİBF Ekonometri Böl.

**Yurtdışı Görevleri** : ---

**Kullandığı Burslar** : ---

**Aldığı Ödüller** : ---

**Üye Olduğu Bilimsel ve Meslekî Toplantılar** :---

**Editör veya Yayın Kurulu Üyelikleri** : ---

**Yurt İçi ve Yurt Dışında katıldığı Projeler** : ---

**Katıldığı Yurt İçi ve Yurt Dışı Bilimsel Toplantılar:**

-Sosyal Hizmetler Sempozyumu: 2003, "Yoksulluk ve Sosyal Hizmetler", Hacettepe Üniversitesi - Akdeniz Üniversitesi, 9-11 Ekim 2003, Antalya.

**Sunulan Bildiri:** Yoksulluğun İntihar Üzerine Etkilerinin Ekonometrik Analizi

-7. Avrupa Kriminoloji Kongresi 26-29 Eylül, Bologna.

**Sunulan Bildiri:** VIOLENT CRIMES Trends of Violence in Bursa/ TURKEY 1994–2007:

A Time Series Perspective

**Yayımlanan Çalışmalar** :

"Poverty Correlates of Suicide Rates In Turkey: A Time-Series Analysis" **International Journal of Criminal Justice**

**Diğer:**---

10.12.2007

Özer ARABACI

