

ÖZET

Doktora Tezi

TAHMİN UYGULAMALARINDA PERFORMANS GELİŞTİRMEK İÇİN KULLANILAN NORMALİZASYON METOTLARINA YENİ BİR YAKLAŞIM

Bayram AKDEMİR

Selçuk Üniversitesi

Fen Bilimleri Enstitüsü

Elektrik-Elektronik Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Salih GÜNEŞ

2009

Jüri:

Prof. Dr. Sadık KARA

Prof. Dr. Mehmet BAYRAK

Prof. Dr. Saadetdin HERDEM

Prof. Dr. Ahmet ARSLAN

Doç. Dr. Salih GÜNEŞ

Bu çalışmada tahmin ve sınıflama problemlerinde kullanılmak üzere yeni bir normalizasyon metodu geliştirilmiş ve önerilmiştir. Önerilen metot özellikle yapay zeka uygulamaları hedef alınarak geliştirilmiş ve geliştirilen metodun testinde günümüz literatüründe hakkında en fazla makale ve dökümantasyon olan ve yaygın

şekilde kullanılan iki yapay zeka tekniği olan Adaptif ağ tabanlı bulanık çıkarım sistemi (ANFIS) ve Yapay sinir ağları (YSA) kullanılmıştır.

Geliştirilen metot yalnızca bir normalizasyon metodu olmanın ötesinde veri normalizasyonuna yeni bir bakış açısı getirmektedir. Satır tabanlı olan normalizasyon modelinin performansının kayda değer şekilde klasik modellerle yarışması önerilen modelin de literatürde kullanılan klasik normalizasyon metotlarına yeni bir alternatif sunabileceğinin işaretidir. Satır tabanlı normalizasyon yapılırken, bir veri seti içerisindeki özellik vektörleri farklı birimlerde olabileceği göz önüne alınarak ilk aşamada bu özellikler birimsizleştirilmiş daha sonra normalize edilmiştir. Bu sayede birim problemi ortadan kalkmış ve değerler salt birer sayı haline dönmüştür.

Önerilen metodun test edilmesinde yalnızca tahmin veri setleri kullanılmamış, sınıflama veri setlerindeki olası sonuçları görmek için de literatürde yaygın olan sınıflama veri kümeleri ile denenmiştir. Sınıflama veri setlerindeki başarısı da modelin bu tür veri yapılarında da kullanılabileceğini göstermektedir.

Geliştirilen metodun performansının klasik metotlarla olan farklılığı literatürde kullanılan matematiksel ve istatistiksel değerlendirme kriterleri ile kıyas edilmiş ve değerlendirilmiştir. Ortalama devinim, ortalama karesel hata ve iki kat çapraz geçerlilik testi uygulanmıştır. Çalışmada önerilen metodun, ANFIS ve YSA performanslarını arttırması ve sonuçlarının literatürde kullanılan klasik modellerle yarışabilmesi bu metodun başarısı olarak sayılabilir.

Anahtar Kelimeler: Normalizasyon, Satır tabanlı normalizasyon, Öklid uzaklığı, Tahmin verileri, Yapay zeka, Yapay sinir ağları, Adaptif ağ tabanlı bulanık çıkarım sistemi

ABSTRACT

PhD. Thesis

A NEW APPROACH TO NORMALIZATION METHODS FOR IMPROVING PERFORMANCE ON THE PREDICTION APPLICATIONS

Bayram AKDEMİR

Selcuk University

Graduate School of Natural and Applied Sciences

Department of Electrical-Electronics Engineering

Supervisor: Assoc.Prof.Dr. Salih GÜNEŞ

2009

Jury:

Prof. Dr. Sadık KARA

Prof. Dr. Mehmet BAYRAK

Prof. Dr. Saadetdin HERDEM

Prof. Dr. Ahmet ARSLAN

Assoc. Prof. Dr. Salih GÜNEŞ

In this study, a novel normalization method has been developed and proposed to be used in prediction and classification problems. The proposed method has been developed aiming particularly artificial intelligence applications. During testing of the developed method, two widely used artificial intelligence techniques namely

adaptive network based fuzzy inference system (ANFIS) and artificial neural networks (ANN), which are perhaps the two most widely reported and documented subjects in literature, have been used.

The developed method is not only a normalization method but also brings the data normalization a different point of view. Significant competing of the performance of line-based normalization model with other classical models gives a hint that the proposed model can offer a new alternative to classical normalization methods used in literature. During the implementation of line-based normalization, by taking into account that the feature vectors can be in different dimensions in a dataset, in the first stage the features have been made dimensionless and normalized afterwards. Therefore, the dimension problem has been avoided and each value has been transformed into sole numbers.

During the test of proposed method, besides the prediction dataset, classification dataset that are widely available in literature have been tried to observe possible results for classification dataset. The success of the model in classification dataset demonstrates that it can be used in this kind of data structures.

The performance of the developed method have been evaluated and compared with the classical methods by employing mathematical and statistical evaluation criteria used in literature. Mean-variance, mean-square-error and two-fold cross-validation tests have been performed. Providing an increase on the performance of ANFIS and ANN and competing of the results with other classical models used in literature can be regarded as the success of this method.

Keywords – Normalization, line-based normalization, Euclid distance, Prediction dataset, Artificial intelligence, Artificial neural networks, Adaptive network based fuzzy inference system

TEŞEKKÜR

Çalışmama verdiği destek ve gösterdiği anlayıştan dolayı danışmanım sayın Doç.Dr. Salih GÜNEŞ'e, tez süresince belirttikleri görüşler ve önerilerle tezin yönlenmesinde yardımcı olan sayın Prof.Dr. Mehmet BAYRAK ve sayın Prof.Dr. Ahmet ARSLAN'a, Bölüm Başkanımız Prof. Dr. Saadetdin HERDEM'e, tıbbi veriler konusunda sağladığı desteklerden dolayı sayın Prof. Dr. Sadık KARA ve Prof. Dr. Bülent ORAN'a, Selçuk Üniversitesi Mühendislik-Mimarlık Fakültesi Elk.-Elt. Müh. Bölümü çalışma arkadaşlarıma ve özellikle manevi desteğini hiç bir zaman esirgemeyen her zaman ve her konuda hep yanımda olan aileme desteklerinden dolayı teşekkür ederim.

Bayram AKDEMİR

2009, Konya

İÇİNDEKİLER

ÖZET	i
ABSTRACT	iii
TEŞEKKÜR	v
İÇİNDEKİLER	vi
SİMGELER VE KISALTMALAR.....	ix
ŞEKİLLER LİSTESİ.....	x
TABLolar LİSTESİ.....	xiii
1. GİRİŞ	16
1.1 Literatür Taraması	18
1.2 Çalışmanın Temel Amacı ve Literatüre Katkısı	20
1.3 Çalışmanın Kapsamı ve Organizasyon	21
2. NORMALİZASYON METOTLARI.....	23
2.1 Literatürde Kullanılan Normalizasyon Metotları.....	24
2.1.1 Minimum-maksimum metodu.....	24
2.1.2 Ondalık ölçekleme metodu	26
2.1.3 Z-skör metodu	27
2.2 Önerilen Normalizasyon Metodu.....	29
2.2.1 Birimsizleştirilmiş satır tabanlı normalizasyon metodu	29
2.2.2 Veri setinin birimsizleştirilmesi	30
2.2.3 Önerilen normalizasyon metodunun matematiksel tanımı	31
3. ÖNERİLEN NORMALİZASYON METODUNUN UYGULANMASI.....	34
3.1 Satır Normalizasyonunun Veri Üzerindeki Etkisi	37
3.2 Geri Çözümleme.....	38
3.2.1 Minimum-Maksimum metodu	38
3.2.2 Ondalık Ölçekleme metodu.....	39
3.2.3 Z-skör metodu	39
3.2.4 LBNM metodu	40
4. YAPAY ZEKA.....	41
4.1 Yapay Sinir Ağları.....	41
4.1.1 Yapay sinir ağları tarihi.....	42
4.1.2 Yapay sinir ağlarının yapısı	44
4.1.3 Yapay sinir ağlarının yapısı	45

4.1.4	Geriye yayılım algoritması.....	49
4.1.5	Levenberg-Marquardt öğrenme metodu.....	50
4.2	Adaptif Ağ Tabanlı Bulanık Çıkarım Sistemi (ANFIS)	52
4.2.1	ANFIS'in Yapısı	53
5.	VERİ SETLERİ.....	63
5.1	Tahmin verileri	65
5.1.1	Beton verisi	65
5.1.2	Kalp M-Mode aort ve sistol verileri.....	66
5.1.3	Diş anterior temporal ve üst dudak verileri	67
5.2	Sınıflama verileri	68
5.2.1	SPECT (Single Photon Emission Computed Tomography) görüntüleri kalp hastalığı teşhisi verileri	68
5.2.2	Kalp hastalığı verisi	68
5.2.3	VEP verisi	69
5.2.4	Doppler sinyali ile damar sertliği veri kümesi	70
5.2.5	Hepatit verisi	70
6.	PERFORMANS KRİTERLERİ	72
6.1	Tahmin verilerinde kullanılan değerlendirme kriterleri.....	72
6.1.1	Ortalama Devinim.....	72
6.1.2	Ortalama Karesel Hata	73
6.1.3	R^2 değeri.....	73
6.1.4	2-Kat Çaprazlama	74
6.2	Sınıflama verilerinde kullanılan değerlendirme kriterleri.....	75
6.2.1	Sınıflama doğruluğu, duyarlılık ve seçicilik değerleri	75
6.2.2	ROC eğrisi altında kalan alan	76
6.3	Tahmin Verilerinin Değerlendirilmesi.....	79
6.3.1	Kalp aort verisi.....	80
6.3.2	Beton Verisi	92
6.3.3	Kalp Sistol Verisi	106
6.3.4	Diş anterior temporal verisi.....	117
6.3.5	Diş üst dudak verisi.....	130
6.4	Sınıflama Verilerinin Değerlendirilmesi	141
7.	ÖZ SONUÇLAR VE ÖNERİLER.....	144
7.1	Özsonuçlar	144
7.1.1	Aort verilerinin değerlendirilmesi	145
7.1.2	Beton verilerinin değerlendirilmesi.....	146
7.1.3	Sistol verilerinin değerlendirilmesi	146

7.1.4	Diş anteior temporal veri setinin değeriendirilmesi	147
7.1.5	Diş üst dudak veri setinin değeriendirilmesi	148
7.2	Sınıflama verilerinin değeriendirilmesi	148
7.3	Öneriler.....	149
8.	KAYNAKLAR:	151

SİMGELER VE KISALTMALAR

AD	: Average Deviation (Ortalama Devinim)
ANN	: Artificial Neural Network (Yapay Sinir Ağları)
ANFIS	: Adaptive Network Based Fuzzy Inference System (Adaptif Ağ Tabanlı Bulanık Çıkarım Sistemi)
Cor	: Korelasyon
EBNM	: Euclidean Based Normalization Method (Öklid Tabanlı normalizasyon)
LBNM	: Line Based Normalization Method (Satır Tabanlı Normalizasyon Metodu)
Min-Mak	: Minimum-Maksimum
MLP	: Multi Layer Perceptron (Çok Katmanlı Nöron)
MSE	: Mean Squared Error (Ortalama Karesel Hata)
p	: İstatistiksel olasılık
R^2	: İstatistiksel değerlendirme kriteri
UCI	: University of California, Irvine
Var	: Varyans
YSA	: Yapay Sinir Ağları

ŞEKİLLER LİSTESİ

Şekil 2.1 Minimum-Maksimum metodunun kullanılış şeklini gösteren diyagram.....	25
Şekil 2.2 Ondalık ölçeklemenin nasıl yapıldığını anlatan akış diyagramı.....	26
Şekil 2.3 Z-skor hesaplanmasını gösteren akış diyagramı	28
Şekil 2.4 Önerilen ve klasik metodun veri setine uygulanışının gösterimi	30
Şekil 2.5 Önerilen metodun satır ve sütun boyunca nasıl kullanıldığını gösteren nümerik bir örnek (a) sütun boyunca uygulanan normalizasyon (b) satır boyunca uygulanan normalizasyon	33
Şekil 4.1 Biyolojik sinir hücresinin yapısı	44
Şekil 4.2 Sentetik nöronun modelini gösteren resim (MATLAB 7.0 Help dosyasından alınmıştır)	46
Şekil 4.3 YSA transfer fonksiyonlarına ait yaygın transfer fonksiyonları (a) logaritmik sigmoid transfer fonksiyonu (b) tanjant sigmoid transfer fonksiyonu (c) basamak transfer fonksiyonu (d) liner transfer fonksiyonu (e) saturasyona uğratılmış liner transfer fonksiyonu (f) pozitif transfer fonksiyonu	47
Şekil 4.4 MLP yapısının çalışmasını gösteren bağlantı diyagramı (MATLAB 7.0 Help dökümanından alınmıştır)	48
Şekil 4.5 Bir YSA ağında geriye yayılım algoritmasının uygulanması gösteren akış şeması	49
Şekil 4.6 İki girişli ANFIS yapısını gösteren temsili resim (a) üyelik fonksiyonlarının (b) iki girişli sugeno yapılı-ANFIS modelinin gösterimi.....	54
Şekil 4.7 Katman 1'e ait iki girişli Sugeno yapılı ANFIS'in hesaplanmasını gösterimi	55
Şekil 4.8 ANFIS yapısı içerisindeki adaptif düğümleri gösteren temsili şekil	59
Şekil 4.9 Dilsel ifadelere ait parametrelerin değişimini gösteren şekil (a) çan eğrisi üyelik fonksiyonuna ait parametreleri gösteren eşitlik (b) parametre a 'nin üyelik fonksiyonuna etkisi (c) parametre b 'nin üyelik fonksiyonuna etkisi (d) parametre c 'nin üyelik fonksiyonuna etkisi	60
Şekil 4.10 Giriş uzayındaki Grid bölümlendirmesini anlatan resim.....	61
Şekil 6.1 İki fold çaprazlama geçerlilik testinin blok diagramı.....	74
Şekil 6.2 Farklı AUC değerlerine sahip ROC eğrileri.....	77
Şekil 6.3 Ham kalp aort verisine ait (a) YSA eğitme hata grafiği (b) YSA test hata grafiği.....	81
Şekil 6.4 LBNM metoduna ait kalp aort verilerinin (a) YSA eğitme hata grafiği (b) YSA test hata grafiği	82
Şekil 6.5 Min.-Mak. Metoduna ait kalp aort verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği	83
Şekil 6.6 Ondalık ölçeklemeye ait kalp aort verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği	84
Şekil 6.7 Z-skor metoduna ait kalp aort verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği	85
Şekil 6.8 Ham kalp aort verilerinin (a) ANFIS eğitme hata grafiği (b) ANFIS test hata grafiği.....	87

Şekil 6.9 LBNM metoduna ait kalp aort verilerinin (a) ANFIS eğitme hata grafiği (b) ANFIS test hata grafiği	88
Şekil 6.10 Min.-Mak. metoduna ait kalp aort verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği.....	89
Şekil 6.11 Ondalık ölçeklemeye ait kalp aort verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği.....	90
Şekil 6.12 Z-skör metoduna ait kalp aort verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği	91
Şekil 6.13 Ham beton verilerinin (a) YSA eğitme hata grafiği (b) YSA test hata grafiği.....	93
Şekil 6.14 LBNM metoduna ait beton verilerinin (a) YSA eğitme hata grafiği (b) YSA test hata grafiği	94
Şekil 6.15 Min.-Mak. metoduna ait beton verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği	95
Şekil 6.16 Ondalık ölçeklemeye ait beton verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği	97
Şekil 6.17 Z-skör metoduna ait beton verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği	98
Şekil 6.18 Ham beton verilerinin (a) ANFIS eğitme hata grafiği (b) ANFIS test hata grafiği.....	100
Şekil 6.19 LBNM metoduna ait beton verilerinin (a) ANFIS eğitme hata grafiği (b) ANFIS test hata grafiği	101
Şekil 6.20 Min.-Mak. Metoduna ait beton verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği	102
Şekil 6.21 Ondalık ölçekleme metoduna ait beton verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği	103
Şekil 6.22 Z-skör metoduna ait beton verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği	104
Şekil 6.23 Ham kalp sistol verilerinin (a) YSA eğitme hata grafiği (b) YSA test hata grafiği.....	106
Şekil 6.24 LBNM metoduna ait kalp sistol verilerinin (a) YSA eğitme hata grafiği (b) YSA test hata grafiği	107
Şekil 6.25 Min.-Mak. Metoduna ait kalp sistol verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği	108
Şekil 6.26 Ondalık ölçekleme metoduna ait kalp sistol verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği.....	109
Şekil 6.27 Z-skör metoduna ait kalp sistol verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği	110
Şekil 6.28 Ham kalp sistol verilerinin (a) ANFIS eğitme hata grafiği (b) ANFIS test hata grafiği ...	112
Şekil 6.29 LBNM metoduna ait kalp sistol verilerinin (a) ANFIS eğitme hata grafiği (b) ANFIS test hata grafiği.....	113
Şekil 6.30 Min.-Mak. Metoduna ait kalp sistol verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği.....	114
Şekil 6.31 Ondalık ölçekleme metoduna ait kalp sistol verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği	115
Şekil 6.32 Z-skör metoduna ait kalp sistol verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği	116

Şekil 6.33 Ham dış anterior temporal verilerinin (a) YSA eğitme hata grafiği (b) YSA test hata grafiği	118
Şekil 6.34 LBNM metoduna ait dış anterior temporal verilerinin (a) YSA eğitme hata grafiği (b) YSA test hata grafiği	119
Şekil 6.35 Min.-Mak. metoduna ait dış anterior temporal verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği.....	120
Şekil 6.36 Ondalık ölçeklemeye ait dış anterior temporal verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği.....	121
Şekil 6.37 Z-skor metoduna ait dış anterior temporal verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği.....	122
Şekil 6.38 Ham dış anterior temporal verilerinin (a) ANFIS eğitme hata grafiği (b) ANFIS test hata grafiği	124
Şekil 6.39 LBNM metoduna ait dış anterior temporal verilerinin (a) ANFIS eğitme hata grafiği (b) ANFIS test hata grafiği.....	125
Şekil 6.40 Min.-Mak. Metoduna ait dış anterior temporal verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği.....	126
Şekil 6.41 Ondalık ölçeklemeye ait dış anterior temporal verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği.....	127
Şekil 6.42 Z-skor metoduna ait dış anterior temporal verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği	128
Şekil 6.43 Ham dış üst dudak verilerinin (a) YSA eğitme hata grafiği (b) YSA test hata grafiği.....	130
Şekil 6.44 LBNM metoduna ait dış üst dudak verilerinin (a) YSA eğitme hata grafiği (b) YSA test hata grafiği.....	131
Şekil 6.45 Min.-Mak. Metoduna ait dış üst dudak verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği.....	132
Şekil 6.46 Ondalık ölçekleme metoduna ait dış üst dudak verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği.....	133
Şekil 6.47 Z-skor metoduna ait dış üst dudak verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği	134
Şekil 6.48 Ham dış üst dudak verilerinin (a) ANFIS eğitme hata grafiği (b) ANFIS test hata grafiği.....	136
Şekil 6.49 LBNM metoduna ait dış üst dudak verilerinin (a) ANFIS eğitme hata grafiği (b) ANFIS test hata grafiği	137
Şekil 6.50 Min.-Mak. Metoduna ait dış üst dudak verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği	138
Şekil 6.51 Ondalık ölçekleme metoduna ait dış üst dudak verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği.....	139
Şekil 6.52 Z-skor metoduna ait dış üst dudak verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği.....	140

TABLÖLER LİSTESİ

Tablo 3.1 Tek değişkenli Diş Anterior Temporal verisinin ham, minimum-maksimum ve satır tabanlı normalizasyonlarına ait sonuçları aşağıdaki gibidir.....	34
Tablo 3.2 Tek değişkenli Diş Anterior Temporal verisinin normalizasyon metotlarına göre giriş ile çıkış arasındaki korelasyonun değişimini gösterir tablo.	36
Tablo 3.3 Tek değişkenli Diş Anterior Temporal verisinin normalizasyon metotlarına göre giriş ile çıkış arasındaki <i>p-olasılık</i> değişimini gösterir tablo.	37
Tablo 3.4 Korelasyon değerlerini gösterir tablo.....	37
Tablo 3.5 Veri setlerine ait <i>p-olasılık</i> değerlerini gösterir tablo.....	38
Tablo 4.1 Adaptif parametrelerin nasıl optimize edildiğini gösteren tablo	61
Tablo 5.1 Beton sıkışma gerilme verisine ait özellikleri gösteren tablo.....	65
Tablo 5.2 SPECT veri setine ait özelliklerin istatistiki değerleri	68
Tablo 5.3 Kalp hastalığı veri kümesinde bulunan özelliklerin istatistiki değerleri	69
Tablo 5.4 Hepatit veri setinde bulunan özelliklerin istatistiki değerleri.....	71
Tablo 6.1 Karışıklık (confusion) matrisi	75
Tablo 6.2 Ham kalp aort verisine ait YSA performans sonuçları.....	81
Tablo 6.3 LBNM metoduna ait kalp aort verilerinin YSA performans sonuçları	82
Tablo 6.4 Min.-Mak. Metodunun kalp aort verilerinin YSA performans sonuçları.....	83
Tablo 6.5 Ondalık ölçeklemeye ait kalp aort verilerinin YSA performans sonuçları	84
Tablo 6.6 Z-skor metoduna ait kalp aort verilerinin YSA eğitme ve test performans sonuçları.....	85
Tablo 6.7 Tüm normalizasyon metotlarına ait kalp aort verilerinin YSA eğitme performans sonuçları	86
Tablo 6.8 Tüm normalizasyon metotlarına ait kalp aort verilerinin YSA test performans sonuçları...	86
Tablo 6.9 Ham kalp aort ham verilerinin ANFIS performans sonuçları	87
Tablo 6.10 LBNM metoduna ait kalp aort verilerinin ANFIS performans sonuçları.....	88
Tablo 6.11 Min.-Mak. Metoduna ait kalp aort verilerinin ANFIS performans sonuçları.....	89
Tablo 6.12 Ondalık ölçekleme metoda ait kalp aort verilerinin ANFIS performans sonuçları	90
Tablo 6.13 Z-skor metoduna ait kalp aort verilerinin ANFIS performans sonuçları.....	91
Tablo 6.14 Tüm normalizasyon metotlarına ait kalp aort verilerinin ANFIS eğitme performans sonuçları	92
Tablo 6.15 Tüm normalizasyon metotlarına ait kalp aort verisi ANFIS test performans sonuçları	92
Tablo 6.16 Ham beton verilerinin YSA performans sonuçları.....	94
Tablo 6.17 LBNM metoduna ait beton verilerinin YSA performans sonuçları	95
Tablo 6.18 Min.-Mak. Metoduna ait beton verilerinin YSA performans sonuçları	96
Tablo 6.19 Ondalık ölçekleme metoduna ait beton verilerinin YSA performans sonuçları.....	97

Tablo 6.20 Z-skör metoduna ait beton verilerinin performans sonuçları	98
Tablo 6.21 Tüm normalizasyon metotlarına ait beton verisi YSA eğitim performans sonuçları.....	99
Tablo 6.22 Tüm normalizasyon metotlarına ait YSA test performans sonuçları	99
Tablo 6.23 Ham beton verilerinin ANFIS performans sonuçları	100
Tablo 6.24 LBNM metoduna ait beton verilerinin ANFIS performans sonuçları	101
Tablo 6.25 Min.-Mak. Metoduna ait beton verilerinin ANFIS performans sonuçları.....	102
Tablo 6.26 Ondalık ölçekleme metoduna ait beton verilerinin ANFIS performans sonuçları	103
Tablo 6.27 Z-skör metoduna ait beton verilerinin ANFIS performans sonuçları.....	105
Tablo 6.28 Tüm normalizasyon metotlarına ait beton verilerinin ANFIS eğitim performans sonuçları	105
Tablo 6.29 Tüm normalizasyon metotlarına ait beton verilerinin ANFIS test performans sonuçları.	105
Tablo 6.30 Ham kalp sistol verilerinin YSA performans sonuçları	107
Tablo 6.31 LBNM metoduna ait kalp sistol verilerinin YSA performans sonuçları	108
Tablo 6.32 Min.-Mak. Metoduna ait kalp sistol verilerinin YSA performans sonuçları.....	109
Tablo 6.33 Ondalık ölçekleme metoduna ait kalp sistol verilerinin YSA performans sonuçları	110
Tablo 6.34 Z-skör metoduna ait kalp sistol verilerinin YSA performans sonuçları.....	110
Tablo 6.35 Tüm normalizasyon metotlarına ait kalp sistol verilerinin YSA eğitim performans sonuçları	111
Tablo 6.36 Tüm normalizasyon metotlarına ait kalp sistol verilerinin YSA test performans sonuçları	111
Tablo 6.37 Ham kalp sistol verilerinin ANFIS performans sonuçları.....	112
Tablo 6.38 LBNM metoduna ait kalp sistol verilerinin ANFIS performans sonuçları	113
Tablo 6.39 Min.-Mak. Metoduna ait kalp sistol verilerinin ANFIS performans sonuçları	114
Tablo 6.40 Ondalık ölçekleme metoduna ait kalp sistol verilerinin ANFIS performans sonuçları	115
Tablo 6.41 Z-skör metoduna ait kalp sistol verilerinin ANFIS performans sonuçları	116
Tablo 6.42 Tüm normalizasyon metotlarına ait kalp sistol verilerinin ANFIS eğitim performans sonuçları	117
Tablo 6.43 Tüm normalizasyon metotlarına ait kalp sistol verilerinin ANFIS test performans sonuçları	117
Tablo 6.44 Ham dış anterior temporal verilerinin YSA performans sonuçları.....	119
Tablo 6.45 LBNM metoduna ait dış anterior temporal verilerinin YSA performans sonuçları	120
Tablo 6.46 Dış anterior temporal verisi Min.-Mak. metoduna ait YSA performans sonuçları	121
Tablo 6.47 Ondalık ölçekleme metoduna ait dış anterior temporal verilerinin YSA performans sonuçları	122
Tablo 6.48 Z-skör metoduna ait dış anterior temporal verilerinin YSA performans sonuçları	123
Tablo 6.49 Tüm normalizasyon metotlarına ait dış anterior temporal verilerinin YSA eğitim performans sonuçları	123
Tablo 6.50 Tüm normalizasyon metotlarına ait dış anterior temporal verilerinin YSA test performans sonuçları	123
Tablo 6.51 Ham dış anterior temporal verilerinin ANFIS performans sonuçları	125

Tablo 6.52 LBNM metoduna ait dış anterior temporal verilerinin ANFIS performans sonuçları.....	126
Tablo 6.53 Min.-Mak. metoduna ait dış anterior temporal verilerinin ANFIS performans sonuçları	127
Tablo 6.54 Ondalık ölçekleme metoduna ait dış anterior temporal verilerinin ANFIS performans sonuçları	128
Tablo 6.55 Z-skör metoduna ait dış anterior temporal verilerinin ANFIS performans sonuçları.....	129
Tablo 6.56 Tüm normalizasyon metotlarına ait dış anterior temporal verilerinin ANFIS eğitme performans sonuçları	129
Tablo 6.57 Tüm normalizasyon metotlarına ait dış anterior temporal verilerinin ANFIS test performans sonuçları	129
Tablo 6.58 Ham dış üst dudak verilerinin YSA performans sonuçları.....	131
Tablo 6.59 LBNM metoduna ait dış üst dudak verilerinin YSA performans sonuçları	132
Tablo 6.60 Min.-Mak. Metoduna ait dış üst dudak verilerinin YSA performans sonuçları	132
Tablo 6.61 Ondalık ölçekleme metoduna ait dış üst dudak verilerinin YSA performans sonuçları...	133
Tablo 6.62 Z-skör metoduna ait dış üst dudak verilerinin YSA performans sonuçları	134
Tablo 6.63 Tüm normalizasyon metotlarına ait dış üst dudak verilerinin YSA eğitme performansı sonuçları	135
Tablo 6.64 Tüm normalizasyon metotlarına ait dış üst dudak verilerinin YSA test performansı sonuçları	135
Tablo 6.65 Ham dış üst dudak verilerinin ANFIS performans sonuçları	136
Tablo 6.66 LBNM metoduna ait dış üst dudak verilerinin ANFIS performans sonuçları.....	137
Tablo 6.67 Min.-Mak. Metoduna ait dış üst dudak verilerinin ANFIS performans sonuçları.....	138
Tablo 6.68 Ondalık ölçekleme metoduna ait dış üst dudak verilerinin ANFIS performans sonuçları	139
Tablo 6.69 Z-skör metoduna ait dış üst dudak verilerinin ANFIS performans sonuçları.....	140
Tablo 6.70 Tüm normalizasyon metotlarına ait dış üst dudak verilerinin ANFIS eğitme performans sonuçları	140
Tablo 6.71 Tüm normalizasyon metotlarına ait dış üst dudak verilerinin ANFIS test performans sonuçları	141
Tablo 6.72 SPECT veri setine ait YSA ve ANFIS performans sonuçları	141
Tablo 6.73 Kalp hastalığı verisine ait YSA ve ANFIS sınıflama performansı sonuçları	142
Tablo 6.74 Doppler veri setine ait YSA ve ANFIS performans sonuçları	143
Tablo 6.75 VEP veri setine ait YSA ve ANFIS performans sonuçları.....	143
Tablo 6.76 Hepatit veri setine ait YSA ve ANFIS performans sonuçları.....	143

1. GİRİŞ

Hızla gelişen teknolojinin beklide beraberinde getirdiği en büyük sorunlardan biri daha çok veri işleme gereksinimidir. Artık ürünlerin daha az enerji harcamaları, daha konforlu hale gelmeleri ve daha fazla işlevsel olmalarının altında yatan etken bu ürünlerin yapım, kontrol ve kullanım alanlarındaki daha fazla verinin yorumlanmasında yatmaktadır. Tıbbi cihazlardan havacılık sektörüne, tarım sektöründen kimya sektörüne gibi her alanda insan yorumuna gerek kalmadan prosedürlerin uygulanması, yorumlanması ve uygulanması gerekmektedir. geniş çeşitlilik ve yapısal farklar verilerin işlenmesi sırasında zorlukları da beraberinde getirir. Veri seti içerisindeki semboller ve sayısal olmayan büyüklükler gibi nedenler veri işlemek için aşılması gereken ilk engellerdir. Özellikle yapay zeka uygulamaları doğrusal olamayn (non-linear) problemlerin çözümünde oldukça gözde yöntemler olmasına rağmen veri yapısı sonucu direk etkileyen faktörlerden biridir. veri yapısının yapay zeka uygulamasına katılabilabilmesi için normalize edilmesi gerekmektedir. normalizasyon sonucunda veri seti değerleri yapay zeka (özellikle yapay sinir ağları) transfer fonksiyonuna uygun hale gelir.

Yapay zekâ pek çok disiplinlerde uygulama alanı bulmuştur. Özellikle nonlinear problem çözümlerdeki kayda değer performansı ve çözümü kolaylaştırması yapay zekâyı öne çıkaran hususlardandır. Deterministik modeller problemin çözümünde kesin sonuç önermekle birlikte, şartların değişmesi bulunan çözümü demode hale getirebilir. Oysa sezgisel (heuristik) modeller şartların değişmesi karşısında da uygun bir veri üretilip sistemin çalışmasına ya da toparlanarak devamında ortama uyum sağlamasına ön ayak olabilirler.

Yapay sinir ağları (YSA) ve Adaptif Ağ tabanlı bulanık çıkarım sistemi (Adaptive Neural Fuzzy Inference System – ANFIS), yapay zeka modelleri arasında oldukça popüler modellerdir. Yapay sinir ağları, insan beynindeki sinirlerin

modellenmesi ile oluşturulmuş paralel veri işleyebilen ve pek çok literatürde uygulama alanı bulmuş bir modeldir.

ANFIS, bulanık mantığın ileri bir modelidir. İlk defa Lotfi Zadeh tarafından tarif edilen bulanık mantık'dan (Fuzzy) bu yana, dilsel ifadeler bulanık mantığın çalışmasının temel taşı olsa da literatürde sürekli oluşturulan eklentiler sayesinde ANFIS oldukça güçlü bir çıkarım sistemi haline gelmiştir.

YSA ve ANFIS ilk evrelerini başarı ile tamamlamış ve pek çok disiplinde uygulama alanı bulmasından dolayıyla da güçlü bir yapay zeka metodu olarak kabul edilmiştir.

Yapay zeka uygulamasından önce veri üzerinde bazen bir ön işleme gerekli olabilir. Örneğin bir not sistemi yapay zekaya uygulanmak isteniyorsa öncelikle notların harf ile ifade ettiği dilsel ifadelerden ya da renklerle ifade edilen bir problem varsa bunun sayısal ifadelere çevrilmesi gerekir. Sayısal ifadelere döndürmenin şartı nedir ya da hangi aralıktaki verileri işlemek sonuca daha doğru yaklaşılması anlamına gelir sorusunun cevabı veri ön işleminde gizlidir. Normalizasyon ise veri setinin yapay zekaya uygulanmaya hazır hale getirilmesi için yapılan ön-işlemlerden biridir.

Yapay zeka uygulamalarında, yapay zeka tekniklerinin transfer fonksiyonuna bağlı tanım aralığı da verilerin normalizasyona tabi tutulmasını gerektirebilir. Bununla beraber, büyük değerdeki verilerin daha işlenebilir sayılara indirgenmesi, verilerin ötelenmesi, sıkıştırılması ve aykırı değerlerin indirgenmesi gibi ihtiyaçlar da normalizasyonun yapılması için geçerli birer sebeptir. Örneğin YSA bu bağlamda sigmoid transfer fonksiyonu kullanıldığında oldukça başarılı sonuçlar veren bir model olmasına rağmen sigmoid fonksiyonu çeşitleriyle ilgili olarak 0-1 yada -1 ile +1 arasında YSA'nın daha performanslı çalışmasını sağlamaktadır. Yukarıda bahsedilen sebepler zinciri ve benzerleri verilerin normalize edilmesi ihtiyacını doğurur. Bununla birlikte bazı disiplinlerde belirli normalizasyon türleri daha fazla anlam içerebilir. Özellikle tıbbi ve istatistiksel uygulamalarda z-skorunun oldukça yaygın bir kullanım alanı vardır.

Pek çok uygulamada veya deneysel verilerde değerler süreksiz ve noktasal parçalar halinde olmasına rağmen aranan, gerçekte sürekli bir fonksiyon ya da onun

bir parçasıdır. Hedef, eğer tahmine dayalı bir değerlendirme ise problemin çözümü ve doğru ön işleme ayrı bir önem kazanabilir. Bununla birlikte pek çok mühendislik problemi doğrusal olmayan sonuçları içermenin yanında ikinci dereceden veya daha yüksek çözümleri ihtiva edebilir. Giriş ile çıkış arasında basit bir deterministik ifade bulmak imkansız ya da çok parçalı olabilir.

Mühendislik literatüründe en yaygın normalizasyon tekniği veri setinin her bir özelliğini 0-1 aralığına taşıyan minimum-maksimum normalizasyon metodudur. Gereklilik ya da zorunluluk arasındaki ince çizgi normalizasyonu bir ön işleme prosedürü haline getirmekle kalmayıp aynı zamanda sonucu etkilemesine bağlı olarak yeni modellerin geliştirilmesini de ihtiyaç haline getirmiştir. Ondalık ölçekleme minimum-maksimum normalizasyonundan sonra literatürde kullanılmış bir uygulamadır.

Bu tez çalışmasında, literatürde kullanılan klasik normalizasyon metotlarından farklı olarak Öklid uzaklığının normalizasyon metodu olarak kullanılabileceğın ve bu konuda yapılan uygulamalar anlatılmıştır.

Önerilen metot tahmin amaçlı veri setlerini hedef alınarak geliştirilmiş olmasına rağmen sınıflama verilerinde de denenmiş ve sonuçları çalışmaya eklenmiştir. Elde edilen sonuçlar literatürde kullanılan performans kriterlerine göre kıyas edilmiş ve 2 kat çapraz geçerlilik testi uygulanmıştır. Önerilen metodun performansı hedeflenen veri tiplerinde denenmiş, sonuçlar olumlu bulunmuş ve bu bağlamda yapılan çalışmalar literature olan katkı sağlamıştır.

1.1 Literatür Taraması

Gerek tahmin gerekse sınıflama verilerinin mühendislik uygulamalarında özel bir önemi vardır. Özellikle deneysel verilerden elde edilen değerler bir basamak oluşturmakla birlikte örneğın deprem dayanıklılık testleri gibi verilerde aranan çözümler test edilmeyen kombinasyonlarını da içerebilir. Hedeflenen aygıtı ya da düzeneğı her durum için ayrı ayrı test etmek zaman, enerji ve para kaybına sebep olabilir (otomobil kaza testleri gibi). Bunlara ilaveten giriş parametreleri ile çıkış

arasındaki non-lineerlik deterministik bir çözüm modeli oluşturma zorluğu içerebilir. Bu gibi durumlarda uygun bir yapay zeka tekniği problemin çözümünde bir omurga teşkil edebilir.

Yapay zekanın kullanımı, 1950’li yıllarda ortaya çıkmış ve zamanla doğanın bir çok davranışı yapay zeka tekniklerine esin kaynağı olmuştur. Karıncaların koloni olarak yaşayışını ve yiyecek bulmadaki performanslarından esinlenilen karınca koloni algoritması, insan beynindeki nöronların işleyişini model alan yapay sinir ağları, dilsel ifadelerden esinlenilen ve Lotfi Zadeh’in öncüsü olduğu bulanık mantık, insan genetiğini ve evrim teorisinin modellenmesine çalışılan genetik ve evrim algoritmaları en yaygın yapay zeka modelleri arasındadır (Lee ve Zadeh 1969, Beasley 1993, Yasuhara 1971, Dorigo 1996).

Bunlar arasında özellikle YSA ve bulanık mantığın geliştirilmiş modeli olan ANFIS pek çok mühendislik çözümünde kullanılmış ve kayda değer sonuçlar literatürlerde yayınlanmıştır (Roger Jang 1993). Mamdani’nin ilk dilsel ifadelerinin bir DC motora uygulanmasının ardından bulanık mantık ilgi uyandırmış ve yeni nesil pek çok endüstriyel uygulamada bulanık mantık esin kaynağı olmuştur (Mamdani 1974, Jia 1994, De Gloria 1999).

YSA’nın gerisindeki 50 yıl dikkate alındığında gelişimini tamamlamış bir metot olarak kabul edilebilir. Bu bağlamda YSA oldukça popüler bir yapay zeka modelidir ve literatürde mühendislik, tıp gibi pek çok disiplinde çok sayıda çalışma vardır. Özellikle paralel veri işleme kabiliyeti ve dinamik yapısı bu tercihin önemli nedenlerinden biridir. Görüntü gibi çok uzun metrajlı veya çok parametrelili verilerde denemeye geçecek bir metot olabilir (Özşen 2007, Parichatprecha 2008, Pan 2008).

Ancak yapay zeka tek başına problemin çözümünde etkili olamayabilir. Özellikle verileri yapay zekaya uygulamadan önce yapılacak olan bir ya da birkaç ön işleme; transfer fonksiyonu seçimi, normalizasyon, filtreleme gibi doğru adımlar hedefin yakalanmasında önemli katkı sağlayabilir (Beiu 1996, Minai 1993). Verileri ham halde işlemenin her ne kadar en doğru sonucu vereceği umulsa da veriyi ön işlemeden geçirmenin performansı arttırdığı gözlemlenmiştir. Buna ilave olarak normalizasyonun yapay zekanın hızını arttırdığına dair literatürde çalışma da vardır (Shalabi 2006a).

Bu tez çalışmasında, ön işleme işlemlerden biri olan normalizasyon üzerinde durulmakta ve literatürdeki normalizasyonlara yeni bir bakış açısı getirilmektedir. Literatürde kullanılan üç tür normalizasyon oldukça yaygındır. Verilen 0 ile 1 aralığına lineer olarak sıkıştırıldığı minimum-maksimum metodu, veri ağırlık merkezi sıfır olmak üzere verileri gauss eğrisi üzerinde yayan z-skor metodu ve sayıları 1 den küçük hale onun katlarına bölerek normalize eden ondalık ölçekleme metotları en yaygın normalizasyon metotlarıdır (Shalabi 2006b, Shalabi 2006c).

Yapay zeka ile çözülmesi düşünülen problem bir tahmin verisi olmak zorunda değildir. Çözülmesi istenilen problem bir tahminden öte sınıflama problemi de olabilir. Özellikle görüntü işleme gibi yoğun veri kümelerinde yapılan sınıflamalarda yapay zeka oldukça hızlı cevap verebilir ve bir tercih sebebi olabilir (Cios 1996). Özellikle dış dünyadan alınan tıbbi veri setleri oldukça karmaşık ve birçok parametreye bağlı sağlıklı-hasta ayrımı yapılan sınıflama veri kümelerine bir örnek teşkil edebilir (Güven 2005-2006-2008). Bu tür veriler yapay zeka uygulamalarında oldukça yaygın tercih edilen giri ile çıkışın kolayca sınıflama yapılamadığı veri kümeleridir.

1.2 Çalışmanın Temel Amacı ve Literatüre Katkısı

Yapılan çalışma, literatürde çalışılan normalizasyon tekniklerine yeni bir bakış açısı getirmekte ve bu bakış açısının ileriye taşınması adına yeni bir normalizasyon metodunun da önerilmesine öncülük etmektedir. Klasik uygulamalarda tüm normalizasyon teknikleri normalizasyon türü ne olursa olsun veri setine sütun sütun uygulanmaktadır. Veri seti içerisindeki özelliklerin ayrı ayrı normalize edilmesinin bir sebebi de veri setlerinin farklı özelliklerden oluşmasından dolayı, özelliklerin kendi aralarında bağımsız olarak normalize edilmesi özellikle istatistik açısından önemlidir ve özellikleri ayrı ayrı normalizasyonunu gerektirir.

Normalizasyonun sütun bazında değil de satır satır yapılmasının performansı daha da arttırabileceği teması bu çalışmanın çıkış noktalarından birisidir. Geliştirilen metot hem sütun hem de satır bazında verilere uygulanmıştır. Satır tabanlı normalize

yapılırken veri seti içerisindeki özellikler birimsizleştirilmiştir. Özellikle tahmin amaçlı veriler için uygulanması düşünülen bu önerilen metot, literatürdeki pek çok tahmin (prediction) amaçlı verilerin yanında sınıflama verileri ile de denenmiş ve normalizasyonun satır tabanlı olmasının avantajlı olabileceği ile ilgili varsayımı destekleyen veri grubunun denenenler içerisinde azımsanmayacak oranda olduğu çalışmalar esnasında görülmüştür.

Çalışmada sunulan Öklid tabanlı normalizasyonun özellikle satır tabanlı kullanımında literatürdek kullanılan yaygın normalizasyon metotlarına göre daha başarılı sonuçlar verdiği görülmüştür. Bu başarının tüm veri setleri için geçerli olmayabileceği unutmamalıdır.

Yapılan çalışmanın sonuçları değerlendirildiğinde, önerilen satır tabanlı metodun ve normalizasyon türünün, yapay zekanın tahmin verilerinin işlenmesindeki performansını kayda değer şekilde artırabilmesi bu metodun literatüre katkısı olarak yorumlanabilir.

1.3 Çalışmanın Kapsamı ve Organizasyon

Tez çalışması 8 bölümden oluşmaktadır, Bölüm-1’de, literatür çalışmaları değerlendirilmiş ve önerilen metodun yapay zeka ile kullanılmasının getirdiği avantajlarından bahsedilerek literatüre olan katkısı anlatılmıştır.

Bölüm-2’de, literatürde kullanılan yaygın normalizasyon metotları ve önerilen normalizasyon metotları ve uygulamalarına ait ayrıntılar yapay bir veri seti üzerinde açıklanmıştır. Ayrıca bu bölüm önerilen normalizasyon metodunun hem sütun bazında hemde satır bazında kullanılabileceğine ait önermeyi de içermektedir.

Bölüm-3’de, önerilen yeni metot ve gerçek bir veri üzerindeki uygulaması ile normalizasyon metodunun veri seti üzerindeki etkisi korelasyon ve p-olasılık değerleri ile açıklanmıştır. Bunlara ek olarak, orijinal veri değerine geri dönebilmek için neler yapılması gerektiğine değinilmiştir.

Bölüm-4’de, yapay zeka tekniklerinden Yapay Sinir Ağları (YSA) ve Adaptif ağ tabanlı bulanık çıkarım sistemi (Adaptive Neural Fuzzy Inference System -

ANFIS)’den bahsedilmiştir. Bu iki popüler metodun tahmin verilerindeki performansı ve kullanım şekilleri üzerinde durulmuş parametrelerine ait açıklamalarda bulunulmuştur.

Bölüm-5’de ise, tahmin ve sınıflama verileri üzerinde durulmuştur. Kullanılan sınıflama ve tahmin veri setleri açıklanmış ve kaynakları ile ilgili açıklayıcı bilgilere yer verilmiştir.

Bölüm-6’da, tahmin ve sınıflama veri setlerinin nasıl değerlendirileceği ile ilgili kriterlere atıf yapılmış ve bu kriterler açıklanarak veri setleri değerlendirilmiştir.

Bölüm-7 öz sonuçlar ve öneriler olarak düzenlenmiştir. Verilere ait değerlendirme sonuçlarının yorumlanmasını, önerilen metot ile ilgili olarak ileriye dönük çalışmaların neler olabileceğini ve sonuç değerlendirmesini kapsamaktadır.

Bölüm 8’de son olarak bu çalışmada faydalanan kaynaklar verilmiştir.

2. NORMALİZASYON METOTLARI

Normalizasyon yapay zeka performansının artırılması adına oldukça popüler bir ön işleme haline gelmiştir. Normalizasyon $m \times n$ veri boyutuna sahip bir veri setini bir uzaydan başka bir uzaya taşır. Bu taşımada yeni maksimum ve minimum noktalar oluşur ancak veri setininin $m \times n$ olan boyutu değişmez. Bu taşıma sonunda beklenen yapay zekanın daha performanslı bir çözüm sunmasıdır. Yapılan normalizasyon yapay zekanın transfer karakteristiğine uygun ise elde edilen sonuçlar daha başarılıdır. YSA'nın *logsig* transfer fonksiyonu buna bir örnek olarak gösterilebilir. Özellikle YSA tabanlı çözüm arayışlarında transfer fonksiyonunun özel bir önemi vardır (Widrow 1990). *Logsig* transfer fonksiyonu, YSA uygulamalarındaki en yaygın kullanılan transfer fonksiyonudur ve verilerin normalizasyonunu -1 ile 1 veya (0-1 aralığı da uygundur) arasında olması mecburiyetini getirir (Vong 2005). Benzer şartlar ANFIS içinde geçerlidir. İlerleyen bölümlerde yapılan simülasyonlardaki kıyaslamalar bu durumu daha açık ortaya koymaktadır.

Normalizasyonun gerekliliğinden sonra sorgulanan bir diğer olgu ne tür bir normalizasyon olmalı ve nasıl kullanılmalı sorusudur. Veri setlerinin içerikleri birbirlerinden farklı ölçü birimindeki özelliklerden oluşabilir. Örneğin tıbbi bir veri seti için birinci sütun yaş değerleri, ikinci sütun vücut ağırlığı ve üçüncü sütun kalp atışı olabilir. Bununla birlikte veri seti sayısal olarak ifade edilmesi mümkün olmayan özelliklerde içerebilir. Örneğin; Az-çok, iyi-kötü, sağ-sol ya da + ve - gibi. Bu durumda bu tür özelliklerin de sayısal olarak anlamlandırılması gerekir.

Normalizasyon türü ne olursa olsun özellikle farklı özellikleri olan veri setlerinde veriler sütun sütun birbirlerinden bağımsız olarak normalize edilmektedir. Ayrı özellikleri olan veriler genel olarak her bir sütun bir özelliği gösterir ve normalizasyon da her bir özelliğe ayrı ayrı uygulanır. Örneğin ilk sütun kalp atışı ise bu değerler kendi içinde normalize edilirler. Devamındaki sütunlar da ayrı ayrı, bir birlerinden bağımsız olarak normalize edilirler. Normalizasyon sonunda her bir sütun

ayrı ayrı normalize edildiği için her bir sütunun maksimum ve minimum değeri aynı olur.

2.1 Literatürde Kullanılan Normalizasyon Metotları

Literatürde pek çok normalizasyon metoduna rastlanılmasına rağmen (gravitasyon, satır tabanlı gen normalizasyonu gibi) yaygın olarak kullanılan üç normalizasyon metodu vardır. Mühendislik uygulamalarında minimum-maksimum metodu ve kısmen de ondalık ölçekleme metodu kullanılmaktadır. Üçüncü metot, Z-skor metodu olup özellikle istatistik ve tıp alanında oldukça yaygın bir biçimde veri işlemede kullanılmaktadır. Şimdi sıra ile bu metotları açıklayalım.

2.1.1 Minimum-maksimum metodu

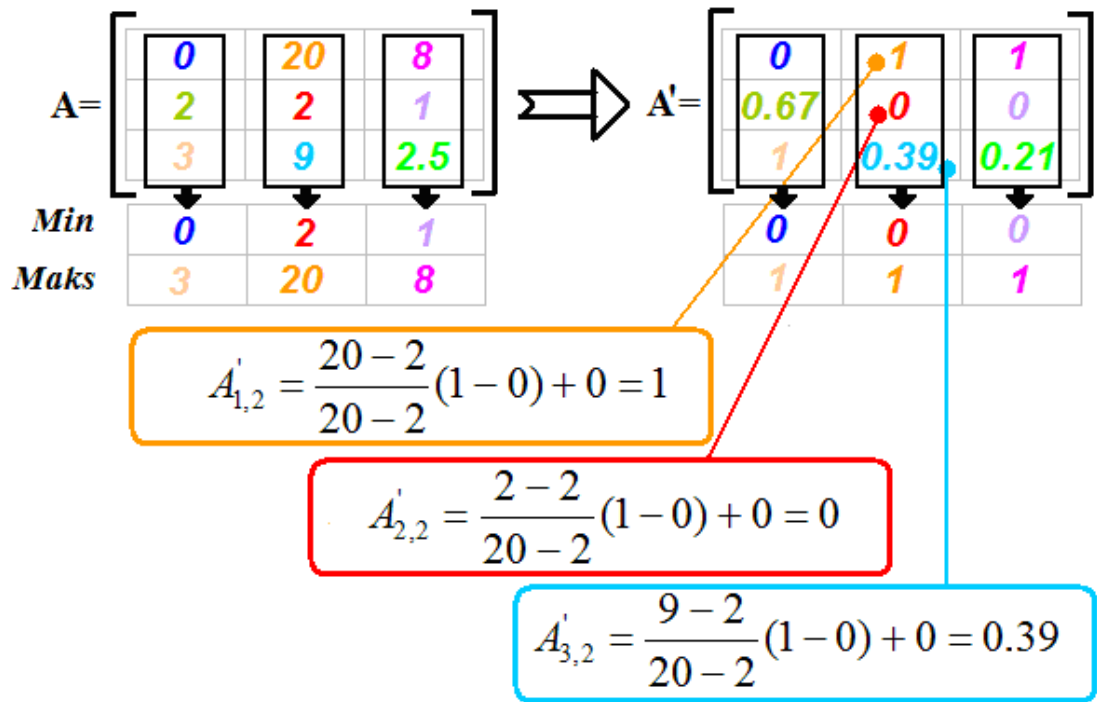
Bu metot özellikle literatürdeki mühendislik problemlerinde kullanılan en yaygın metottur. Bu mühendislik çalışmalarında tıbbi veriler, görüntü verileri gibi kaynağı mühendislik menşeli olmayan veri setlerinde de yaygın olarak kullanılmaktadır. Minimum-maksimum modelinde yeni değerler Eşitlik 2.1'den hesaplanabilir.

$$v' = \frac{v - \min_A}{\max_A - \min_A} (\text{yeni_maks_A} - \text{yeni_min_A}) + \text{yeni_min_A} \quad (2.1)$$

burada v eski ve v' yeni değer olmak üzere, \min_A değeri veri içerisindeki minimum değeri, \max_A değeri maksimum değeri ve yeni_maks_A ve yeni_min_A ise verinin taşınacağı aralığın alt ve üst değerlerini göstermektedir. Burada veri eğer 0 ile 1 arasına taşınmak isteniyorsa yukarıdaki ifade Eşitlik 2.2'ye dönüşür.

$$v' = \frac{v - \min_A}{\max_A - \min_A} \quad (2.2)$$

Bu metot veri setinin doğrusal bir transformasyonunu sağlar. Herhangi iki değer arasındaki fark ya da değerlerdeki büyüklüklerin orantısallığı değişmez. Sayısal veriler içerisinde negatif işaretli değerler olsa bile bu işaretler normalizasyon sonunda ortadan kalkar ve değerler pozitif işaretli hale gelirler. Bu metodun nasıl uygulandığı ile ilgili sayısal bir matris örneği Şekil 2.1’de gösterilmiştir.



Şekil 2.1 Minimum-Maksimum metodunun kullanılış şeklini gösteren diyagram

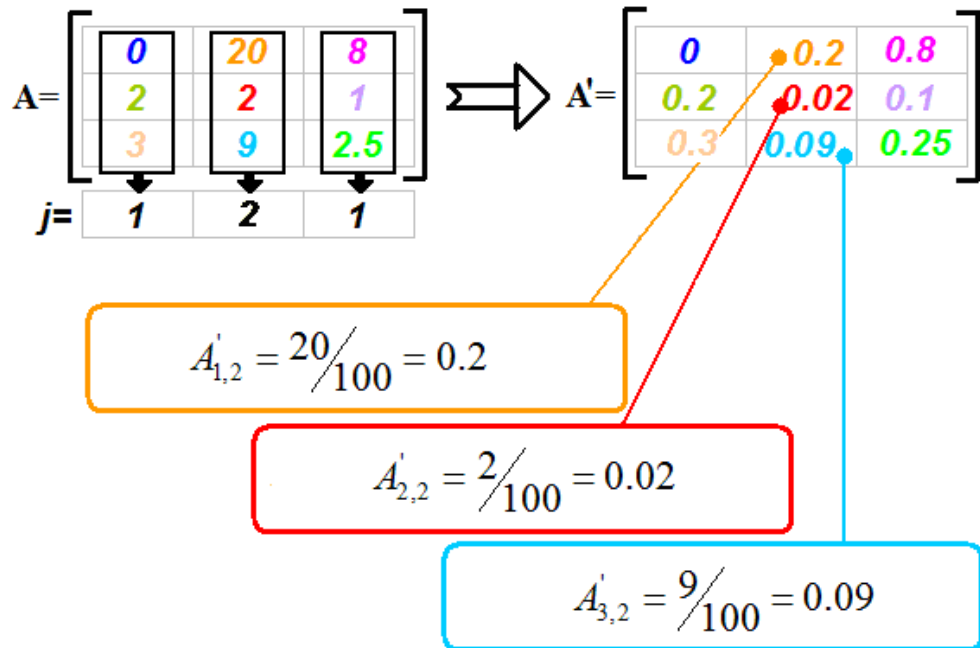
Şekil 2.1’den de görüldüğü gibi A matrisi 3x3 elemanlı bir kare matristir. Metodun uygulanabilmesi için öncelikle sütunlara ait maksimum ve minimum değerlerinin bulunması gerekir. A matrisinin minimum-maksimum metoduna göre normalize edilmiş haline dikkate edilirse her sütun için bir sıfır ve bir değerinin olduğu görülür. Bu metotta elde edilen bir maksimum ve minimum noktası tüm özellikler aynı olacak şekilde minimum noktası 0, maksimum nokta da 1 olur.

2.1.2 Ondalık ölçekleme metodu

Ondalık ölçeklemeye literatürlerde rastlanmasına rağmen, Minimum-Maksimum metodu kadar yaygın olarak kullanılmayan bir metottur. Genel düşünce sayıları 1 den küçük hale getirmek için 10 veya 10'un kuvvetlerine bölme felsefinden oluşmaktadır. Eşitlik 2.3'den de görülebileceği üzere ondalık normalizasyonun tanımı j de yatmaktadır ve j , A değerini birden küçük yapan en küçük değer olarak tanımlanabilir.

$$A' = \frac{A}{10^j} \quad (2.3)$$

Sonuçta aralık -1 ve +1 arasındadır. Şekil 2.2 örnek bir matrisin nasıl ondalık ölçekleme yapıldığını göstermektedir.



Şekil 2.2 Ondalık ölçeklemenin nasıl yapıldığını anlatan akış diyagramı

Öncelikle sütunlardaki maksimum değeri 1 den küçük yapan j değerleri bulunur ve ilgili sütunun, j değerine tüm elemanlarının bölümü ile yeni değerler elde edilmiş olur. Örnekteki matris için 1. ve 3. özellik için $j=1$ ve orta sütun için bulunan j değeri 2'dir ve Eşitlik 2.3'de yerin yazıldığında ikinci sütun için bölen olarak 100, 1. ve 3. sütun için ise 10 elde edilmiş olur. Daha sonra tüm sütuna ait elemanların ilgili katsayıya bölünerek yeni yeni $A_{(j,n)}$ değerleri elde edilmiş olur.

2.1.3 Z-skor metodu

Z-skor istatistik değerlendirilmelerde sıklıkla kullanılan bir ön işleme metodudur. Değerlerin normalize edilebilmesi için standart sapma ve ortalama değerlerin öncelikle hesaplanması gerekir. Z-skor verilerin ortalama değeri ile bu değerlerin ortalamadan olan uzaklıkları ile ilgili yeni değerler üretir¹. Z-skor normalizasyondan sonra ortalaması sıfır olan bir gauss dağılımı sunar. Standart sapma, varyansın kare köküdür ve Eşitlik 2.4'deki gibi hesaplanabilir.

$$S = \left(\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{\frac{1}{2}} \quad (2.4)$$

Bununla birlikte standart sapma ile bereber verinin özelliğine ait ortalama değerinde hesaplanması gerekir. Ortalama değer formülü Eşitlik 2.5'deki gibi hesaplanır.

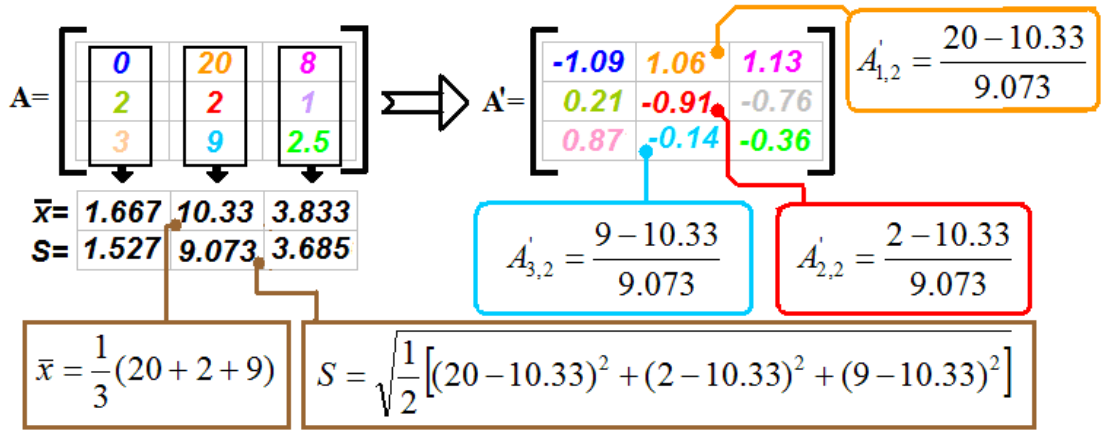
$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.5)$$

¹ http://en.wikipedia.org/wiki/Standard_score

Eşitlik 2.6, \bar{x} ortalama değer ve S standart sapma olmak üzere, verilen değerleri z-skor değerlerine çevirmiş olur.

$$A' = \frac{A - \bar{x}}{S} \quad (2.6)$$

Burada A olan değer ve A' ise z-skor normalizasyonuna göre elde edilen yeni değerdir. Z-skor, her ne kadar $-\infty$ ile $+\infty$ arasında tanımlı olsada dağılım genellikle -1.5 ile $+1.5$ arasında gerçekleşir. Şekil 2.3 Z-skora ait bir örnek matrisin nasıl hesaplandığını göstermektedir.



Şekil 2.3 Z-skor hesaplanmasını gösteren akış diyagramı

Bu ifadelerde her bir sütuna ait ortalama (\bar{x}) ve standart sapma (S) değerleri ayrı ayrı hesaplanmaktadır. Z-skor metodu normalizasyon adına bu kadarı kullanılmakla birlikte, istatistik disiplininde devamına ait uygulamalar vardır. Z-skor dağılımından veriye ait gauss dağılımı ya da gauss dağılımının yeri ile ilgili olarak güvenlik aralığı hesaplama gibi daha ileri uygulamaların da başlangıcıdır. Literatürde özellikle istatistiksel çözümlemelerde oldukça yaygın bir kullanımı vardır (Lawlor 2006, Nelson 2004, Jiang-Ming 2007).

2.2 Önerilen Normalizasyon Metodu

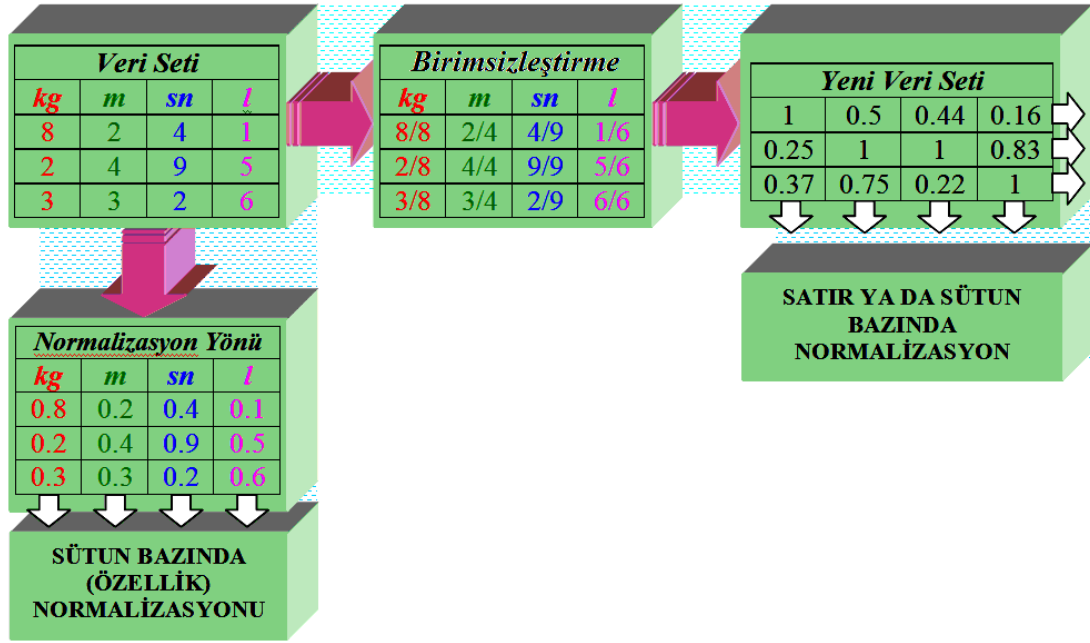
Çalışmada kullanılan ve önerilen normalizasyon metodu klasik normalizasyonlardan farklı olarak Öklid uzaklığını normalizasyon ölçütü olarak kullanmaktadır. Önerilen model hem satır hemde sütunda uygulanabilmektedir. Yalnız satır bazında uygulanabilmesi için veri tipinin aynı birimde olması veya birimsizleştirme işleminden geçmesi gerekmektedir.

2.2.1 Birimsizleştirilmiş satır tabanlı normalizasyon metodu

Literatürde minimum-maksimum, z-skor ve ondalık ölçekleme yaygın olarak kullanılmaktadır. Önerilen metod literatürde yaygın olarak kullanılan modellere yeni bir bakış açısı getirmekte; Öklid uzaklığının normalizasyon işleminde kullanarak yapay zekânın işleyişine ait performans verilerinin artabileceği tüm veri türleri için geçerli olmasa da yapılan çalışmalarda görülmüştür. Önerilen metod hem satır hemde sütun bazında kullanılabilir yapıdadır.

Satır bazındaki uygulamalarda dikkat edilmesi gereken noktalardan biri; veri yapısı farklı birimlere ait özellikler içeriyorsa, veri seti içerisindeki farklı birimlerden kurtulmak ve veri yapısını birimsizleştirmek gerekmektedir. Matematik bilminde farklı birimlerdeki sayıların toplanıp ya da çıkarılması yanlıştır. Şekil 2.4 veri setinin nasıl normalize edildiğini göstermektedir.

Farklı özellikler içeren bir veri seti özellik-özellik normalize edilmesi için normalizasyonun sütun bazında yapılması yeterlidir ve her bir sütun aynı birimdeki sayı türlerinden oluşmaktadır. Örneğin tüm özellik *kg* cinsinden elde edilmiş değerlerdir. Veri seti birimlerden arındırıldığı zaman normalizasyon işlemi artık hem satır hem de sütun bazında gereksinime göre kullanılabilir ve veri seti içerisindeki değerler basit birer sayılar haline dönüşmüş olur. Böylece normalizasyonun farklı birimlerden oluşan veri setleri için satır bazında uygulanabilmesinin önü açılmış olur.



Şekil 2.4 Önerilen ve klasik metodun veri setine uygulanışının gösterimi

2.2.2 Veri setinin birimsizleştirilmesi

İçerisinde farklı birimlerdeki özelliklerden oluşan bir veri seti üzerinde işlem yapabilmek için veri setindeki özelliklerin birimlerinden ayıklanması gerekir. Matematikte en çok duyulan sözlerden bir tanesi “elma ile armutlar toplanmaz” sözüdür. Elmalar ile armutların toplanabilmesi mümkünmüdür? Bu sorunun cevabı şartlı olarak “evet” olabilir. Eğer elma ile armuttan bahsedilirken her ikisinde kg ise toplamakta bir mahzur yoktur (Munson 2002).



Yapılan çalışmalarda birimlerin arındırılması için veri özelliği, özelliğin en büyük değerine bölünerek veri içerisindeki değerlerin salt birer sayı haline

dönüştürülmesi sağlanmıştır. Birimsizleştirmek için Eşitlik 2.7, özelliğin her bir değerine birimsiz oranlar atamak için kullanılabilir.

$$A'_{(i,j)} = \frac{A_{(i,j)}}{A_{(j_{maks})}} \quad (2.7)$$

Burada $A_{(i,j)}$ değeri gözlem değeri, $A_{(j_{maks})}$ ise gözlemler içerisindeki maksimum değer ve $A'_{(i,j)}$ ise birimsizleştirilmiş değerdir. İşlem sonunda veriye ait her bir özelliğe aşağıdaki kurallar doğrultusunda yeni değerler atanmış olur. Oluşan yeni değerler -1 ile +1 arasında olur. Eğer ki özellik kümesi 0 ve negatif işaretli sayılar içermiyorsa ortaya çıkan aralık $0 < A'_{(i,j)} \leq 1$ şeklinde olacaktır. Yalnız tanımlı olan bölge $-1 \leq A'_{(i,j)} \leq 1$ arasındadır.

2.2.3 Önerilen normalizasyon metodunun matematiksel tanımı

Yapılan tez çalışmasında önerilen normalizasyon modeli özelliklerin Öklid uzaklıklarını normalizasyonun bir enstrümanı olarak kullanmakta ve normalize edilmiş verinin elde edilmesinde Öklid uzaklığının kullanılmasını önermektedir. Öklid uzaklığı Eşitlik 2.8'de ifade edildiği gibidir.

$$Z = \sqrt{A_{(1)}^2 + A_{(2)}^2 + \dots + A_{(n)}^2} \quad (2.8)$$

Uzaklığın hesaplanmasından sonra satır yada sütun bazında normalizasyonun yapılması mümkündür. Her yeni değer özelliğin değerinin uzaklığa oranı şeklinde oluşacaktır ve her yeni değer Eşitlik 2.9'daki gibi hesaplanabilir. Veri seti 10 özellikli bir veri ise 10 adet Öklid uzaklığı oluşmuş olur.

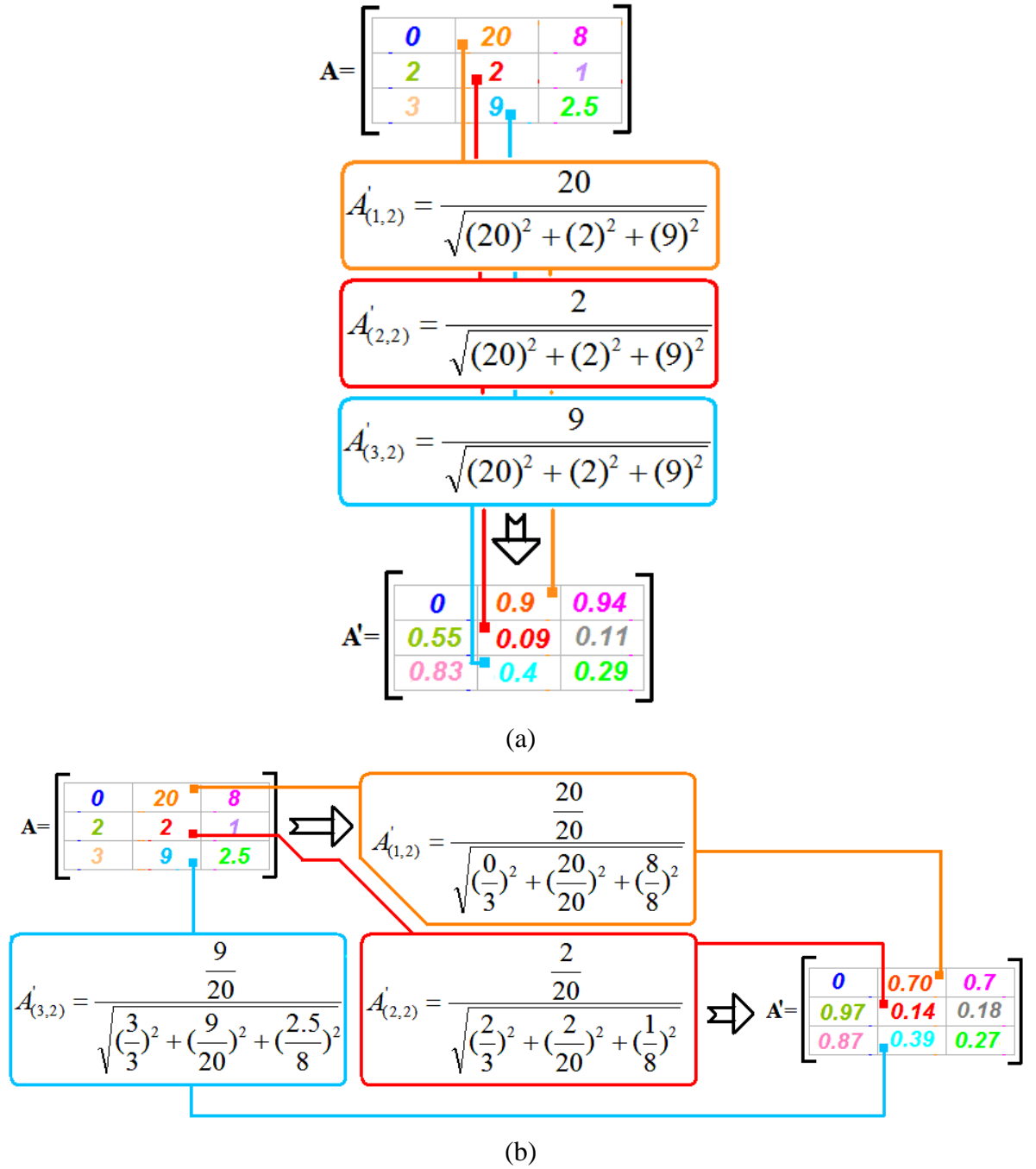
$$A'_{(i,j)} = \frac{A_{(i,j)}}{Z} \quad (2.9)$$

Burada $i=1,2,...n$ ve $j=1,2,...n$ dir. Veri setinde farklı özellikler içeren değerler var ise bu durumda veri setinin birimsizleştirilmesi gerekmektedir. bu durumda önerilen normalizasyon eşitliğinin içerisine birimsizleştirilmesi eşitliği de girmiş olur. Bu, Eşitlik 2.2 ile Eşitlik 2.9'un bir bileşimidir ve Eşitlik 2.10 ile ifade edilebilir. $m \times n$ boyutlu bir veri seti m adet $1 \times n$ boyutlu özelliklerden oluşur. normalizasyon sonunda üretilecek olan normalize değerler Eşitlik 2.10'dan hesaplanabilir.

$$A'_{(i,j)} = \frac{\frac{A_{(i,j)}}{A_{(j_{maks})}}}{\sqrt{\sum_{j=1}^j \left(\frac{A_{(i,j)}}{A_{(j_{maks})}} \right)^2}} \quad (2.10)$$

burada, $A'_{(i,j)}$ elde edilen normalize edilmiş değeri, $A_{j_{maks}}$ ise her sütuna ait maksimum değeri göstermektedir.

Oluşan yeni değerler birimsiz olup normalize edilmiş değerlerdir. Tez çalışmasında önerilen Eşitlik 2.9 sütun bazında normalizasyon için, Eşitlik 2.10 ise satır bazında normalizasyon için kullanılmış ve bu bağlamda değerlendirilmiştir. Bununla birlikte tüm veri seti aynı birimdeki değerlerden oluşmakta ise bu durumda satır boyunca Eşitlik 2.9'da kullanılabilir. Şekil 2.5'de önerilen normalizasyon metodunun uygulanışı gösteren sayısal bir örnek vardır. Uygulama yönü satır yada sütun olabilir. Satır bazında yapılan çalışmalara İngilizce adına itafen “**Line Based Normalization Method**” (LBNM) denilmiştir (Akdemir 2008a, Akdemir 2008b).



Şekil 2.5 Önerilen metodun satır ve sütun boyunca nasıl kullanıldığını gösteren nümerik bir örnek (a) sütun boyunca uygulanan normalizasyon (b) satır boyunca uygulanan normalizasyon

Tez içerisinde satır tabanlı uygulama için LBNM kodu ve aynı formülün sütun bazında kullanılmasına ise EBNM (Euclidean Based Normalization Method) kısaltması kullanılmıştır.

3. ÖNERİLEN NORMALİZASYON METODUNUN UYGULANMASI

Önerilen normalizasyon metodunun veri üzerindeki etkisini inceleyelim. Bunun için veri setini minimum-maksimum ve satır bazında önerilen modele göre normalize edelim. Örnek veri seti için diş anterior temporal veri seti kullanılmıştır. Tek giriş ve tek çıkışı olduğundan önerilen metodun gösteriminin anlaşılabilirliği açısından diş verileri tercih edilmiştir. Dolayısı ile tek girişli tek çıkışlı verilerde verilerde birim problemi yoktur. Kullanılan veri setine ait ham ve normalize edilmiş değerler Tablo 3.1’de verildiği gibidir. Veri seti Erciyes Üniversitesi Diş Hekimliği Fakültesi’nde Prof. Dr. Sadık Kara ve ekibinin elde ettiği ve literatürde kullandığı bir veridir.

Tablo 3.1 Tek değişkenli Diş Anterior Temporal verisinin ham, minimum-maksimum ve satır tabanlı normalizasyonlarına ait sonuçları aşağıdaki gibidir.

Ham Veriler		Min.-Mak. Metodu		Satır Normalizasyonu	
Giriş	Çıkış	Giriş	Çıkış	Giriş	Çıkış
1.484640	1.999270	1.000000	0.783531	0.787152	0.616759
0.000040	0.000020	0.000007	0.000000	0.960192	0.279341
0.000040	0.000050	0.000007	0.000012	0.808723	0.588189
0.000100	0.000150	0.000047	0.000051	0.753410	0.657551
0.000110	0.267600	0.000054	0.104868	0.000706	1.000000
0.000200	0.423000	0.000115	0.165771	0.000813	1.000000
0.000310	0.000290	0.000189	0.000106	0.878319	0.478075
0.000390	0.004320	0.000242	0.001685	0.153323	0.988176
0.000910	1.005020	0.000593	0.393872	0.001556	0.999999
0.001210	0.003670	0.000795	0.001430	0.492999	0.870030
0.001230	0.107510	0.000808	0.042127	0.019659	0.999807
0.001450	0.074730	0.000956	0.029280	0.033329	0.999444
0.001500	0.019700	0.000990	0.007713	0.129757	0.991546
0.001800	0.796000	0.001192	0.311955	0.003886	0.999992
0.003810	0.042700	0.002546	0.016727	0.151580	0.988445
0.004730	0.298310	0.003166	0.116904	0.027241	0.999629
0.005580	0.223050	0.003738	0.087408	0.042956	0.999077

0.013400	0.308710	0.009006	0.120979	0.074395	0.997229
0.014610	0.506600	0.009821	0.198535	0.049505	0.998774
0.015810	0.424330	0.010629	0.166292	0.063905	0.997956
0.016390	0.012090	0.011020	0.004730	0.918938	0.394403
0.020060	0.025920	0.013492	0.010151	0.799303	0.600928
0.022040	0.064130	0.014825	0.025126	0.508575	0.861017
0.024430	0.542000	0.016435	0.212409	0.077236	0.997013
0.025420	0.072400	0.017102	0.028367	0.516656	0.856193
0.033500	1.926900	0.022545	0.755168	0.029867	0.999554
0.033700	0.046120	0.022679	0.018067	0.782286	0.622919
0.039700	0.072640	0.026721	0.028461	0.684642	0.728880
0.061120	0.901400	0.041149	0.353262	0.115752	0.993278
0.076110	0.367140	0.051246	0.143879	0.335623	0.941996
0.092820	0.358500	0.062501	0.140493	0.406551	0.913628
0.094900	0.125470	0.063902	0.049165	0.792608	0.609732
0.095200	1.025300	0.064104	0.401820	0.157586	0.987505
0.099260	0.105670	0.066839	0.041406	0.850124	0.526583
0.150880	0.211620	0.101609	0.082929	0.774755	0.632261
0.304450	0.336180	0.205050	0.131745	0.841321	0.540536
0.312640	0.433550	0.210567	0.169906	0.778255	0.627948
0.465600	0.540110	0.313598	0.211668	0.828865	0.559448
0.635530	2.551610	0.428059	1.000000	0.393530	0.919312
0.000030	0.000070	0.000000	0.000020	0.593059	0.805159

Veri setinin normalizasyonunda veri boyutunda bir değişme söz konusu değildir. Veri seti sahip olduğu $m \times n$ boyutunu korumaktadır. Minimum-maksimum normalizasyonu Eşitlik 2.2'ye göre satır tabanlı normalizasyon ise Eşitlik 2.10'a göre yapılmıştır. Elde edilen normalize edilmiş veri setlerinin normalizasyon türlerine göre giriş ile çıkış arasındaki ilişki katsayısı olan korelasyonuna ve p -değerine bakılmıştır. p -değeri 0.05'den küçük olduğu durumda giriş ile çıkış arasında olası istatistiksel bir bağlantıdan bahsedilebilir.² p değerinin uygunluğu güçlü bir istatistik

² Matlab Ver. 7.4.0.287 On-line help tools

bağıntıyı açıklar (Onbaşılı 2002, Saltık, 1997). Korelasyon doğrusal bir aitin ölçüsüdür.

Korelasyon sonucu;

a) +1; Bu durumda iki değişken arasında %100 artan doğrusal bir ilişki vardır. Örneğin $y=2x$ ve $y=x$ doğrusu gibi

b) 0; Bu durumda iki değişken arasında doğrusal bir ilişki oluşturabilecek bir bağıntı söz konusu değildir.

c) -1; Bu durumda iki değişken arasında %100 azalan doğrusal bir ilişki vardır. Örneğin $y=2x$ ve $y=-2x$ doğrusu gibi

Tablo 3.2 iki özellikli tek giriş ve tek çıkışlı diş anterior verisine ait normalizasyon metotlarına bağlı olarak oluşan korelasyon değerlerini göstermektedir. Tablo 3.2’den tüm sütun bazında uygulanan modellerde doğrusal ilişki aynı kalmasına karşın satır bazında yapılan normalizasyonda doğrusal ilişki katsayısı (korelasyon) değeri kayda değer şekilde arttığı gözlemlenebilir.

Tablo 3.2 Tek değişkenli Diş Anterior Temporal verisinin normalizasyon metotlarına göre giriş ile çıkış arasındaki korelasyonun değişimini gösterir tablo.

	Normalizasyon metoduna göre korelasyon değerleri					
	Ham veri	EBNM	LBNM	Min.-Mak.	Ondalık	Z-Skor
Anterior	0.478	0.478	-0.9538	0.478	0.478	0.478

Tablo 3.3’de tüm normalizasyon modellerine göre istatistiksel bir aitin varlığı kaçınılmazdır. Ancak önerilen metoda göre olan normalizasyon şeklinde bu bağıntının daha belirgin olduğu sayının sıfıra daha da yaklaşmasından bellidir. Ayrıca Tablo 3.3 veri setinde giriş ile çıkış arasındaki olası istatistiksel bağıntıya işaret eden *p-olasılık* değerleri verilmiştir. *p*-değerinin 0.05 ten küçük olması kuvvetli bir istatistiksel bağıntıya işaret eder.

Tablo 3.3 Tek değişkenli Diş Anterior Temporal verisinin normalizasyon metotlarına göre giriş ile çıkış arasındaki *p-olasılık* değişimini gösterir tablo.

	Normalizasyon metoduna göre p-değeri sonuçları					
	Ham veri	EBNM	LBNM	Min.-Mak.	Ondalık	Z-Skor
Anterior	0.0018	0.0018	0	0.0018	0.0018	0.0018

Korelasyon katsayısının artması verinin giriş ile çıkış arasındaki doğrusal aitin arttığının birleşimidir. Doğrusal aitin artması ve *p-olasılık* değerinin daha da kuvvetlenmesi yapay zekanın daha performanslı sonuç üretmesinin bir sebebi kabul edilmiştir.

3.1 Satır Normalizasyonunun Veri Üzerindeki Etkisi

Tezde çalışılan veri setlerinin önerilen normalizasyon metodu sonunda korelasyon ilişkisinin sonucu etkileyecek miktarda artması ve yapay zeka performansının da daha başarılı sonuçlar bulması umulmaktadır. Ancak bunun tüm veri setleri için geçerliliğinin olmayabileceğini gözardı etmemekle birlikte korelasyonun artması sonucu yapay zekanın problemi daha az hata ile çözdüğü yargısına varılabilir. Tablo 3.4, Tez çalışmasında kullanılan veri setlerinin normalizasyon sonucunda oluşan korelasyon değerlerini göstermektedir.

Tablo 3.4 Korelasyon değerlerini gösterir tablo

	Ortalama Korelasyon Değerleri					
	Ham veri	EBNM	LBNM	Min.-Mak.	Ond	Z-Skor
Diş Anterior	0.478	0.478	-0.9538	0.478	0.478	0.478
Diş üst dudak	0.2306	0.2306	-0.9521	0.2306	0.2306	0.2306
Kalp Sistol	0.427867	0.427867	0.462967	0.427867	0.427867	0.427867
Kalp Aort	0.546733	0.546733	0.590467	0.546733	0.427867	0.427867
Beton	0.228344	0.228344	0.303167	0.228344	0.228344	0.228344

Korelasyon doğrusal bir ilişkiyi sembolize ederken, p-istatistik değeri ise istatistiksel bir bağın varlığını simgeler. İstatistiksel bir bağdan bahsedebilmek için p değerinin 0.05'den küçük olması gerekmektedir. Korelasyon değeri ise bire yakın olmak koşulu ile negatif ya da pozitif değer almasında bağıntının güçlenmesi adına

bir fark yoktur. Tablo 3.5’de veri setinin normalizasyon sonunda oluşan *p-olasılık* değerleri verilmiştir.

Tablo 3.5 Veri setlerine ait *p-olasılık* değerlerini gösterir tablo

	P olasılık değerleri					
	Ham veri	EBNM	LBNM	Min.-Mak.	Ond	Z-Skor
Diş Anterior	0.0018	0.0018	0	0.0018	0.0018	0.0018
Diş üst dudak	0.1522	0.1522	0	0.1522	0.1522	0.1522
Kalp Sistol	0.023967	0.023967	0	0.023967	0.023967	0.023967
Kalp Aort	0.001367	0.001367	0	0.001367	0.023967	0.023967
Beton	0.0000778	0.0000778	0.000111	0.0000778	0.0000778	0.0000778

3.2 Geri Çözümleme

Diğer bir önemli konu normalize edilmiş verinin tekrar orijinal (ham) haline dönüştürülebilmesidir. Kullanılan her metodun beraberinde getirdiği avantajlar ve dezavantajlar olabilir. Normalizasyon metodundan beklenen, veriler normalize olduktan sonra geriye de çözümlenebilmeli, kapalı bir kutu haline dönüşmemelidir. Bununla birlikte sonradan da veri ilave edilebilmelidir. Şimdi sıra ile yaygın olan normalizasyon metotlarının olası zayıf ve güçlü taraflarının irdeleyelim.

3.2.1 Minimum-Maksimum metodu

Minimum-Maksimum metodu literatürde en çok kullanılan metottur. Uygulama şekli de göz önünde tutulduğunda minimum-maksimum metodunda verilerin geri elde edilebilmesi için her sütuna ait maksimum ve minimum değerlerinin bilinmesi gerekir. Bununla birlikte bu metodun en büyük dezavantajı sonradan veri ilave etme sırasında ortaya çıkmaktadır. İlave edilmesi gereken veri önceden normalize edilmiş ham veri aralığının dışında ise tüm sistem alt üst olmaktadır. Önceden işlenmiş tüm veri baştan tekrar edilmesi gerekmektedir. Bu

yüzden mümkün mertebe bu normalizasyon türünde olası en yüksek verinin de setin içerisinde olmasında fayda vardır.

3.2.2 Ondalık Ölçekleme metodu

Bu metodun özelliği veri seti içerisinde en büyük basamaklı değeri 1 den küçük hale getiren 10'un katından oluşan bir sayıya bölmekte yatar. Bu yüzden olası bir taşma meydana gelmesi minimum-maksimum metoduna göre daha zordur. Bununla birlikte geri çözümleme oldukça kolaydır. Normalize edilmiş değerin sadece katsayısına geri çarpılması orijinal değerlere dönülmesini sağlar. Ancak uzun soluklu verilerde minimum-maksimum metodunda olan taşma burada da meydana gelebilir. Örneğin finansman verilerinin işlendiği bir yapay zeka metodu için ondalık ölçeklemenin kullanıldığını varsayalım. Teorik olarak finansman verisi 10 kattan daha fazla artarsa yine tüm normalizasyonu baştan yapmak gerekecektir. Dolayısıyla ondalık ölçekleme kullanılırken bu tür bir ihtimal göz ardı edilmemelidir. Bu metotta sonradan veri eklenilmesi kolay olmakla birlikte eklenecek veri değeri ondalık bölme işlemi yapıldığı halde 1'den büyük bir girdi oluşturuyorsa bu durumda önceki normalizasyonların değiştirilmesi gerekecektir.

3.2.3 Z-skor metodu

Bu metodun problemleri de yukarıda sayılan minimum-maksimum ve ondalık normalizasyon metodlarının problemleri ile aynıdır. Normalizasyon tamamlandıktan sonra yeni bir değer ilave etmek neredeyse imkansızdır. Özellikle veri setinin dışında olan bir değeri eklemek çok daha zordur. Z-skor veriyi ortalaması sıfır olacak şekilde bir gauss zarfı üzerinde dağıtır. Bununla birlikte z-skor normalizasyonundan sonra bu veri setine veri eklemek işlemi başa almak demektir. Z-skor normalizasyonunda formüle ait değerlerin bilinmesi ham veriye dönüş sağlar.

3.2.4 LBNM metodu

Satır tabanlı normalizasyonun en büyük kolaylıklarından biri gözlemin eklenecek gözlem değerleri bilindiğinde kolayca eklenebilmesidir. Bununla birlikte geri çözümleme o kadar kolay değildir ve her özelliğe ait maksimum ve minimum değerlerinin bilinmesinin yanında her gözleme ait bir parçanın bilinmesi gerekliliği vardır.

4. YAPAY ZEKA

Yapay zeka kavramını tanımlayabilmek için öncelikle zeka kavramını tanımlamak yerinde olacaktır. Zeka, bireyin amaçlı bir biçimde hareket edebilme, mantıklı düşünebilme ve çevresine uyum sağlayabilen yetilerinin tamamı olarak tanımlamak mümkündür (Sağıroğlu 2003).

Zeka kavramının tanımı ışığında yapay zeka kavramını insan zekasına özgü olan bilişsel fonksiyonları, algılama, öğrenme, çoğul kavramları bağdaştırabilme, fikir yürütme, sorun çözme, iletişim kurma, çıkarımsama yapabilme yeteneklerini sergilemesi beklenen yapay işletim sistemi olarak tanımlamak mümkündür³. Başdöndürücü bir hızla gelişen bilgisayar ve teknolojiye rağmen yapay zekanın şimdiki yakalayabildiği zeka seviyesi henüz 3-4 yaşlarındaki bir çocuğunki kadardır. Bununla birlikte yapay zekanın hedefi yetişkin bir insanın zekasıdır.

Bu çalışmada yapay zeka modellerinden yapay sinir ağları (YSA) ve ANFIS kullanılmıştır. Her iki yapay zeka algoritmaları literatürde oldukça yaygın olarak tercih edilen ve neredeyse gelişimlerini tamamlamış algoritmalarlardır. YSA ağırlıklandırılmış ileri beslemeli bir öğrenme yapısı varken ANFIS dilsel tanımlamaları kullanan daha çok şartlı kurallar zincirini kullanan bir algoritması vardır.

4.1 Yapay Sinir Ağları

Yapay sinir ağları, yapay zeka modelleri içerisinde en çok rağbet görenlerinden bir tanesidir. Anlaşılabilir yapısı ve zorlu problemlere olan kolay

³ http://tr.wikipedia.org/wiki/Yapay_zeka

uygulanabilirliği YSA'nın çekiciliklerinden yalnızca bir kaçıdır. YSA'nın basit ve anlaşılabilirliği sayesinde farklı problemlere yönelik farklı türlerinin de ortaya çıkmasına yol açmıştır. Örneğin zamanla değişen ve zamanın bir parametre olduğu veriler için dinamik YSA ve parametrelerin hızlı tahmin edilmesi için Levenberg-Marquart öğrenme gibi pek çok çeşit YSA modeli ortaya çıkmıştır. YSA'nın diğer modellere göre örneğin karınca koloni algoritması, ANFIS gibi daha eski bir geçmişi vardır.

4.1.1 Yapay sinir ağları tarihi

İlk araştırmalar ve yapay sinir ağları idealize edilmiş tanımıyla yapay zeka konusundaki ilk çalışmalardan biri McCulloch ve Pitts tarafından yapılmıştır (McCulloch 1943, McCulloch 1947). Bu araştırmacıların önerdiği, yapay sinir hücrelerini kullanan hesaplama modeli, önermeler mantığı, fizyoloji ve Turing'in hesaplama kuramına dayanıyordu. Herhangibir hesaplanabilir fonksiyonun sinir hücrelerinden oluşan ağlarla hesaplanabileceğini ve mantıksal ve-veya işlemlerinin gerçekleştirilebileceğini gösterdiler. Bu ağ yapılarının uygun şekilde tanımlanmaları halinde öğrenme becerisi kazanabileceğini de ileri sürdüler. Hebb, sinir hücreleri arasındaki bağlantıların şiddetlerini değiştirmek için basit bir kural önerince, öğrenebilen yapay sinir ağlarını gerçekleştirmek de olası hale geldi (Hebb 1949).

Çalışmalarını Princeton Üniversitesi'nde sürdüren Mc Carthy, Minsky, Shannon ve Rochester'le birlikte 1956 yılında Dartmouth'da iki aylık bir açık çalışma düzenledi. Bu toplantıda birçok çalışmanın temelleri atılmakla birlikte, toplantının en önemli özelliği Mc Carthy tarafından önerilen Yapay zeka adının konmasıdır (Mc Carthy 1969). İlk kuram ispatlayan programlardan Logic Theorist (Mantık kuramcısı) burada Newell ve Simon tarafından tanıtılmıştır (Gugerty 2007). Daha sonraki yıllarda daha çok mantıksal tabanlı çalışmalar egemen olmuş, YSA ve diğer geliştirilmeye çalışılan yapay zeka modelleri için suni problemler geliştirilmiş ve çözümler üretilmiştir. Belki de zekanın temellerinin tam olarak tanımlanamamasından kaynaklanan sıkıntılardan YSA'nın gerçek problemlere

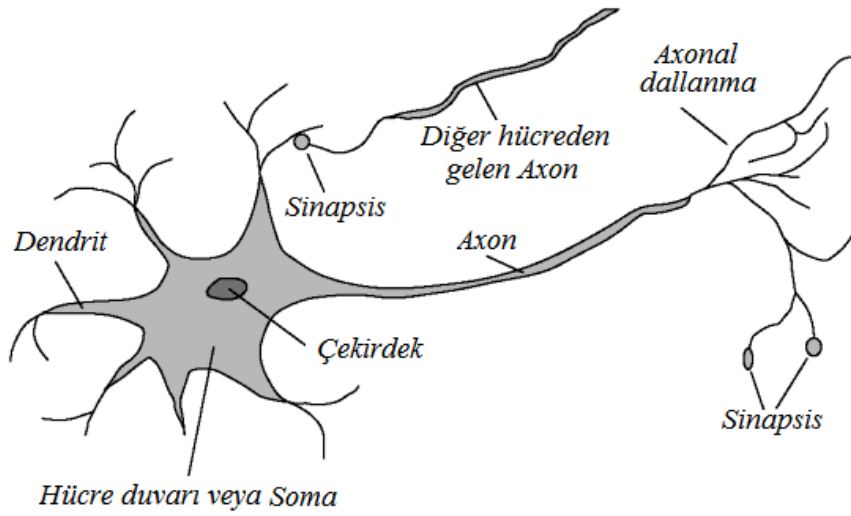
verdiği uyumsuz cevaplar nedeniyle YSA'nın yalnızca suni problemleri çözebildiği, gerçek problemleri çözemediği tezi ortaya atılmıştır. Zeka kavramını simüle etmek için bu çalışmalarda kullanılan temel yapılardaki bazı önemli yetersizliklerin ortaya konmasıyla birçok araştırmacılar çalışmalarını durdurdular. Özellikle Minsky ve Papert'in 1969'da yayınlanan Perceptrons adlı kitaplarında tek katmanlı nöronlarda bazı basit problemleri çözemeyeceğini gösterip aynı kısırlığın çok katmanlı nöronlarda da beklenilmesi gerektiğini söylemeleri ile YSA çalışmaları keskin bir şekilde kesilmiştir (Minsky 1969). Bu kısırlık probleme yönelik noktasal çözümler geliştirme yönüne doğru bir kayma oluşturdu. Bu yönelim “uzman sistemler” olarak genel yapıdan ayrıldı ve problemlere yönelik noktasal başarılı çözümler yapay zekaya olan ilgiyi arttırdı. DEC firmasının müşteri taleplerini yönlendiren R1 uzman sisteminin şirkete 40 milyon dolarlık tasarruf yaptırması yapay zeka ve özellikle YSA'nın önünü açmış oldu. Birden bire yapay zekaya olan talep arttı ve 1988'de yapay zeka endüstrisi 2 milyar dolara çıktı⁴.

YSA'nın kolay anlaşılabilir yapısı ve zorlu problemler de olsa kolay uygulanabilirliği cazibesini arttırmıştır. YSA veri işlemede yeni ufuklar açmıştır. YSA, yerel bir hafıza yapısı tanımlayabilen ve birbirleri ile çeşitli şekillerde bağlantılı olan veri işleme elemanlarının dağıtılmış veriyi işleyebildikleri paralel bir yapıdır. Veri işleme veya işlemci elemanların (nöronların) bir ağ yapısı tanımlayacak şekilde birbirleri ile bağlanmaları fikri beyin hücrelerinin birbirleri ile olan bağlantılarından esinlenilerek geliştirilmiştir. Beyinde her bir nöronun kendinden sonraki nöronlarla olan bağlantısı vardır. YSA içerisindeki nöronların fazlalığı insan beyninin hasarlara karşı gösterdiği yapısal esnekliği YSA'ya da kazandırmaktadır.

⁴ http://tr.wikipedia.org/wiki/Yapay_zek%C3%A2#Tarih.C3.A7e

4.1.2 Yapay sinir ağlarının yapısı

Yapay sinir ağlarının çıkış noktası insan beyninin çalışma felsefesidir. Bu bağlamda insan beynindeki nöronların nasıl çalıştığını ve iletişimin nasıl sağlandığını irdelemek YSA'nın çalışmasını anlamak adına iyi bir başlangıç kabul edilebilir. Biyolojik sinir ağını oluşturan bir nöronun resmi Şekil 4.1'de gösterilmiştir ve temelde nöronu 3 bölgeye ayırmak mümkündür.



Şekil 4.1 Biyolojik sinir hücresinin yapısı

- Soma
- Akson
- Dendrit

Bu değişik bölgelerin her biri bilgi alış verişinde belirli bir rol oynamaktadır. Soma olarak adlandırılan hücre gövdesi hücreyi denetler ve hücre etkinliklerinin tümünden sorumludur. Hücre gövdesinde iki tür uzantı çıkmaktadır: Dendritler ve aksonlar.

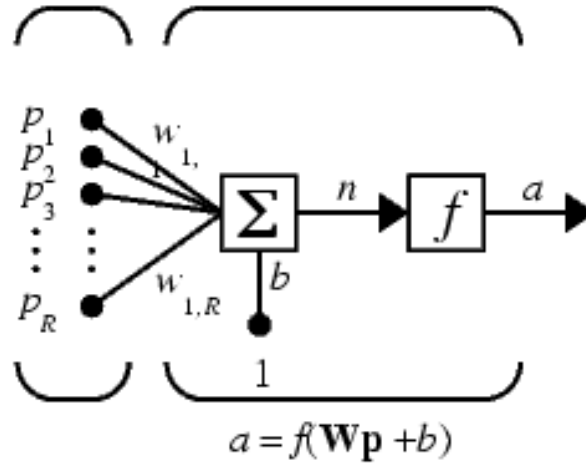
Dendritler, hücre duvarından dışarı çıkan ağaç dalları şeklindedirler. Dendritleri bilgiyi iletim hatları olarak kullanılan uzun fiberler oluşmuş aksonlar

boyunca diğer nöronlardan alır ve hücre gövdesine taşırlar. Dendritlerde yoğun bir metabolizma ve protein sentezi vardır. Aksonlar ise bilgiyi gövdedeki bilgiyi diğer nöronlardan dendritlere taşımakla görevlidirler. İki nöron birbirlerine sinapslar yardımı ile bağlıdırlar. Biyolojik olarak anlatılan bu hücre birimleri mikro metrelerle ölçülen boyuttadırlar. Sinirler arasında sinyal elektro-kimyasal bir düzenekte aktarılmaktadırlar ve elektriksel bir dalga hareketi kimyasal tepkime ile sıra sıra taşınır. Bu taşıma tabi olarak ışık hızında değildir. Beyin mesajlarının ortalama hızının saatte 580 km'ye çıkabildiğini söyleyen yazılar vardır (Nabiyev 2005). Öyleyse beyni bilgisayarlardan daha hızlı kılan ya da cazip hale getiren nedir? Bunun cevabı nöronların fazlalığıdır. İnsan beyninde hesaplamada yaklaşık olarak 10^{11} adet nöron kullanılmaktadır. Bu kadar sayıdaki nöronun paralel şekilde çalışması beynin hızını inanılmaz ölçüde arttırmaktadır.

4.1.3 Yapay sinir ağlarının yapısı

Nöronlar arasındaki sinaptik bağlantılar mesajların taşındığı yer olduğu için son derece önemlidir. Bir nöron belirli bir eşiği aştığında tetiklenir ve nöron boşalır. Alınan ateşleme uyarısı akson boyunca birçok terminallere iletilir. Son derece basitleştirilmiş sanal nöronun yapısı Şekil 4.2'de gösterilmiştir. İlk nöronun biçimsel tanımı Mc Culloch-Pits tarafından gösterilmiştir.

Bütün sinir ağı modellerinde p_i giriş değerleri w_i sinir eklemi ağırlıklandırma katsayıları ile çarpılarak toplanır. Toplanan girişler nöron üzerindeki bir bias değeri ile de toplanır. Toplanan değerler nöronun eşik seviyesini simüle eden bir transfer fonksiyonundan sonra net denilen çıkışı verir. Bir netin çıkışını Eşitlik 4.1'in yardımı ile hesaplanır.



Şekil 4.2 Sentetik nöronun modelini gösteren resim (MATLAB 7.0 Help dosyasından alınmıştır)

$$net = y = f(t_{bias} + \sum_{i=1}^n p_i w_i) \quad (4.1)$$

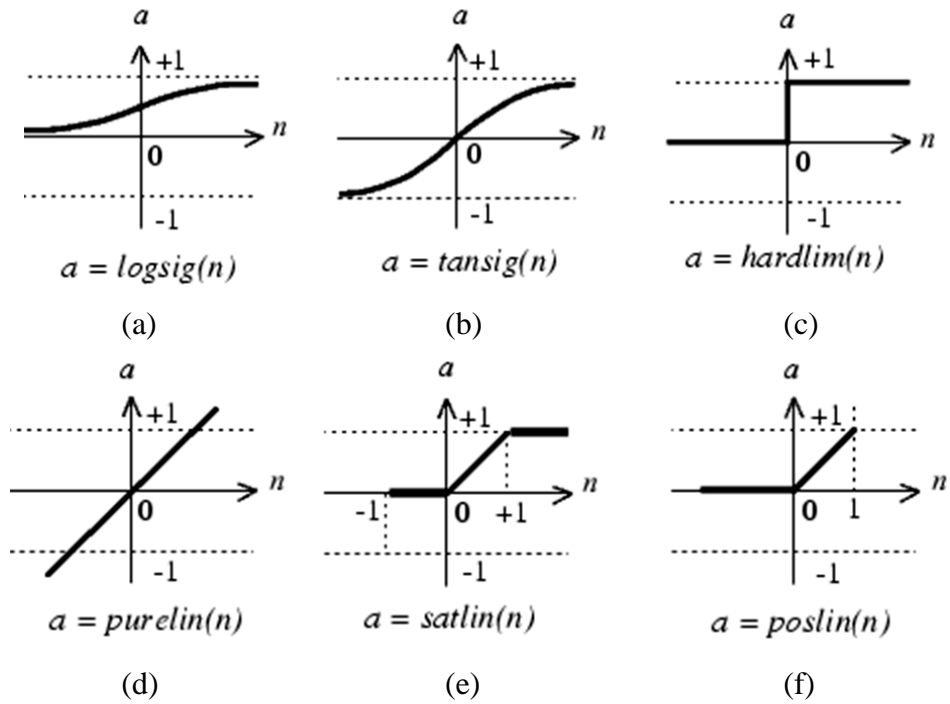
Burada p ve w 'ler uygun girişlerin ve ağırlıkların vektörel ifadesi olarak kullanılmaktadır. R ise giriş sayısıdır. Bu ağırlıklandırılmış girişlerin toplamı sentetik nöron a 'nın çıkış değerini verir. Netin değeri ikilik şekilde ifade edileceği Eşitlik 4.2'de ifade edildiği gibi yalnızca iki konumu olan bir fonksiyon olur. Belirlenen eşik seviyesine göre çıkış yalnızca iki konumdan birinin değerini alır. Değer eşğin altında kaldığında çıkış 0 , eşğin üstünde ise çıkış 1 olur.

$$y = \begin{cases} 1, & \sum p_i w_i \geq T \\ 0, & \sum p_i w_i < T \end{cases} \quad (4.2)$$

Burada eşik seviyesi T 'dir. Bu transfer fonksiyonuna basamak (step) transfer fonksiyonu denilir. Gerçek uygulamalar ve problemler yalnızca ikilik sayısal sistemlerinden oluşmaz ve bu basamak fonksiyonu da gerçel sayılarda işe yaramayacaktır. Bu yüzden farklı transfer fonksiyonları literatüre yer edinmiştir. Bunlardan en yaygınlarına ait transfer fonksiyonları Şekil 4.3'de gösterilmiştir.

YSA içeriğindeki her bir net bu transfer fonksiyonlarından birini içerir. Ancak tüm yapı içerisinde farklı transfer fonksiyonları kullanılabilir. Örneğin girişte *logaritmik sigmoid* çıkışta *liner transfer fonksiyonu* tercih edilebilir. Genel yaklaşım dijital veri işlemek için basamak transfer fonksiyonu tercih edilmekle birlikte gerçel sayılarda *logaritmik sigmoid* fonksiyon oldukça yaygındır. Logaritmik fonksiyonun matematiksel ifadesi aşağıda Eşitlik 4.3’de gösterildiği gibidir.

$$\log \text{sig}(n) = 1/(1 + e^{-n}) \quad (4.3)$$

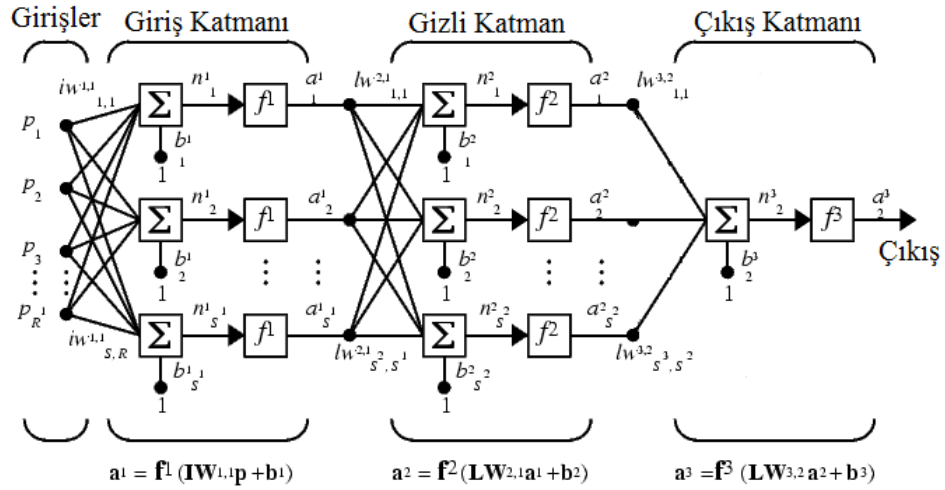


Şekil 4.3 YSA transfer fonksiyonlarına ait yaygın transfer fonksiyonları (a) logaritmik sigmoid transfer fonksiyonu (b) tanjant sigmoid transfer fonksiyonu (c) basamak transfer fonksiyonu (d) liner transfer fonksiyonu (e) saturasyona uğratılmış liner transfer fonksiyonu (f) pozitif transfer fonksiyonu

Bir YSA yalnızca bir nöron kullanmaz. Pek çok katmanı da bir arada içerebilir. YSA’nın bu uygulanaşına çok katmanlı algılama (Multi Layer Perceptrons, MLP) denir. Bir MLP bir giriş katmanı, çıkış katmanı ve giriş ile çıkış arasında kalan en az bir gizli katmandan oluşur. Ancak performansın sonucuna ve problemin

zorluğuna bağlı olarak gizli katman sayısı değiştirilebilir. Şekil 4.4'de bir MLP yapısı gösterilmiştir.

YSA her adımda ileriye doğru giden bir yapısı vardır. Genel anlamda geriye dönülüp bakılmaz bir değerlendirme yapıldı mı artık geriye dönüş yoktur ve bu değer bir sonraki net için devamını oluşturan bir giriş haline dönüşmüş olur. Bu yüzden YSA tarifinde yapısına ileri yön beslemeli yapıda dendiği olur. Yalnızca ileriye giderken geriye doğru ağırlık katsayıları nasıl ayarlanabilir? Bu soruya cevap verebilecek farklı metot ve öğrenme metotları geliştirilmiştir. Bunlardan iki tanesi olan Geriye Yayılım (backpropagation) algoritması ve Levenberg-Marquart öğrenme metodu en yaygın olanlardır.

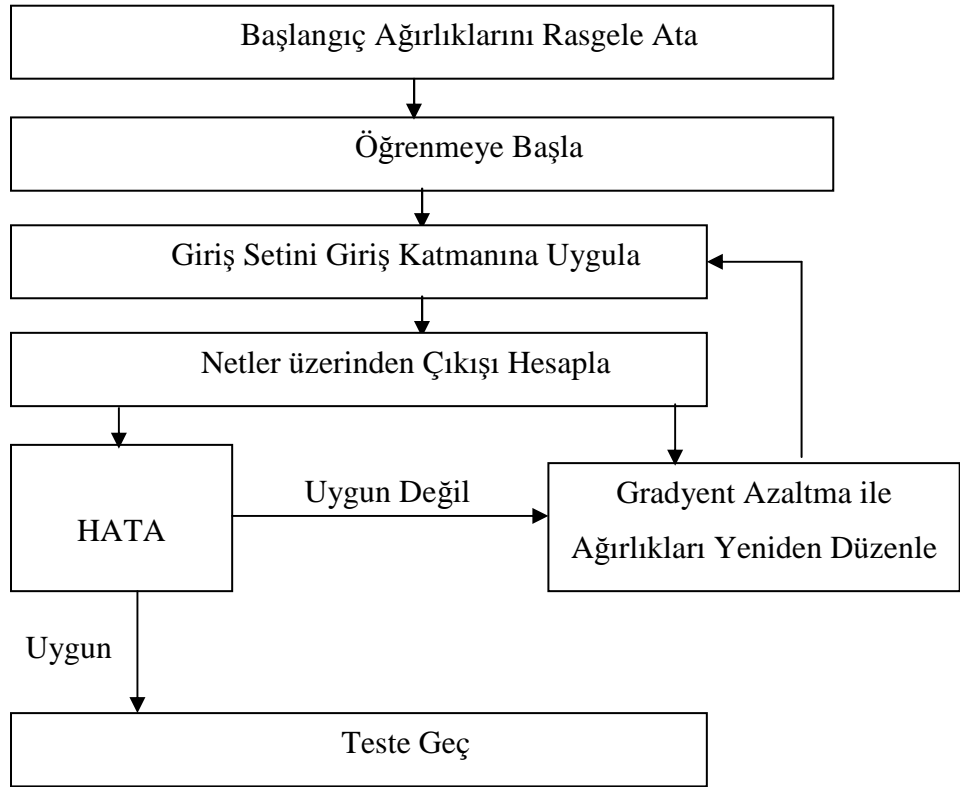


Şekil 4.4 MLP yapısının çalışmasını gösteren bağlantı diyagramı (MATLAB 7.0 Help dökümanından alınmıştır)

Bir YSA yapısının performans değerlerinin tatminkâr olabilmesi için çoklu katmanlı bir yapı tercih edilebilir. Ancak buraya kadar olan YSA'nın belkide kolay tarafıdır. Şimdiki sorun doğru değeri verebilecek ideal ağırlıkların ve bias değerlerinin nasıl hesaplanması gerektiği ve bunların çözümlerinin nasıl olacağı konusundadır.

4.1.4 Geriye yayılım algoritması

Uygulamalarda en yaygın kullanılan öğretim algoritmasıdır. Anlaşılması kolay ve matematiksel olarak kolay ispatlanabilir olmasından dolayı oldukça popülerdir (Rumelhart 1986a, Rumelhart 1986b). Bu algoritma hatayı çıkıştan girişe doğru azaltarak geldiğinden “Geriye Yayılım” adını almıştır. MLP eğitimlerinde kullanılan en temel algoritmalarından biridir. Bu algoritmanın akış şeması Şekil 4.5’de verilmiştir.



Şekil 4.5 Bir YSA ağında geriye yayılım algoritmasının uygulanması gösteren akış şeması

Eğitme işlemi ve eğitmeden sonraki test işlemi bu akışa göre yapılır. Bu algoritma ile i ve j kat işlem elemanları arasındaki ağırlıklardan $\Delta w_{ij}(t)$ değişikliği Eşitlik 4.4’deki gibi hesaplanır.

$$\Delta w_{ij}(t) = \eta \delta_i j_i + \alpha \Delta w_{ji}(t-1) \quad (4.4)$$

Eşitlikte η öğrenme katsayısı, α momentum katsayısı ve δ_j ara veya çıkış katmanındaki bir j nöronuna ait faktördür.

Çıkış katı için bu faktör Eşitlik 4.5'deki gibi ifade edilebilir.

$$\delta_j = \frac{\partial f}{\partial net_j} (y_j^t - y_j) \quad (4.5)$$

Burada $net_j = \sum x_j w_{ji}$ ve $y_j^{(t)}$ ise j işlemci elemanının hedef çıkışıdır. Ara katlardaki nöronlar için Eşitlik 4.6'daki gibi ifade edilebilir.

$$\delta_j = \left[\frac{\partial f}{\partial net_j} \right] \sum w_{qi} \delta_q \quad (4.6)$$

Ara katlardaki nöronlar için herhangi bir hedef çıkış olmadığından Eşitlik 4.5 yerine Eşitlik 4.6 kullanılır. Bu durumda çıkış katından başlayarak δ_j faktörü, bütün katlardaki nöronlar için hesaplanır. Daha sonra Eşitlik 4.4'deki denklemden ağırlıklar yeniden oluşturulur. Bu bir kademeli olarak azalma algoritması olup, MLP yapılarını eğitmede kullanılan bir algoritmadır. Temelde istenilen çıkış ile ağ çıkışı arasındaki farkın ağırlıklara bağlı olarak düşürülmesi prensibine dayanır. Başarılı ve hızlı bir eğitim süreci için η ve α değerlerinin iyi ayarlanması gerekmektedir.

4.1.5 Levenberg-Marquardt öğrenme metodu

Temel olarak bu algoritma maksimum komşuluk fikri üzerine kurulmuş en az kareler hesaplama metodudur (Levenberg 1944). Bu algoritma gradyent arama

metotlarından Gauss Newton ve Steepest-Descent algoritmalarının en iyi özelliklerinden oluşur. Genel olarak bu metot yavaş yakınsama problemi yaşamaz. Bunun en büyük nedeni bu metot Hessian matrisini kullanmadan 2. derece türevlendirilmiş değerlerden ağırlıkları hesaplar. Dolayısıyla levenberg-Marquardt metodu oldukça hızlı yakınsayan bir metot olarak karşımıza çıkar. Bununla birlikte rasgele atanan ilk başlangıç sayıları önemlidir.

$E(w)$ 'nin bir amaca yönelik hata değeri olduğunu varsayarsak m tane veri Eşitlik 4.7'de verilir;

$$E(w) = \sum_{i=1}^m e_i^2 = \|f(w)\|^2 \quad (4.7)$$

Bu eşitlikte $e_i^2(w) \equiv (y_i - yd_i)^2$ dir. Burada amaç fonksiyonu $f(.)$ ve onun jakobiyeni J 'nin bir noktada w bilindiği kabul edilir. LM metodunda hedef, parametre vektörü w 'nin $E(w)$ minimum iken bulunmasıdır. Yeni vektör $w_{(k+1)}$, Eşitlik 4.8'de de gösterildiği gibi kabul edilen vektör w_k 'dan hesaplanır.

$$w_{k+1} = w_k + \delta w_k \quad (4.8)$$

dır. burada δw_k aşağıdaki Eşitlik 4.9 da verilmiştir.

$$(J_k^T J_k + \lambda I) \delta w_k = -J_k^T f(w_k) \quad (4.9)$$

Bu eşitlikte;

J_k : f in w_k değerlerilmiş Jakobyeni

λ : Marquardt parametresi,

I : birim ya da tanımlama matrisi

dir.

Levenberg-Marquardt metodu aşağıdaki aşağıdaki yönergeyi kullanır.

- (i) $E(w_k)$ 'yı hesapla
- (ii) Küçük bir λ değeri ile başla (0.01 gibi)
- (iii) δw_k için Eşitlik 4.26'yı çöz ve $\delta w_{k+1} = w_k + \delta w_k$ değerini hesapla
- (iv) Şayet $E(w_k + \delta w_k) \geq E(w_k)$ ise λ 'yı 10 kat arttır ve iii'ye git
- (v) Şayet $E(w_k + \delta w_k) < E(w_k)$ ise λ 'yı 10 kat azalt.
- (vi) $w_k : w_k \leftarrow w_k + \delta w_k$ 'yı güncelleştir ve adım iii'ye git.

Hedef çıkışı hesaplamak için bir MLP'nin LM metodu ile eğitilmesi w_0 değerlerinin atanması ile başlar. Bu atama ilk anda rastgele değerlerden oluşur. Daha sonra hatalarının karelerinin toplamını sıfır yapmak için metot uğraşır. Bu adımlama yukarıda bahsedildiği şekilde sürer.

4.2 Adaptif Ağ Tabanlı Bulanık Çıkarım Sistemi (ANFIS)

Bulanık Mantık ve Sinir ağları yapısı birbirini doğal olarak tamamlayan birinin eksiği diğerinde olan yapay zeka sistemleridir. Sinir ağları düşük seviyeli hesaplama yapılarında iyi performans verirken, ham verilerde bulanık mantık dilsel yorumlarla daha yüksek seviyeli yorumlar yapabilir. Ancak bulanık mantık sistem yeni çevre şartları için kendini yenileyemez ve öğrenme yetisi de yoktur. Diğer taraftan sinir ağlarının öğrenme kabiliyeti vardır. Bu iki yapının birleştirilmesi bulanık mantık çıkarım sistemine öğrenme kabiliyeti eklemiş ve bulanık mantık çıkarım sistemini daha etkili hale getirmiştir. Ortaya çıkan yapı sadece öğrenme kabiliyeti olarak açıklamak yetersiz kalır. Buna ilaveten bu birleşme sonucu bulanık mantık parametrelerinin optimizasyonunda da gelişme olmuş, en küçük kareler metodu ve gradyent metotlarının kazanımı ile bulanık mantık sinir ağları metodundan daha hızlı yakınsayan bir metot haline dönüşmüştür.

Öncelikle ANFIS'in ne olduğunu tanımlayabilmek için sinir ağı bulanık mantık (Neuro-Fuzzy)'nin ne olduğuna bakmakta fayda vardır. Bir nero-fuzzy sistem, öğrenme parametreleri, bulanık mantık (fuzzy) setlerinin ve bulanık mantık kurallarının elde edilmesinde yapay sinir ağıları teoreminden esinlenilmiş bir çıkarım sistemini kullanır. Modern Neuro-Fuzzy sistemler genellikle özel yapılı ileri beslemeli çok katmanlı yapay sinir ağıları yapısıdır (Lin 1996).

4.2.1 ANFIS'in Yapısı

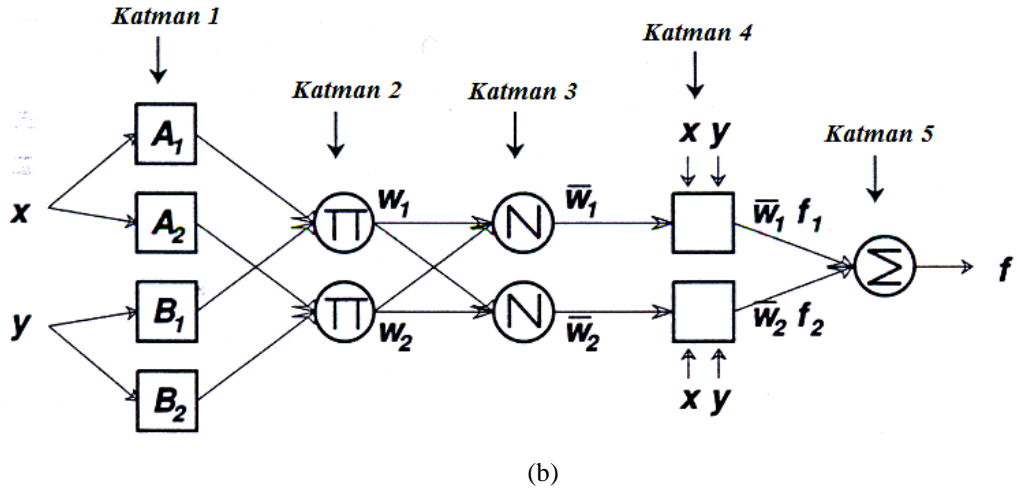
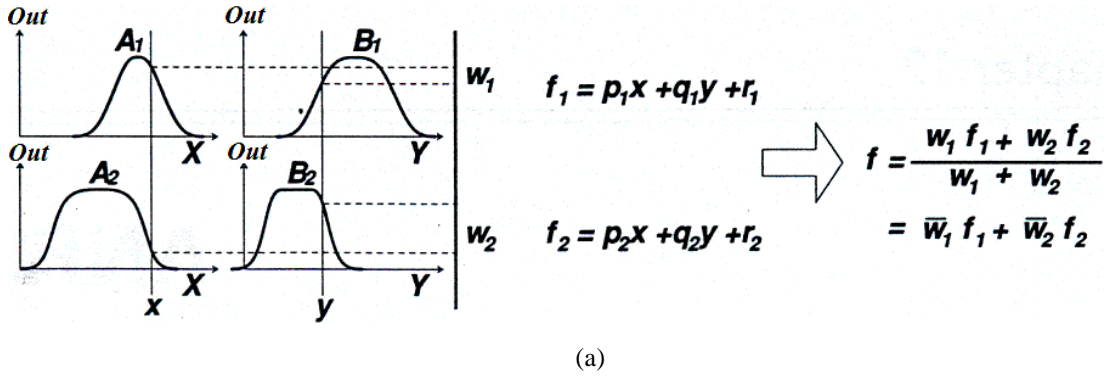
ANFIS 1993'lü yıllarda Jang tarafından ortaya atılmış yapay sinir ağıları gibi bir network yapısı olan ve kendi parametrelerini YSA algoritmalarındaki gibi güncelleyen eden bir Fuzzy yapısıdır (Jang 1993, Buckley 1994). ANFIS'in yapısı da YSA gibi ileri beslemeli bir yapıdır. Üretilen her çıktı geriye bakılmaksızın bir sonraki katman için işlenmesi gereken yeni bir veriyi oluşturur. Genel olarak bir ağ tabanlı bulanık mantık çıkarım sistemi bir giriş, bir çıkış ve 3 gizli katman olmak üzere 5 katmandan oluşur. 3 gizli katman, üyelik fonksiyonlarını ve bulanık mantık kurallarını içerir.

Bulanık mantık ilk olarak 1950'li li yıllarda Lotfi Z. Tarafından, çözüm için dilsel ifadeleri kullanan bir yapay zeka tekniği olarak ortaya atılmıştır. Temel yapı aynı kalmakla birlikte dilsel ifadelere olan bakış açıları farklılığı bulanık mantığında türevlerinin oluşmasının önünü açmıştır. Mamdani ve Sugeno modeli olmak üzere iki adet berraklaştırma (defuzzication) modeli vardır (Wang 1996). Ebrahim Mamdani'nin önermiş olduğu geri çözümleme sisteminde hem giriş hemde çıkış üyelik fonksiyonları fuzzy setlerinden oluşur (Mamdani 1975). Diğer yöntem ise TSK (Tagaki-Sugeno-Kanga) metodu olarak bilinir ve bu berraklaştırma yönteminde, çıkış fonksiyonu liner değişkenlerden ya da sabit sayılardan oluşur. Bu yönüyle Sugeno modeli Mamdaninin önerisinden daha basit ve etkili kalmaktadır (Sugeno 1985). Sugeno, Takagi gibi önerenlerin isimleri ile bulanık mantık çalışmaları anılır olmuştur. Fuzzy'nin temelindeki dilsel ifadelerin kullanımı ANFIS'inde temelidir. Ancak bu dilsel ifadelerin optimizasyonu bir network ağı ile yapılmakta bu da

ANFIS'in işlem gücüne paralel veri işleme kabiliyeti ile beraber öğrenme yetisi katmaktadır. ANFIS uygulamalarında Sugeno en yaygın tercih edilen bulanık mantık modelidir. Sugeno ile Tsukama metotlarında ANFIS'in katman sayıları arasında bir fark yoktur. Yalnızca Tsukama modeli katman 2'den katman 4'e ileri yönde bir ilave besleme yapar.

ANFIS yapısını basitleştirmek için Şekil 4.6'da iki girişli bir Sugeno modeli ve onunla ilgili olan ANFIS yapısı ve çalışması verilmiştir.

Katman 1: Bu katmandaki her bir düğüm, düğüm fonksiyonu x veya y 'nin A_i ve B_i ile ifade edilen dilsel ifadelerinin adaptif bir şeklini oluşturur. Bu katmanın çıkışı her bir i değeri için Eşitlik 4.10'dan hesaplanır. Bu katman bulanıklaştırma katmanıdır.



Şekil 4.6 İki girişli ANFIS yapısını gösteren temsili resim (a) üyelik fonksiyonlarının (b) iki girişli sugeno yapıları-ANFIS modelinin gösterimi

Bu katmandaki düğümler giriş ifadesini bulanıklaştırır. Jang'ın modelinde bulanıklaştırma fonksiyonu olarak çan eğrisi fonksiyonu kullanılmıştır. Birinci katmana ait çıkış fonksiyonu Eşitlik 4.10'da gösterilmiştir.

$$\begin{aligned} O_{1,i} &= \mu_{A_i}(x), & i &= 1, 2, \dots \\ O_{1,i} &= \mu_{B_{i-2}}(y), & i &= 3, 4, \dots \end{aligned} \quad (4.10)$$

$O_{1,i}$, bulanık üyelik fonksiyonu olan A ve B'ye ait ilişkililik derecesi gösteren bir belirteçtir. Bu ifade üyelik fonksiyonu tipinden hesap edilir. Örneğin Gauss ya da trapezoidal ya da üçgen bir transfer özdeşliği olan üyelik fonksiyonuna göre ifade edilecek olan $O_{1,i}$ değerleri farklı olacaktır. Örneğin çan yapısındaki dilsel bir ifade için μ_A değeri Eşitlik 4.11'den hesaplanabilir.

$$\mu_A = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2b}} \quad (4.11)$$

Burada a_i , b ve c_i set parametreleridir. Bu parametrelerin değişmesi çan fonksiyonunun da değişmesine neden olur. Şekil 4.7'deki Sugeno modeli için eğer $x=3$ ve $y=4$ olsun ve başlangıç parametreleri olarak tanımlanan a_i , b ve c_i parametrelerinin tanımlandığını varsayalım. Bu Eşitlik 4.11'den; katman 1'de oluşan çıkış değerleri Şekil 4.7'deki gibi oluşacaktır.

$$\begin{aligned} x & \begin{cases} A_1 \rightarrow \mu_{A_1}(x) = \mu_{A_1}(3) = \frac{1}{1 + \left| \frac{3-1}{2} \right|^2} = 0.5 \\ A_2 \rightarrow \mu_{A_2}(x) = \mu_{A_2}(3) = \frac{1}{1 + \left| \frac{3-9}{2} \right|^2} = 0.1 \end{cases} \\ y & \begin{cases} B_1 \rightarrow \mu_{B_1}(y) = \mu_{B_1}(4) = \frac{1}{1 + \left| \frac{4-2}{2} \right|^2} = 0.5 \\ B_2 \rightarrow \mu_{B_2}(y) = \mu_{B_2}(4) = \frac{1}{1 + \left| \frac{4-14}{2} \right|^2} = 0.25 \end{cases} \end{aligned}$$

Şekil 4.7 Katman 1'e ait iki girişli Sugeno yapıları ANFIS'in hesaplanmasını gösterimi

Katman 2: Bu katmandaki her bir düğüm Π ile gösterilen sabit düğümlerdir. Bu düğümün çıkışları kendisine gelen tüm sinyallerden türetilir. Eşitlik 4.12 bu sinyallerin nasıl türetildiğini göstermektedir.

$$O_{2,i} = \omega_i = \mu_{A_i}(x) \mu_{B_{i-2}}(y), \quad i = 1, 2, \dots \quad (4.12)$$

Her bir çıkış düğümü kendisinden sonraki katmanda bulunan düğümdeki kuralın tetiklenmesi ile ilgili gerilmeyi etkiler. Genel olarak her bir T-norm operatörü bulanık sistemindeki VE operatörüne karşılık geldiği söylenilebilir. Katman 1’de sunulan örneği devam ettirirsek; Katman 2’ye ait ağırlık değerleri Eşitlik 4.12’den hesaplanırsa,

$$O_{2,1} = \omega_1 = \mu_{A_1}(x) \mu_{B_1}(y) = 0.5 \times 0.5 = 0.25 \quad i = 1$$

$$O_{2,2} = \omega_2 = \mu_{A_2}(x) \mu_{B_2}(y) = 0.1 \times 0.25 = 0.025 \quad i = 2$$

Üretilen bu değerler ω_i ’ye ait ağırlıkları oluşturur.

Katman 3: Bu katmanda bir önceki katmanda oluşan ağırlıklar kendi içlerinde normalize edilirler. Normalize işlemi, i .inci ağırlığın toplam ağırlığa oranı şeklinde Eşitlik 4.13’de gösterildiği gibi hesaplanır.

$$O_{3,i} = \varpi_i = \frac{\omega_i}{\omega_1 + \omega_2} \quad i = 1, 2 \quad (4.13)$$

Burada ϖ_i ağırlıklandırılmış yeni değerdir. Bu uyumlandırmanın sonunda bu katmanın çıkışı ağırlıklandırılmış katsayılarla düzeltilmiş şekliyle yeninden elde edilir.

Bu katmanla da ilgili olarak örneğimizi devam ettirirsek Eşitlik 4.13’den ağırlıklar,

$O_{3,1} = \overline{w}_1 = \frac{0.25}{0.25 + 0.025} = 0.91$ ve $O_{3,2} = \overline{w}_2 = \frac{0.025}{0.25 + 0.025} = 0.09$ olarak elde edilirler.

Katman 4: Bu katmanda daha önceden belirlenen fonksiyonlarla üretilen ağırlıklar kullanılarak son katman öncesi çıkışlar elde edilirler. ANFIS için kullanılan fonksiyonlar burada devreye girer ve ağırlıklarla beraber yeni bir değer hesaplanır. Hesaplanan değere ait Eşitlik 4.14’de verilmiştir.

$$f_i = p_i x_i + q_i y_i + r_i \quad (4.14)$$

Burada p ve q sabit katsayılarıdır. Katman 4’e ait çıkışlar Eşitlik 4.15’deki şekliyle ifade edilebilirler.

$$O_{4,i} = \overline{w}_i f_i \quad (4.15)$$

Katman 4’ün sonunda oluşan çıkışlar normalize edilmiş ağırlıklı çarpanların ilgili fonksiyon ile çarpılması sonucu hesaplanır. Fonksiyon $f = ax + b$ şeklinde tanımlı ise bu ANFIS yapısına 1. dereceden Sugeno tanımlı denir.

Yukarıda anlatılan örneğe devam edersek;

1. dereceden tanımlı fonksiyonlar aşağıdaki gibi olsun

$$f_1 = 0.1x + 0.1y + 0.1$$

$$f_2 = 10x + 10y + 10$$

Bu durumda yukarıda açıklanmaya devam edilen örnek için katman 4 çıkışlarını aşağıda gösterildiği gibi hesaplanabilir.

$$O_{4,1} = \varpi_1 f_1 = 0.91 \times (0.1 \times 3 + 0.1 \times 4 + 0.1) \Rightarrow O_{4,1} = 0.73$$

$$O_{4,2} = \varpi_2 f_2 = 0.09 \times (10 \times 3 + 10 \times 4 + 10) \Rightarrow O_{4,2} = 4.5$$

Bu katmanda yapılan işlemi berraklaştırma (de-fuzzication) olarak da tanımlamak mümkündür.

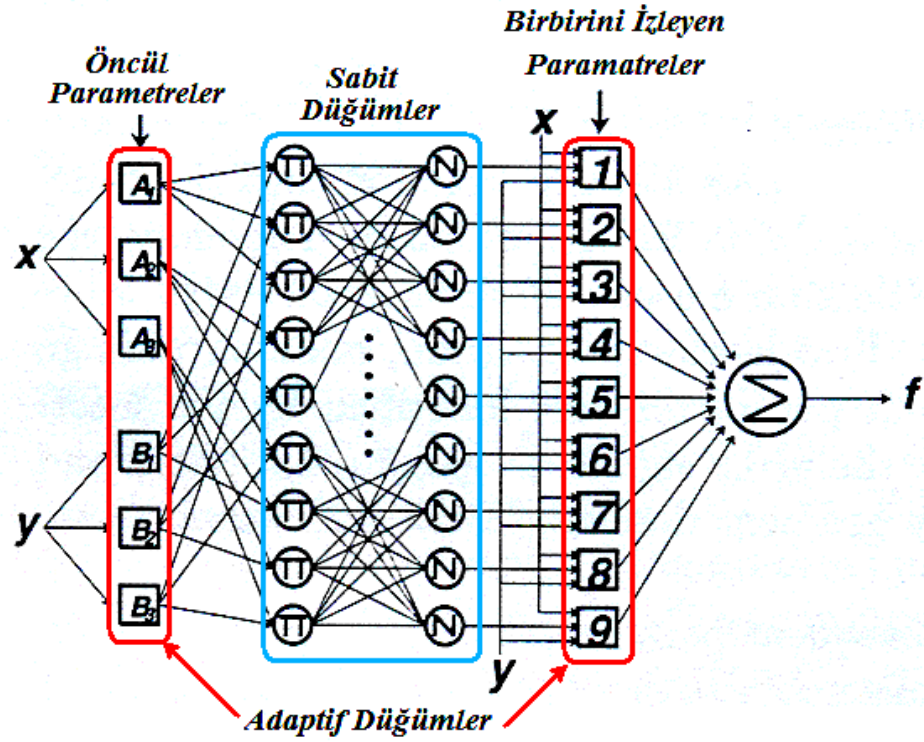
Katman 5: Bu katman yalnızca toplam sembolünü içeren tek düğümlü bir katmandır. Bu girişe gelen tüm değerler bir toplama işlemine tabi tutularak berraklaştırılan değerlerden ANFIS sonucu hesaplanır. Çıkışa ait değer Eşitlik 4.16'dan hesaplanabilir.

$$O_5 = \sum_{i=1}^n O_{4,i} \quad (4.16)$$

daha açık hali ile ifade edersek Eşitlik 4.16, Eşitlik 4.17'ye dönüşür.

$$O_5 = \sum_{i=1}^n O_{4,i} = \sum_{i=1}^n \varpi_i f_i \quad (4.17)$$

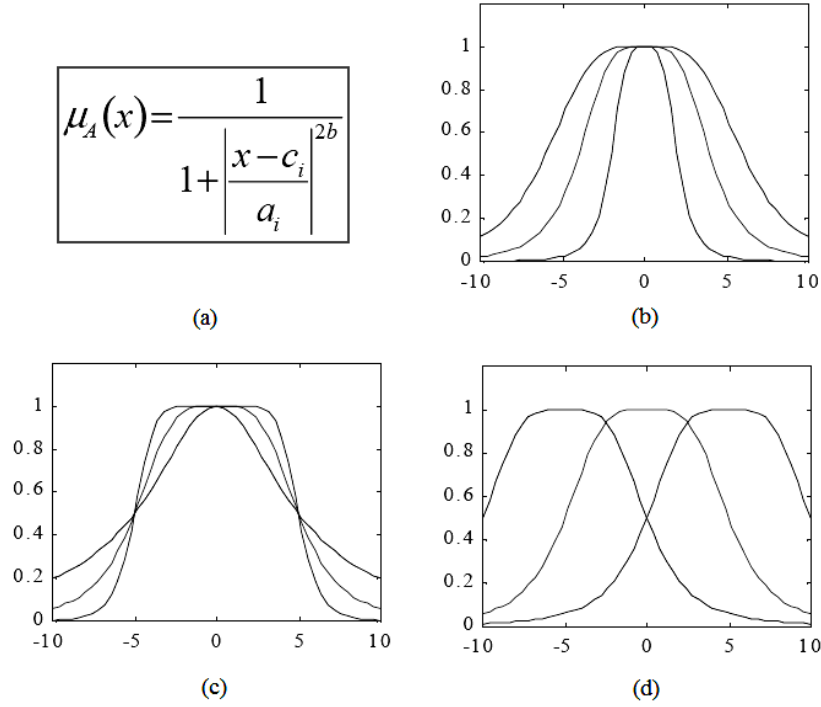
Anlatılan örneğin katman 5 için değeri hesaplanırsa; $O_5 = 4.5 + 0.73 = 5.23$ değeri elde edilir. Buraya kadar anlatılanlar ANFIS'in çıkış değerini nasıl ürettiği ile ilgilidir. Ancak ANFIS'i asıl cazip kılan bu değerleri hesaplayabilmesi değil bu üyelik ve (burada anlatılan şekli ile birinci dereceden Sugeno) fonksiyonlara ait parametrelerin oldukça hızlı bir şekilde hesaplayabilmesidir. ANFIS de YSA gibi ileri yön beslemeli bir ağ yapısına sahiptir. ANFIS yapısı eğitildiği zaman, A ve B bulanık kural yapısında olan fonksiyonların Şekil 4.8'den de görüleceği üzere adaptif olan düğümlerin gerekli parametreleri değişir.



Şekil 4.8 ANFIS yapısı içerisindeki adaptif düğümleri gösteren temsili şekil

Örneğin bulanık mantık üyelik fonksiyonları olarak çan eğrisi üyelik fonksiyonu seçilmiş olsun. Bu fonksiyon Eşitlik 4.11’de verilmiştir. Çan eğrisi parametrelerinin optimizasyonu sonunda, parametre değişimine göre üyelik fonksiyonunun değişimi Şekil 4.9’da gösterilmiştir.

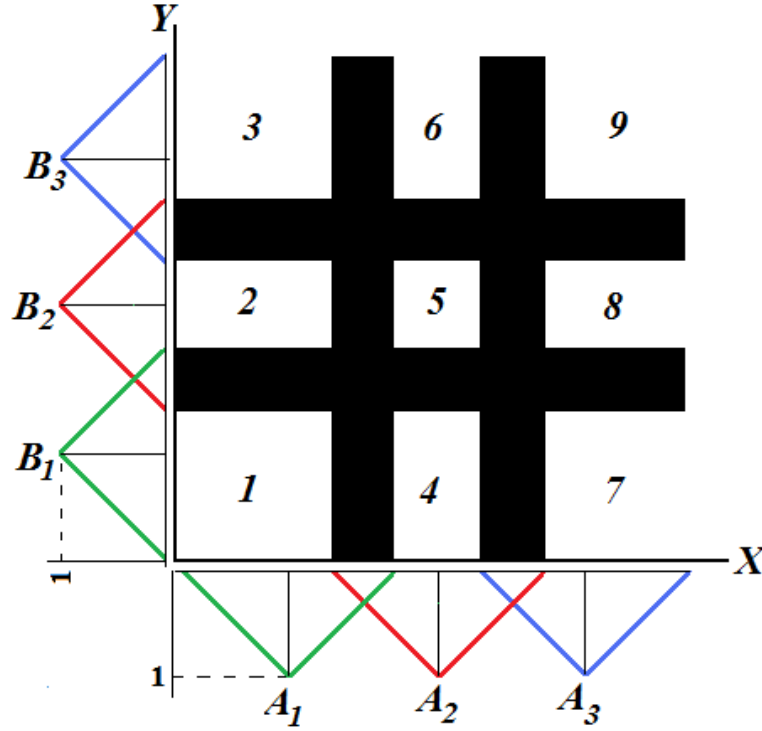
Parametrelerden a ’nın değişimi eğrinin x ekseninde kapsadığı alanı değiştirirken, b parametresi üyelik fonksiyonunu daha dik hale getirmektedir. Üçüncü parametre c ise üyelik fonksiyonunun ötelenmesini sağlamaktadır. Bu üç parametrenin optimizasyonu sonunda üyelik fonksiyonu, doğru çıktının üretilmesi adına en uygun pozisyona gelmiş olur.



Şekil 4.9 Dilsel ifadelere ait parametrelerin değişimini gösteren şekil (a) çan eğrisi üyelik fonksiyonuna ait parametreleri gösteren eşitlik (b) parametre a 'nin üyelik fonksiyonuna etkisi (c) parametre b 'nin üyelik fonksiyonuna etkisi (d) parametre c 'nin üyelik fonksiyonuna etkisi

Bu optimizasyon sonunda oluşan değerlerden A ve B dilsel ifadelerini bir eksen üzerinde gösterirsek Şekil 4.10'u elde ederiz. Bu uzay gösterimdeki desene ızgara (Grid) bölümlendirme denir. Farklı veri yapıları için geliştirilmiş ağaç bölümlendirme ve dağınık bölümlendirme olarak gösterilen iki farklı yapısı daha vardır.

ANFIS'in çalışması boyunca adaptif olarak değiştirebileceği bir diğer parametre seti de şartlı kurallara bağlı olunan fonksiyona ait parametre değerleridir. Bu ayarlanması gereken parametreler çıkıştan bir önceki katmanda yer alan ve fonksiyonun katsayılarıdır. Yukarıda bahsedilen giriş uzayına ait parametre tahmini Öncül Parametreler olarak kabul edilirler. Bununla birlikte 4. katmanda anlatılan parametrelerin de optimizasyonu gereklidir.



Şekil 4.10 Giriş uzayındaki Grid bölümlendirmesini anlatan resim

Bu şart fonksiyonunun optimize edilmesinde ileriye doğru en küçük kareler metodu kullanılarak yapılır. ANFIS yapısının ileriye ve geriye doğru hangi parametreleri optimize ettiği Tablo 4.1’de verilmiştir.

Tablo 4.1 Adaptif parametrelerin nasıl optimize edildiğini gösteren tablo

	İleriye Doğru	Geriye Doğru
Öncül Parametreler	Sabit	Gradyent
Sıralı Parametreler	En Küçük Kareler tahmin	Sabit
Sinyaller	Düğüm Çıkışları	Hata Sinyali

ANFIS bir defa yürütüldüğünde ileriye doğru belirlenen giriş uzayı için en doğru çıkışı verecek fonksiyonu yakalamaya çalışır. Bununla birlikte oluşan hataya bakarak da Gradyentinden giriş uzayını normalize etmeye çalışır. Gradyent hesaplamalarda fonksiyon parametrelerine dokunulmaz ancak giriş uzayı değiştikten sonra ileriye doğru hesap başladığı zaman optimize olmuş giriş uzayı desenine göre en az hata ile fonksiyona ait sıralı parametreleri tekrar hesaplayarak çıkışı yakalamaya çalışır. Sıralı parametrelerin optimizasyonunda düğüm çıkışları

kullanılırken giriş uzayının optimizasyonunda ANFIS'in çıkışına ait hatadan yararlanılır. Böylece bir turda hem üyelik hem de fonksiyonu parametrelerine ait bir düzeltme oluşur. Hem gradyentin hem de en küçük kareler metodunun bir arada kullanılması algoritmasına hibrit öğrenme denilmektedir.

Bu hibritlik ANFIS'in hızını kayda değer şekilde arttırmaktadır. YSA'nın Levenberg-Marquardt öğrenmeli modelinde binli iterasyonlarla bahsedilen öğrenme iterasyonlarından bahsedilirken bu rakam ANFIS için yüzölçü sayılarla ifade edilebilir. Bununla birlikte YSA her başlangıç parametresi için ayrı bir ağırlıklar dengesi bulur oysa ANFIS her zaman aynı hata değerini bulur. YSA'nın Levenberg-Marquardt öğrenme modelinde karşılaşılan başlangıç parametrelerinin önemini ANFIS'de yoktur. Her iterasyonundaki toplam iki düzeltme ANFIS'in daha hızlı yakınsamasına olanak tanır.

5. VERİ SETLERİ

Literatürde kullanılan verileri farklı sınıflara ayırmak mümkündür. University of California, Irvine (UCI) gibi açık veri kaynaklarının bulunduğu bilimsel kabul görmüş on-line siteler veri kaynaklarını; tahmin, sınıflama olmak üzere ikiye ayırmaktadır. Tahmin ve sınıflama veri türlerinin herhangi bir disipline has olduğunu söylemek imkânsızdır. Her iki tür veriye de her alanda rastlamak mümkündür. Örneğin girilen veriler ışığında iyi-kötü ya da eski-yeni ayrımı yapan bir sistem sınıflama amaçlı kullanılmış olur. Bununla birlikte ileriye dönük nüfus tahmini yapan bir sistem ya da deneysel verilerden elde edilen sonuçlarla denenmeyen kısımları arayan bir yapay zekanın da tahmin amaçlı kullanıldığını söylemek daha doğru olacaktır. Mühendislik uygulamalarında tahmin verilerinin ayrı bir önemi vardır. Örneğin betonun depreme dayanıklılığını teste eden, kimyasal bir tepkimeye ait değerleri zaman bazında kayıt altına alan, eldeki nüfus dağılımı ile 10 yıl sonrası için nüfus tahmin yapan ya da elektrikte akım ile gerilim arasındaki aitin incelenmesi gibi pek çok çalışma noktasal verilerden oluşmasına rağmen aslında sürekli tanımlı bir sistemin anlık ifadeleridir. Bu tür verilere tahmin verileri denir. Ara değerleri ancak tahmin etmek mümkündür. Sonuçta her ihtimali denemek zaman, enerji ve ekonomik kayıp anlamına gelebilir. Örneğin bir helikopter firmasının araç omurgasının sağlamlığı ile ilgili kaza testleri yaptığını düşünelim. Böyle bir oluşumda her deneyde olabildiğince çok verinin alınması ve etkili bir biçimde değerlendirilmesi gerekmektedir. Her geliştirme için olmuş mu olmamış mı değerlendirilmesinin bir helikopterin feda edilmesi ile denemek oldukça pahalı bir metot olacaktır. Bu yüzden eldeki verilerin işaret ettiği sebep sonuç ilişkilerinin çok iyi tahmin edilmesi ve arzulanan yapı konseptine en kısa sürede ulaşılmasının hedeflenmesi gerekmektedir.

Tahmin verilerinin değişken sayısı, içeriğinin sayısal olup olmaması gibi pek çok nedenle ilişkili olarak çözümü çok basit olabilmekle birlikte yüksek dereceden nonlinear çözümlü de olabilir. Örneğin arz ve talep arasında minimum maliyetli

üretim yapmaya çalışan bir firma yöneticisi için hedefler ve kısıtlar arasındaki çözüm genelde doğrusaldır^{5,6}. Bu tür denklemler simplex metodu ile kolayca çözülebilirler (Sungmook 2002, Nabih 1977). Ancak özellikle mühendislik uygulamalarında basit yapılı veri türlerine rastlamak neredeyse imkânsızdır. Bir betonun depreme dayanıklılığının sadece içindeki kimyasal elemanların eşitliğiyle ifade etmek oldukça güçtür ya da bir teldeki ısıнын yayılımını diferansiyel denklem dışında ifade etmek neredeyse imkânsızdır (Loh 1984). Veri içeriklerinin sonuçla olan doğrusal olmayan ilişkileri, veri özelliklerinin fazla olması gibi çözümü zorlaştıran sebeplerin çokluğu yapay zeka ile çözüm aranmasını popüler hale getirmiştir. Mühendislik alanında yapay zeka oldukça yaygın olarak kullanılmaktadır (Van-Tsai 2004, Govindaraj 2007, Gonzlez 2007).

Çalışmada önerilen normalizasyon metodunu hedeflediği veriler, süreksizlik içermeyen tahmin amaçlı veri setleridir. Veri kaynağı olarak bilimsel geçerliliği olan farklı disiplinlere ait veri setleri kullanılmıştır. Verilerin ortak özellikleri sınıflama amacı olmayan tahmin verilerinden oluşmasıdır. Veri kaynağı olarak Selçuk Üniversitesi Çocuk Kardiyoloji Bölümünden alınan kalp verileri, UCI web tabanlı bilimsel geçerliliği olan veri havuzu, Erciyes Üniversitesi, Diş Hekimliği Fakültesi, Prof. Dr. Sadık KARA ve ekibinden alınan diş dataları ve CBA veri tabanında bulunan tıbbi veri setlerinde önerilen normalizasyon metodu yapay zeka öncesi kullanılarak performans kriterlerine eklenmiştir.

Bununla birlikte yalnızca tahmin amaçlı veri setleri değil sınıflama verileri üzerindeki etkisinin görülmesi için literatürdeki sınıflama verileri de eklenmiştir. Sınıflama verileri olarak spect, kalp hastalığı, doppler, VEP ve hepatit veri seti kümeleri kullanılmıştır.

⁵ http://en.wikipedia.org/wiki/Simplex_algorithm

⁶ http://en.wikipedia.org/wiki/Nelder-Mead_method

5.1 Tahmin verileri

Tahmin verileri için, UCI'den alınmış olan beton verileri, Tıp literatüründen dış ve kalp verileri bununla birlikte Çevre verisi olarak orman dataları ve kimya disiplinine ait gerilim ve akım veri setleri kullanılmıştır. Toplam 5 adet veri kullanılmış ve yorumlanmıştır. Kullanılan veri setlerine ait kaynak açıklamaları aşağıdaki verilmiştir.

5.1.1 Beton verisi

Beton, inşaat mühendisliğindeki en önemli materyallerden biridir. Betona ait sıkışma gerilimi dayanımı betonun yaşı ve içerdiği materyallerle yüksek derecede ilintili olmasına rağmen bu ilişki lineer bir bağıntı ile ifade edilemez. Bu veri seti kaynağı Univesity of California, Irvine (UCI)'dir (UCI 2007 arşivi⁷. Beton verisine ait özelliklerin listesi Tablo 5.1'de verilmiştir.

Tablo 5.1 Beton sıkışma gerilme verisine ait özellikleri gösteren tablo

Veri Seti Karak.	Çok deęişk.	Örnek Sayısı	1030	Alanı	Fiziksel
Özellik Karak.	Gerçek	Özellik Sayısı	9	Ek. Tarihi	2007/08/03
İlgili Amaç	Regresyon	Eksik Veri	Yok	İnd. Sayısı	1312

Orijinal Kaynak Sahibi: Prof. I-Cheng Yeh Bilgi Yönetimi Bölümü Chung-Hua Üniversitesi Hsin Chu, Tayvan 30067, R.O.C. e-mail:icyeh@chu.edu.tw

Veriye ait özelliklerin açılımı aşağıdaki gibidir:

1. Özellik: Çimento nicelięi: m³ içerisindeki kg cinsinden miktarı

⁷ <http://archive.ics.uci.edu/beta/datasets.html>

2. Özellik: Fırınlanma sonucu ortaya çıkan cüruf nicelliği: m^3 içerisindeki kg cinsinden miktarı
3. Özellik: Karışımdaki uçucu kül miktarı: m^3 içerisindeki Litre cinsinden miktarı
4. Özellik: Karışımdaki su miktarı: m^3 içerisindeki Litre cinsinden miktarı
5. Özellik: Karışımdaki superplastikleştirici miktarı: m^3 içerisindeki Litre cinsinden miktarı
6. Özellik: Karışımdaki kaba malzeme miktarı: m^3 içerisindeki Litre cinsinden miktarı
7. Özellik: Karışımdaki ince malzeme: m^3 içerisindeki Litre cinsinden miktarı
8. Özellik: karışımın yaşı: Gün cinsinden
9. Özellik: Betonun Sıkışma Dayanımı Gerilmesi: Mega Paskal

5.1.2 Kalp M-Mode aort ve sistol verileri

Bu veriler Selçuk Üniversitesi, Meram Tıp Fakültesi Çocuk Kardiyolojisi Bölümü öğretim üyeleri Prof. Dr. Sevim KARAASLAN ve Prof. Dr. Bülent ORAN tarafından yaklaşık 7 yıllık bir araştırmanın sonucu olarak ortaya çıkmış 2700'den fazla sağlıklı insanın kalp verilerinden oluşmaktadır. Bu veriler Orta Anadolu insanının kalp M-mode değerlerine ait geçerli sağlıklı grafiklerin elde edilmesi için yapılmıştır.

Yeni doğan bir bebek doğduğu günden itibaren 18 yaşına kadar çocuk kabul edilmektedir. Bu bağlamda yeni doğduğu günden itibaren 18 yaşına kadar sürekli değişen vücut çapı ve yüzeyi, kalp ile ilgili verilerinde sürekli değişken olmasına neden olur. Yeni doğan bir bebek ile 5 yaşındaki bir çocuğun aort damarının çapı ve sistol değerleri arasında mutlaka fark olacaktır. Bu kalbe ait verilerin her an için bilinmesini sağlayacak yapay zeka çalışması yeni önerilen normalizasyon metodumuz için de bir amaç olmuştur (Akdemir 2008). Verilerin toplanmasındaki

istatistiksel amaç, Türkiye Orta Anadolu bölgesinde yaşayan sağlıklı çocuklara ait kalp M-mode grafiklerini oluşturmaktır.

M-mode kalp sinyalleri göğüs üzerinden ultrasonografi metodu ile çizgisel bir tarayıcı aracılığıyla alınan sinyallerden meydana gelir. Kalbe ait 7 adet veriyi tek sensörle ölçme imkânı vardır. Bunlardan Aort verisi olarak çalışmada kullanılan veri; kalp aort damarının çıkışı ile ilgili çap verisidir. Bu damarın çapı kalbin kasılması ve gevşemesi sıradındaki hareketlerde kalbin oluşturduğu basınç ile ilgili olarak genişler ve daralır. Bu parametre doğal sınırları içerisinde olması sağlıklı bir kalp için hayati önem taşımaktadır.

Bu çalışmada kullanılan diğer bir veri de kalp sistol bilgisidir. Sistol bilgisi kalbin kasılması ve gevşemesine ait verilerdir. Sağlıklı bir kalp; sistol ve diastol hareketlerine bağlı olarak karıncıklar ve onların hareketleri sonucu kalp kanı aort damarı boyunca vücuda yayar. Karıncıkların farklı çaplarda olmaları ya da kalp kaslarının sertleşmesi gibi tıbbi sorunlar neticesinde sistol ve diastol hareketlerinde farklı değerlerin ölçülmesi kalp için sakıncalı durumdur. Bunun olması gereken sistol değerinin bilinmesi kalp için hayati önemi vardır.

5.1.3 Diş anterior temporal ve üst dudak verileri

Bu veriler Erciyes Üniversitesi Diş hekimliği Fakültesinde Tübitak Projesi kapsamında Prof. Dr. Sadık KARA ve Arş. Gör. Şükrü ÖKKESİM tarafından alınan verilerdir. Takılan Diş protezlerinin 6 ay sonraki olası iyileşme durumlarının tahmin edilmesi bu verilerin hedefleri arasındadır. 20 adet hastanın ilk hasta ve sonraki 6 aydaki iyileşmelerine ait parametreyi içeren tek değişkenli bir tahmin veri setidir (Ökkesim 2006).

5.2 Sınıflama verileri

Sınıflama verisi olarak 5 adet veri seti kullanılmıştır. SPECT datası, kalp datası, doppler, hepatit ve VEP verileri kullanılmıştır. Bu verilerin hepsi sınıflama verisi olup çok değişkenli veri setleridir.

5.2.1 SPECT (Single Photon Emission Computed Tomography) görüntüleri kalp hastalığı teşhisi verileri

Genel olarak kalp ile ilgili hareket analizi, koroner arter çevresi bozuklukları veya miyokardinal anormalliklerle ilgili patolojileri tanımlamaya izin verir. Bu tez çalışmasında kullanılan *SPECT* görüntüleri veri kümesi UCI (University of California, Irvine) makine öğrenmesi veritabanından alınmıştır (UCI veritabanı 2008). Tablo 5.2 VEP veri seti hakkında kaynak verileri göstermektedir.

Tablo 5.2 SPECT veri setine ait özelliklerin istatistikî değerleri

Veri Seti Karak.	Çok deęişk.	Örnek Sayısı	267	Alanı	Yaşam
Özellik Karak.	Kategorik	Özellik Sayısı	22	Ek. Tarihi	2001/10/01
İlgili Amaç	Sınıflama	Eksik Veri	Yok	İnd. Sayısı	19828

İlk sahibi: Krzysztof J. Cios, Lukasz A. Kurgan University of Colorado at Denver, Denver, CO 80217, U.S.A

5.2.2 Kalp hastalığı verisi

Kalp hastalığı, kalbin kan odacıklarını veya kalp kasını etkiler. Kalp rahatsızlığı, genellikle, kan odacık problemidir, örnek olarak kan odacığının darlaşması veya kan akışının blokajı gibi problemlerdir (Yan 2006).

Kalp hastalığı veri seti 76 özellikten ve 303 örnekten oluşmaktadır. Bununla birlikte, bütün yayınlanmış deneylerde sadece bu özelliklerden 14 özellik kullanılmıştır. Kalp hastalığı veri setinin iki değişken sınıfı vardır: hastalık yok (% 55.56) ve hastalık var (%44.44) (UCI veritabanı 2008).

Bu veritabanındaki 270 örnekten 150 tanesi “hastalık yok” sınıfına, 120 tanesi de “hastalık var” sınıfına aittir. Tablo 5.3’ de kalp hastalığı veri kümesindeki özelliklerin istatistiki değerleri verilmiştir.

Tablo 5.3 Kalp hastalığı veri kümesinde bulunan özelliklerin istatistiki değerleri

Veri Seti Karak.	Çok değişik.	Örnek Sayısı	303	Alanı	Yaşam
Özellik Karak.	Kategorik, gerçek, tamsayı	Özellik Sayısı	76	Ek. Tarihi	1988/07/01
İlgili Amaç	Sınıflama	Eksik Veri	Var	İnd. Sayısı	14056

5.2.3 VEP verisi

VEP, optik sinir patolojilerinin hassas bir ölçümü olarak tanımlanır (Atilla ve ark. 2006). Kompleks bir dalga, görsel uyarımdan sonra tahmin edilebilir gecikme zamanlarında meydana gelen ayırt edilebilir pozitif ve negatif tepelerle üretilir. VEP, görsel sistemin fonksiyonel birleşimi ile ilgili önemli teşhis bilgisi sağlayabilir ve klinik ve araştırma laboratuvarlarında kullanılır. VEP, optik sinir veya görsel kabuk hastalıklarının yanı sıra kalıtsal Macular bozulması gibi dış retina hastalıklarında anormal olabilir (Jancic-Stefanovic ve ark. 2007, William ve ark. 1993, Güven ve ark. 2008). Genellikle, VEP’ in klinik kullanımı, N75, P100 ve N145’ in gecikmelerine ve tepe genliğine dayanır (Tasman ve ark. 1992). VEP veri seti 63 özellikten oluşmaktadır.

VEP veri seti çıkışı 2 sınıflıdır; hasta ya da değil.

5.2.4 Doppler sinyali ile damar sertliđi veri kümesi

Damar sertliđi (Atherosclerosis) yaygın bir atardamar hastalıđıdır ve hücre dışındaki lipidler, lifli doku, düz kas hücreleri ve kalsiyum toplanması ile ortaya çıkmaktadır. Doppler sonogramları bir zaman serisinin gücünü açıklamakta ve böylece Doppler sinyalinin spektral analizi atardamar içerisindeki hız dağılımı hakkında bilgi sağlamaktadır (Dirgenali 2003, Evans ve ark. 1989).

Damar sertliđi veri kümesi 114 veriden oluşmaktadır. 114 verinin 60'ı hasta, 54'ü sağlıklı kişilerden oluşmaktadır. Veriler Prof. Dr. Sadık Kara ve Dr. Fatma Latiođlu tarafından Erciyes Üniversitesi Tıp Fak. Radyoloji bölümünde Toshiba PowerVision 6000 Doppler Ultrason ünitesi kullanılarak elde edilmiştir. 8 MHz lineer ultrason sondası sol karotis atardamara darbeli ultrason sinyali iletmek için kullanılmıştır. Karotis atardamarlarda Doppler tepkilerini elde etmek için ultrason cihazının ses çıkış portu kullanılmıştır. Ultrason cihazının çıkışındaki ses sinyalleri 44100 Hz'de örneklenmiştir ve bir giriş/çıkış kartı aracılığı ile bilgisayara gönderilmiştir. Elde edilen veri bir ADC (Analog to Digital Converter) kullanılarak sayısal hale getirilmiş ve bir giriş/çıkış kartı ile bilgisayara aktarılmıştır. Doppler veri seti 61 özellikten oluşan iki sınıflı bir veri seti olup hastalık var ya da yok şeklindedir.

5.2.5 Hepatit verisi

Hepatit hastalıđı veri seti 19 özellikten ve 155 örnekten oluşmaktadır. Hepatit hastalıđı veri setinin iki deđişken sınıfı vardır: hastalık yok ve hastalık var (UCI veritabanı 2008). Hepatit veri setine ait kaynak bilgileri Tablo 5.4'de verilmiştir.

Tablo 5.4 Hepatit veri setinde bulunan özelliklerin istatistiki değerleri

Veri Seti Karak.	Çok deęişık.	Örnek Sayısı	155	Alanı	Yaşam
Özellik Karak.	Kategorik, gerçek, tamsayı	Özellik Sayısı	19	Ek. Tarihi	1988/11/01
İlgili Amaç	Sınıflama	Eksik Veri	Var	İnd. Sayısı	8001

6. PERFORMANS KRİTERLERİ

Yapılan çalışmanın değerlendirilmesi için veri türüne göre değerlendirme kriterleri kullanılmıştır. Sınıflama ya da tahmin amaçlı veri setinde kullanılan değerlendirmeler birbirlerinden farklıdır.

6.1 Tahmin verilerinde kullanılan değerlendirme kriterleri

Sınıflama verilerinin değerlendirilmesinde ortalama karesel hata (MEan Squared Error- MSE), ortalama yüzde hata (Average Deviation AD) ve R^2 istatistiği kullanılmıştır.

6.1.1 Ortalama Devinim

Ortalama devinim, “Average Deviation, (AD)” olarak bilinir. Bu formül oldukça kullanışlı bir formüldür. Sayıların küçüklüğünden bağımsız olarak hataların yüzde cinsinden ortalamasıdır. Örneğin değer 5 çıktı ise bunun anlamı ortalama hata %5 demektir anlamını taşımaktadır. AD’nin formülü Eşitlik 6.1’de gösterilmiştir.

$$AD = \frac{1}{n} \sum_{i=1}^n \frac{|X_{im} - X_{ip}|}{X_{im}} \times 100 \quad (6.1)$$

Eşitlikte X_{im} ölçülen değeri, X_{ip} ise tahmin edilen değeri göstermektedir. Bu eşitliğin kullanışlılığının yanında oldukça kayda değer bir problemi vardır. Eğer hedeflenen değerler arasında sıfır var ise formülden de anlaşılaacağı üzere sonuç ∞ ’a gitmektedir. Bu yüzden AD değeri yüksek çıkıyor ise bu durumda hedeflerden arasında sıfır olması ihtimali göz önüne alınmalıdır. AD bir değerlendirme kriteri

olup ortalama yüzde hatası olarak ifade etmek mümkündür (Kaveh 2008, Razmi-Rad 2007).

6.1.2 Ortalama Karesel Hata

Ortalama Karesel Hata, “Mean Squared Error, (MSE)” olarak verilmiştir. Bu formül hatanın kareleri toplamının ortalaması ile ifade edilir. MSE’nin literatürde oldukça yaygın ve geçerli bir kullanımı vardır. MSE Eşitlik 6.2’den hesaplanabilir.

$$E_i = \frac{1}{n} \sum_{i=1}^n (T_{(i)} - P_{(i)})^2 \quad (6.2)$$

Eşitlikte, P tahmin edilen veriyi T ise hedeflenen veriyi sembolize etmektedir.

6.1.3 R^2 değeri

Bu değerlendirmede hedef 1’dir. 1 sayısına yaklaşılmaya bağlı olarak test değerleri ile hedeflenen değerlerin örtüştüğünü söylemek mümkündür. R^2 değeri aşağıdaki Eşitlik 6.3’den hesaplanabilir.

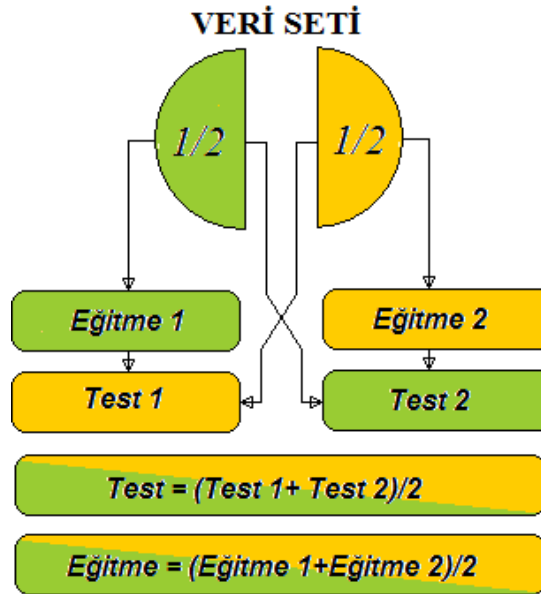
$$R^2 = 1 - \frac{\sum_{m=1}^n (y_{pre,m} - y_{mea,m})^2}{\sum_{m=1}^n (y_{mea,m} - \bar{y}_{mea,m})^2} \quad (6.3)$$

Eşitlikte $y_{pre,m}$ tahmin edilen değer, $y_{mea,m}$ ölçülen değer ve $\bar{y}_{mea,m}$ ise ölçülen değerlerin ortalamasıdır. Bu eşitliği hataların karesinin ortalamaya uzaklığa oranı olarak açıklamak mümkündür. Eğer R^2 değeri büyük sayılar çıkmış ise performansın

tutarsız olduğuna bir işaret olarak kabul edilmelidir. R^2 değerinin bire yaklaştıkça anlamı artar.

6.1.4 2-Kat Çaprazlama

Verileri değerlendirme metotlarından biri olarak da N-kat çaprazlama gösterilebilir. Öyle tür veriler olabilir ki eğitmeyi bozduğu halde test güzel olabilir veya tersi durum söz konusu olabilir. Bu durumda veri seti çalışmada kullanılan şekli ile açıklamak gerekirse %50 test ve %50 eğitme olmak üzere ikiye bölünür. İlk parça eğitme ve ikinci parça test olmak üzere yapay zeka bir defa işletilir (Gongde 2005). Şekil 6.1 'de görüldüğü gibi daha sonra test ile eğitme yer değiştirilerek aynı işlemler tekrarlanır.



Şekil 6.1 İki fold çaprazlama geçerlilik testinin blok diagramı

İki eğitmeden alınan ortalama eğitme hanesine ve iki testin ortalaması da genel test sonucu olarak kabul edilir. Çalışmada YSA ve ANFIS için verilen değerler 2-kat çapraz geçerlilik değerleri sonuçlarıdır. YSA için tüm denemeler 2000 iterasyonluk iki çalışmanın ortalaması, ANFIS için ise standart otomatik atama

sonucu çalışan 200 iterasyonun ortalamasıdır. Çalışmalar MATLAB 7.0 içerisinde yazılmış veya literatürde kullanıma açık uygun Matlab fonksiyon dosyaları ile yapılmıştır.

6.2 Sınıflama verilerinde kullanılan değerlendirme kriterleri

Sınıflama verilerinde bu veri setlerine uygun değerlendirme yöntemleri kullanılmıştır. Çıkışta yalnızca hedeflenen sınıfa ait olma yada olmama gibi belirleyici bir ayırım olduğundan MSE ve AD gibi değerlendirme yöntemlerinin kullanılması yanıltıcı olabilir. Sınıflama veri setlerinin değerlendirilmesinde sınıflama verilerine has değerlendirme kriterleri seçicilik (specificity), duyarlılık (sensitivity) ve ROC eğrisi altında kalan alan (Area Under Curve) kullanılmıştır.

6.2.1 Sınıflama doğruluğu, duyarlılık ve seçicilik değerleri

Bir sınıflama sistemi tarafından yapılan gerçek ve tahmin edilmiş sınıflamalar hakkındaki bilgi, karışıklık matrisi ile verilir. Tablo 6.1 bir karışıklık (confusion) matrisini göstermektedir.

Tablo 6.1 Karışıklık (confusion) matrisi

Gerçek	Tahmin Edilen	
	Negatif	Pozitif
Negatif	DN	YN
Pozitif	YP	DP

Bu matrisde kullanılan terimler şu şekildedir:

- DN, bir örneğin negatif olduğu doğru tahminlerin sayısıdır.
- YP, bir örneğin pozitif olduğu yanlış tahminlerin sayısıdır.
- YN, bir örneğin negatif olduğu yanlış tahminlerin sayısıdır.

- DP, bir örneğin pozitif olduğu doğru tahminlerin sayısıdır.

Sınıflama doğruluğu Eşitlik 6.4'den hesaplanabilir.

$$\text{Sınıflama Doğruluğu (\%)} = \frac{DP + DN}{DP + DN + YN + YP} \times 100 \quad (6.4)$$

Duyarlılık ve seçicilik, bir sınıflama testinin sonuçlarını yorumlamak ve açıklamak için sıklıkla kullanılan bir istatistiksel yöntemdir. Duyarlılık, test edilen bütün pozitif doğrular ve negatif doğruların, tüm doğru ve yanlış çıktıların toplamına oranıdır. Duyarlılık; doğru pozitiflerin, doğru pozitifler ile yanlış pozitiflerin toplamına olan orandır. Seçicilik ise test edilen bütün doğru negatiflerin, doğru negatifler ile yanlış negatiflerin toplamına oranıdır. Duyarlılık ve seçicilik Eşitlik 6.5 ve Eşitlik 6.6'dan hesaplanır.

$$\text{Duyarlılık (\%)} = \frac{DP}{DP + YN} \times 100 \quad (6.5)$$

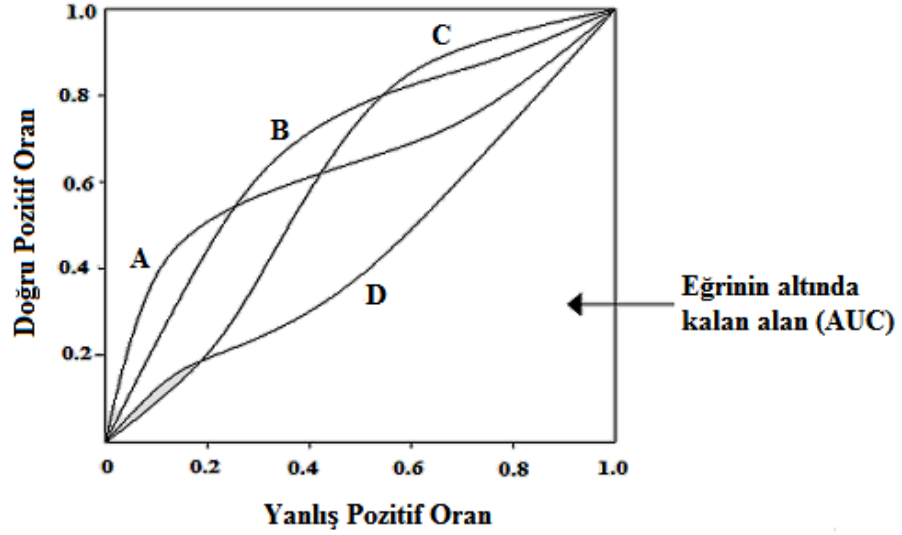
$$\text{Seçicilik (\%)} = \frac{DN}{DN + YP} \times 100 \quad (6.6)$$

6.2.2 ROC eğrisi altında kalan alan

Bir alıcı işlem karakteristiği eğrisi (*ROC*), sınıflandırıcıların performanslarına dayanarak sınıflandırıcıların seçilmesi, organize edilmesi ve görselleştirmesi için kullanılan bir tekniktir (Fawcett 2005). ROC eğrisi altında kalan alan AUC (Area Under of Curve) ile gösterilir.

Bir *ROC* eğrisi, doğru pozitifler ve yanlış pozitifler arasındaki görsel ilişkiyi açıklar. Şekil 6.2 farklı AUC değerlerine sahip *ROC* eğrilerini göstermektedir. *ROC*

eğrisi altında kalan alan ne kadar büyükse sınıflandırıcının performansı da o derece iyidir (Fawcett 2005).



Şekil 6.2 Farklı AUC değerlerine sahip ROC eğrileri

Şekil 6.2’de, A, B, C ve D sınıflayıcılarına ait ROC eğrileri verilmiştir. Şekil 6.2’ den görüleceği üzere eğrinin altında en yüksek alan B sınıflayıcısına, en düşük alan ise D sınıflayıcısına aittir. Yani bu dört sınıflayıcı içerisinde en iyi sınıflayıcı B, en kötü sınıflayıcı ise D’ dir.

ROC eğrileri ve AUC değerinin etkisini görmek için bir örnek verilmiştir. Bu örnekte 5 +, 5 – olmak üzere 2 sınıf ve 10 örnek vardır. Bu veriyi sınıflamak için iki sınıflama algoritması kullanılmıştır. Sınıflandırıcıların elde ettikleri sonuçlar:

İdeal sınıflama durumunda: – – – – – + + + + +
olmaktadır.

1. Sınıflayıcı: – – – – + | – + + + +
2. Sınıflayıcı: + – – – – | + + + + –

sonuçlarını elde etmiştir. Her iki sınıflayıcı için sınıflama doğruluğu %80’ dir. Her iki sınıflayıcının AUC değerleri Eşitlik (6.7)’ ye göre hesaplanır.

$$\hat{A} = \frac{\sum_{i=1}^{n_T} r_i - n_+(n_+ + 1) / 2}{n_+ n_-} \quad (6.7)$$

Burada, r_i , (+) sınıftaki örneklerin sırasını göstermektedir. n_+ ve n_- , sırasıyla pozitif ve negatif kümelerin sayısını göstermektedir. \hat{A} ise ROC eğrisi altında kalan alanıdır. Yukarıdaki örnekte 1. sınıflayıcı için AUC değeri:

$$r_i = 5, 7, 8, 9, 10 \quad n_+ = 5 \quad n_- = 5 \text{ dir.}$$

$$\hat{A} = \frac{\sum_{i=1}^{n_T} r_i - n_+(n_+ + 1) / 2}{n_+ n_-} = \frac{(5 + 7 + 8 + 9 + 10) - (5 \times 6) / 2}{5 \times 5} = \frac{14}{25} = 0.56$$

dir. 2. sınıflayıcı için AUC değeri;

$$r_i = 1, 6, 7, 8, 9 \quad n_+ = 5 \quad n_- = 5 \text{ olması durumunda,}$$

$$\hat{A} = \frac{\sum_{i=1}^{n_T} r_i - n_+(n_+ + 1) / 2}{n_+ n_-} = \frac{(1 + 6 + 7 + 8 + 9) - (5 \times 6) / 2}{5 \times 5} = \frac{16}{25} = 0.64$$

dir. Yukarıdaki sınıflayıcıların sınıflama doğrulukları aynı olmakla birlikte AUC değerleri farklı çıkmıştır. Sınıflayıcıları karşılaştırırken sadece sınıflama doğruluğunun kullanılması başlı başına güvenli olmayabilir, bunun yanında AUC değerlerinin de hesaplanması gerekir.

6.3 Tahmin Verilerinin Değerlendirilmesi

Önerilen metodun denenmesi adına literatürde kaynağı belirgin olan ve kabul görmüş veriler kullanılmıştır. UCI den alınan beton verileri, Selçuk Üniversitesinden alınan kalp verileri ve Erciyes Üniversitesinden alınan diş verileri kullanılmıştır.

Verilerin yapay zeka ile işlenmesine ait parametreler denen en tüm veriler için parametreler aynı değerlerde tutulmaya çalışılmıştır. Yalnız ham veriler için bir istisna söz konusudur ki, ham değerler de YSA transfer fonksiyonu tüm reel sayılar için tanımlı olan *pureline* ile değiştirilmesi zorunluluğu vardır. Normalize edilmiş tüm veriler için parametreler aynı değerler için denenmiştir.

Her iki yapay zeka modeli için de parametrelerin otomatik olarak atanmasına ya da temel ayarlarında (default) kalmasına özen gösterilmiştir. Denemeler sırasında gerekmedikçe (transfer fonksiyonu zorunluluğu gibi) tüm şartların aynı kalmasına özen gösterilmiştir.

YSA için set edilen değerler;

Transfer Fonksiyonu: Logsig

Model: Levenberg-Marquardt

Gizli Düğüm: 1

İterasyon Sayısı: 2000

Hedef Hatası: 0

Öğrenme Oranı. 0.9

Öğrenme Hızını artırma oranı: 1.05

Öğrenme Hızını Azaltma oranı: 0.7

Maksimum Geçerlilik Hataları: 5

Maksimum performans artırılması: 1.04

Momentum Sabiti: 0.9

Minimum Performans Eğimi: 10^{-10}

Eğitme Süresi: ∞

ANFIS için set edilen değerler;

Optimizasyon metodu: Hibrid

İterasyon Sayısı: Durağan hata oluşuncaya kadar

Giriş üyelik fonksiyon tipi sayısı: Otomatik

FIS yapısı: Sub. Clustering

Etkileşim aralığı: 0.5

Saçılma faktörü: 1.25

Kabul oranı: 0.5

Reddetme oranı: 0.15

Önerilen normalizasyon metodunun testi için kullanılan veriler ve bu verilere ait sonuçlar devam eden bölümde anlatılmıştır.

6.3.1 Kalp aort verisi

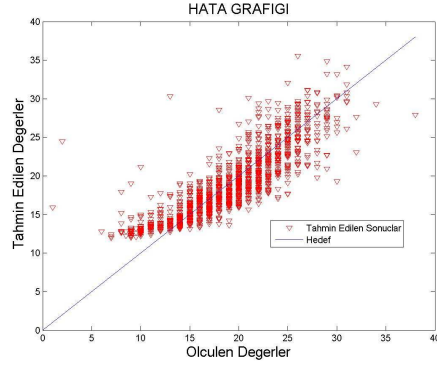
Bu bölümde aort verisine ait YSA ve ANFIS sonuçlar verilmiştir. Sonuçlar eğitime ve teste ait hata grafiklerini ve normalizasyon metotlarına ait sonuçları kapsamaktadır. Hata grafiği, hedeflenen çıkış ile elde edilen çıkış değerlerinin aynı grafik üzerinde gösterilmesinden elde edilmiştir. Hedeflenen değer ile elde edilen değer eşit ise elde edilen grafik bir diagonal çizgi olmalıdır. Bununla birlikte sayısal bir kıyas için normalizasyon metotlarına ait MSE, R^2 ve AD değerleri ayrı ayrı ve toplu halde verilerek kıyaslama yapılmıştır.

6.3.1.1 Kalp aort verileri YSA sonuçları

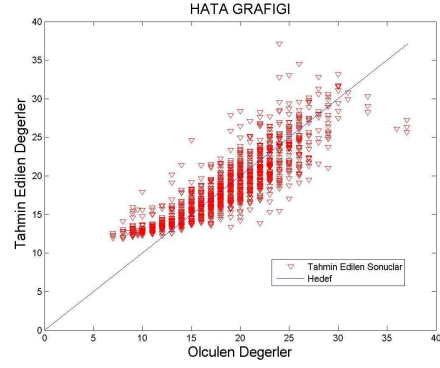
Aort verisine ait YSA sonuçları aşağıda sunulmuştur. Bu sonuçlar önerilen ve literatürde kullanılan metotlarla birlikte verinin ham haline ait performans değerlerini kapsamaktadır.

6.3.1.1.1 Ham kalp aort verilerine ait YSA sonuçları

Kalp aort verisinin, ham halinin YSA ile birlikte işlenmesine ait hata grafiği Şekil 6.3’de gösterilmiştir. Şekil 6.3 (a) aort kalp verisinin eğitime ait hata grafiğini, Şekil 6.3 (b) ise aort kalp verisinin teste ait hata grafiğini göstermektedir.



(a)



(b)

Şekil 6.3 Ham kalp aort verisine ait (a) YSA eğitime hata grafiği (b) YSA test hata grafiği

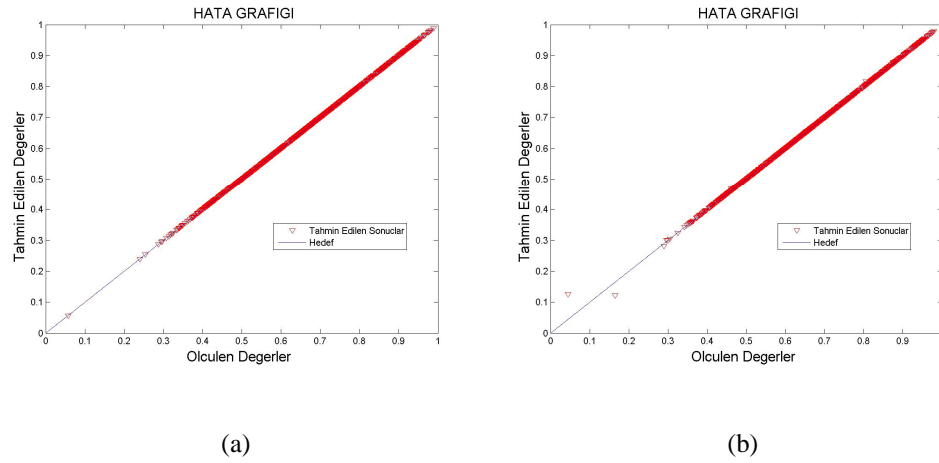
Kalp aort verilerinin normalizasyon uygulanmamış, yalnızca YSA ile işlenmiş haline ait performans kriterlerine ait sayısal veriler Tablo 6.2’de verilmiştir. Ham veriler ile normalize edilmiş verilerin kıyaslanmasında ham veri değerlerinin daha büyük sayılardan oluştuğu dolayısı ile MSE’nin sanal olarak yüksek çıkabileceği göz ardı edilememelidir. Ortalama yüzde hatası % 12.6’dır.

Tablo 6.2 Ham kalp aort verisine ait YSA performans sonuçları

Ham aort verilerinin 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	12.68624	12.71176
MSE	6.863066	6.885417
R ²	0.981972	0.981913

6.3.1.1.2 LBNM metoduna ait kalp aort verisinin YSA performans sonuçları

Kalp aort verisinin, önerilen metot olan LBNM'nin YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.4'de gösterilmiştir. Şekil 6.4 (a) kalp aort verisinin, eğitime ait hata grafiğini, Şekil 6.5 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.4 LBNM metoduna ait kalp aort verilerinin (a) YSA eğitime hata grafiği (b) YSA test hata grafiği

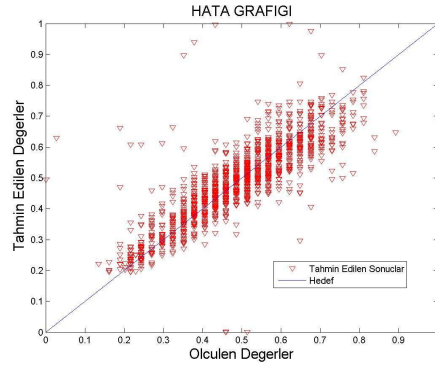
Tablo 6.3 aort verisinin LBNM metodu ile normalize edilmesinden sonra YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir. Eğitim ve test değeri sonuçlarının birbirlerine yakın oluşu ve hata değerlerinin oldukça düşük olması LBNM metodunun bir başarısı olarak kabul edilebilir. R^2 değerlerinin bire yaklaşması bunun bir göstergesidir. Bununla birlikte MSE değerinin sıfır oluşu dikkate değerdir.

Tablo 6.3 LBNM metoduna ait kalp aort verilerinin YSA performans sonuçları

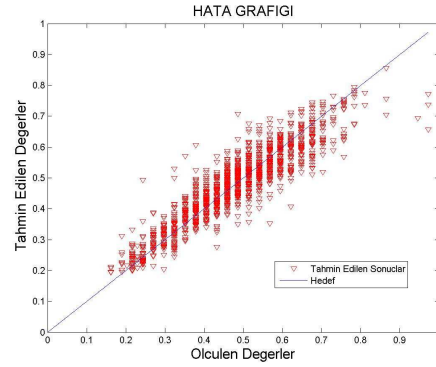
LBNM metodu için 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	0.008410	0.138374
MSE	0.000000	0.000020
R^2	1.000000	0.999967

6.3.1.1.3 Minimum-Maksimum metoduna ait kalp aort verisinin YSA performans sonuçları

Kalp aort verisinin, minimum-maksimum metodunun YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.5’de gösterilmiştir. Şekil 6.5 (a) kalp aort verisinin, eğitime ait hata grafiğini, Şekil 6.5 (b) ise teste ait hata grafiğini göstermektedir.



(a)



(b)

Şekil 6.5 Min.-Mak. Metoduna ait kalp aort verilerinin (a) YSA eğitime hatası grafiği (b) YSA test hatası grafiği

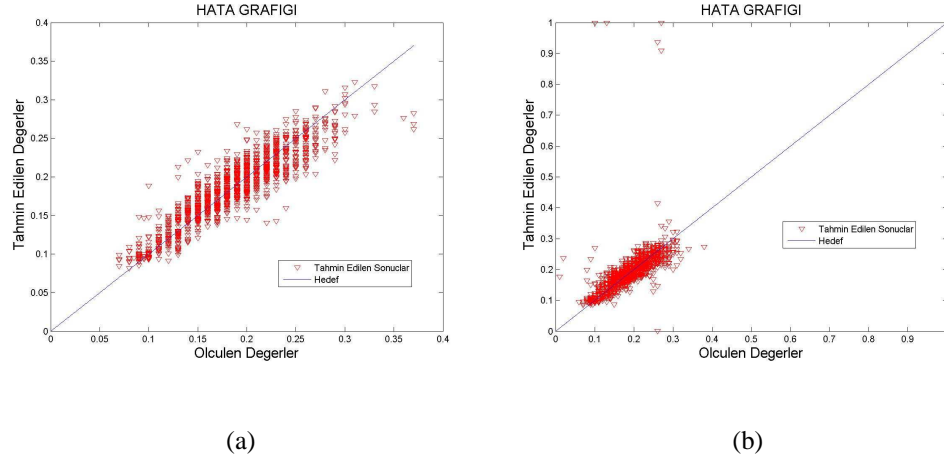
YSA ile kalp aort verisinin minimum-maksimum metoduna göre normalize edilmiş sayısal sonuçları Tablo 6.4’de verilmiştir.

Tablo 6.4 Min.-Mak. Metodunun kalp aort verilerinin YSA performans sonuçları

Min.-Mak. Metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0.003234	0.005256
R ²	0.987115	0.97907

6.3.1.1.4 Ondalık ölçekleme metoduna ait kalp aort verisinin YSA performans sonuçları

Kalp aort verisinin, ondalık ölçekleme metodunun YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.6'da gösterilmiştir. Şekil 6.6 (a) aort verisinin, eğitime ait hata grafiğini, Şekil 6.6 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.6 Ondalık ölçeklemeye ait kalp aort verilerinin (a) YSA eğitime hatası grafiği (b) YSA test hatası grafiği

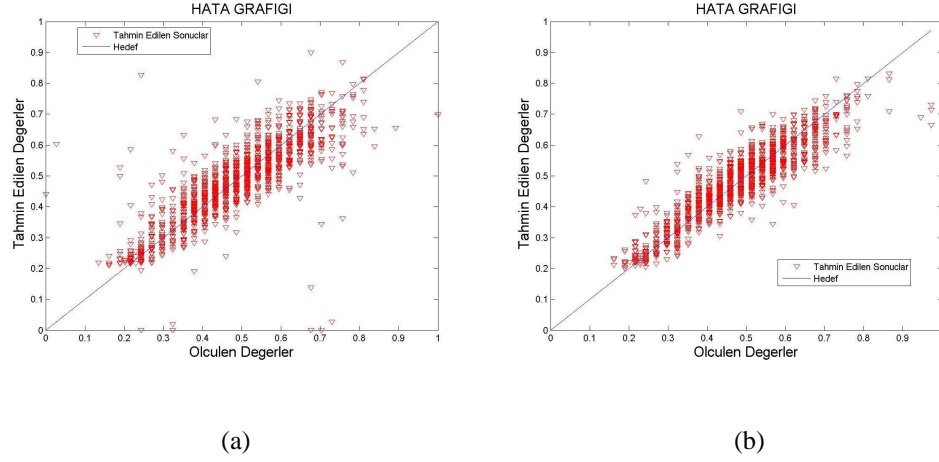
Kalp aort verilerinin ondalık ölçekleme metoduna göre ön işlemeden geçirilmiş ve YSA ile işlenmiş haline ait performans kriterlerinin değerleri Tablo 6.5'de verilmiştir. Tabloda eğitime ve teste ait değerlendirmeler aynı anda verilmiştir.

Tablo 6.5 Ondalık ölçeklemeye ait kalp aort verilerinin YSA performans sonuçları

Ondalık Metoda İlişkin 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	9.442827	11.09331
MSE	0.000437	0.000769
R ²	0.988514	0.979778

6.3.1.1.5 Z-skor metoduna ait kalp aort verisinin YSA performans sonuçları

Kalp aort verisinin, z-skor metodunun YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.7’de gösterilmiştir. Şekil 6.7 (a) aort verisinin, eğitime ait hata grafiğini, Şekil 6.7 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.7 Z-skor metoduna ait kalp aort verilerinin (a) YSA eğitime hatası grafiği (b) YSA test hatası grafiği

Aort verilerinin ondalık ölçekleme metoduna göre ön işlemeden geçirilmiş ve YSA ile işlenmiş haline ait performans kriterlerinin değerleri Tablo 6.6’da verilmiştir. Tabloda eğitime ve teste ait değerlendirmeler aynı anda verilmiştir.

Tablo 6.6 Z-skor metoduna ait kalp aort verilerinin YSA eğitime ve test performans sonuçları

Z-skor metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0.003152	0.004924
R ²	0.987443	0.980393

Aort verisine ait tüm metotların eğitime test değerleri, kıyas yapılabilmesi adına tek bir tabloda verilmiştir. Tablo 6.7 çalışmada önerilen ve literatürdeki yaygın normalizasyon metotlarının kalp aort verisini YSA ile değerlendirilmesine ait eğitime

hatalarının tümünü ait sonuçları göstermektedir. Tablo 6.7 kalp aort verisinin testine ait performans değerlerini göstermektedir.

Tablo 6.7 Tüm normalizasyon metotlarına ait kalp aort verilerinin YSA eğitme performans sonuçları

	EBNM	LBNM	Ham	Z-skör	Ondalık	Min.-Mak.
AD	9.3665	0.00841	12.68624	∞	9.442827	∞
MSE	0.0000004	0	6.863066	0.003152	0.000437	0.003234
R²	0.988495	1	0.981972	0.987443	0.988514	0.987115

Tablo 6.8 Tüm normalizasyon metotlarına ait kalp aort verilerinin YSA test performans sonuçları

	EBNM	LBNM	Ham	Z-skör	Ondalık	Min.-Mak.
AD	12.9744	0.138374	12.71176	∞	11.09331	∞
MSE	0.000097	0.00002	6.885417	0.004924	0.000769	0.005256
R²	0.735514	0.999967	0.981913	0.980393	0.979778	0.97907

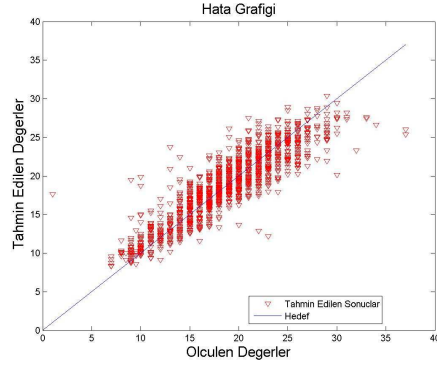
Tablo 6.7 ve 6.8’den görüldüğü gibi, satır normalizasyonunu kullanan YSA uygulaması MSE, AD ve R² kriterlerinin tamamında öndedir. Özellikle test değerleri kendisine en yakın minimum-maksimum metodundan MSE’ye göre 262.8 kat ve R² değerine göre 1.02 kat daha iyidir.

6.3.1.2 Kalp aort verisi ANFIS sonuçları

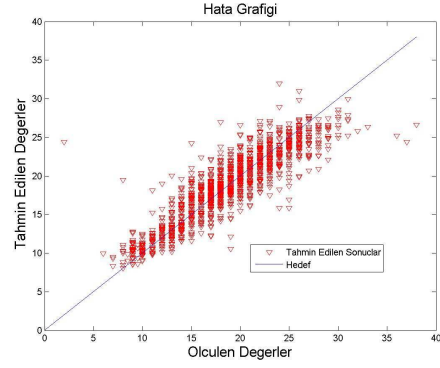
Kalp aort verisine ait ANFIS sonuçları aşağıda sunulmuştur. Bu sonuçlar önerilen ve literatürde kullanılan metotlarla birlikte verinin ham haline ait performans değerlerini kapsamaktadır.

6.3.1.2.1 Kalp aort ham verilerine ait ANFIS sonuçları

Kalp aort verisinin, ham halinin ANFIS ile birlikte işlenmesine ait hata grafiği Şekil 6.8’de gösterilmiştir. Şekil 6.8 (a) aort verisinin, eğitime, Şekil 6.8 (b) ise teste ait hata grafiğini göstermektedir.



(a)



(b)

Şekil 6.8 Ham kalp aort verilerinin (a) ANFIS eğitime hata grafiği (b) ANFIS test hata grafiği

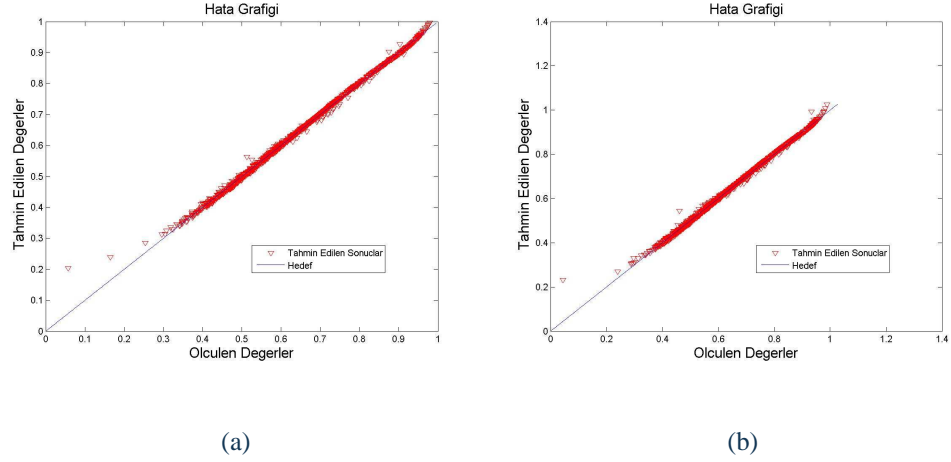
Kalp aort verilerinin herhangi bir ön işlemeden geçirilmeden yalnızca ANFIS ile işlenmesinden elde edilen performans kriterlerinin değerleri Tablo 6.9’da verilmiştir. Tabloda eğitime ve teste ait değerlendirmeler aynı anda verilmiştir.

Tablo 6.9 Ham kalp aort ham verilerinin ANFIS performans sonuçları

Ham verilere ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	10.30585	10.56803
MSE	5.253771	6.094904
R ²	0.986189	0.983974

6.3.1.2.2 LBNM metoduna ait kalp aort verilerinin ANFIS performans sonuçları

Kalp aort verisinin, önerilen metot olan LBNM'nin ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.9'da gösterilmiştir. Şekil 6.9 (a) aort verisinin, eğitime ait hata grafiğini, Şekil 6.9 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.9 LBNM metoduna ait kalp aort verilerinin (a) ANFIS eğitime hata grafiği (b) ANFIS test hata grafiği

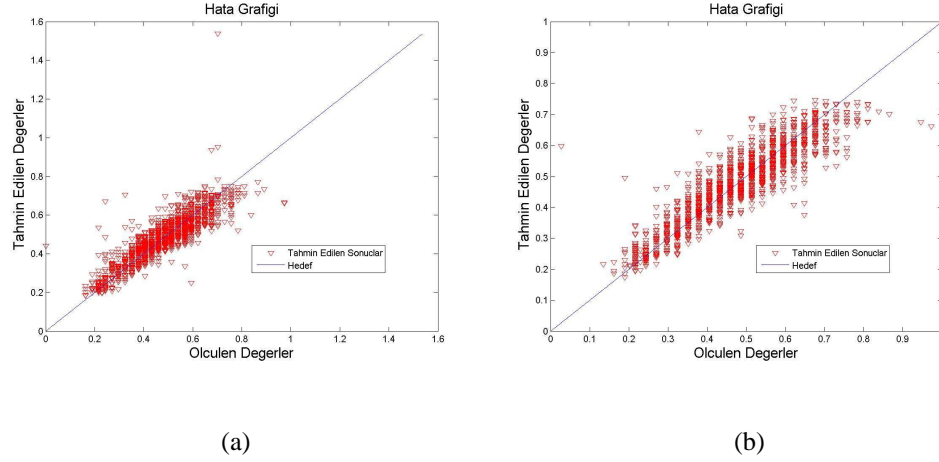
Tablo 6.10 aort verisinin LBNM metodu ile normalize edilmesinden sonra ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir. Eğitim ve test değeri sonuçlarının birbirlerine yakın oluşu ve hata değerlerinin oldukça düşük olması LBNM metodunun bir başarısı olarak kabul edilebilir. R^2 değerlerinin eğitime ve testte ideal değer olan 1'e yakın çıkması bu metodun bir başarısıdır.

Tablo 6.10 LBNM metoduna ait kalp aort verilerinin ANFIS performans sonuçları

LBNM metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	1.281	1.131
MSE	0	0
R^2	1	1

6.3.1.2.3 Minimum-Maksimum metoduna ait kalp aort verileri ANFIS performans sonuçları

Kalp aort verisinin, minimum-maksimum metodunun ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.10'da gösterilmiştir. Şekil 10 (a) aort verisinin, eğitime ait hata grafiğini, Şekil 6.10 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.10 Min.-Mak. metoduna ait kalp aort verilerinin (a) ANFIS eğitime hatası grafiği (b) ANFIS test hatası grafiği

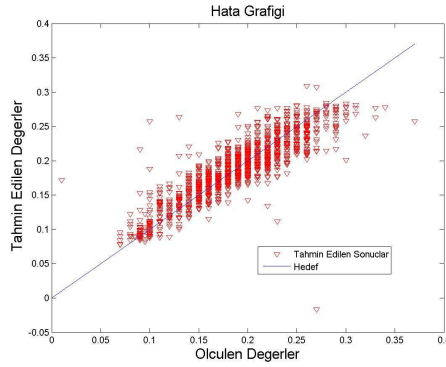
Tablo 6.11 kalp aort verisinin Min.-Mak. metodu ile normalize edilmesinden sonra ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.

Tablo 6.11 Min.-Mak. Metoduna ait kalp aort verilerinin ANFIS performans sonuçları

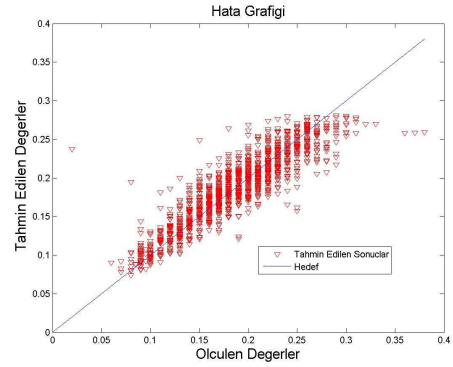
Min.-Mak. metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0.003754	0.004343
R ²	0.98505	0.982699

6.3.1.2.4 Ondalık ölçekleme metoduna ait kalp aort verilerinin ANFIS performans sonuçları

Kalp aort verisinin, ondalık ölçekleme metodunun ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.19'da gösterilmiştir. Şekil 6.11 (a) aort verisinin, eğitime ait hata grafiğini, Şekil 6.11 (b) ise teste ait hata grafiğini göstermektedir.



(a)



(b)

Şekil 6.11 Ondalık ölçeklemeye ait kalp aort verilerinin (a) ANFIS eğitime hata grafiği (b) ANFIS test hata grafiği

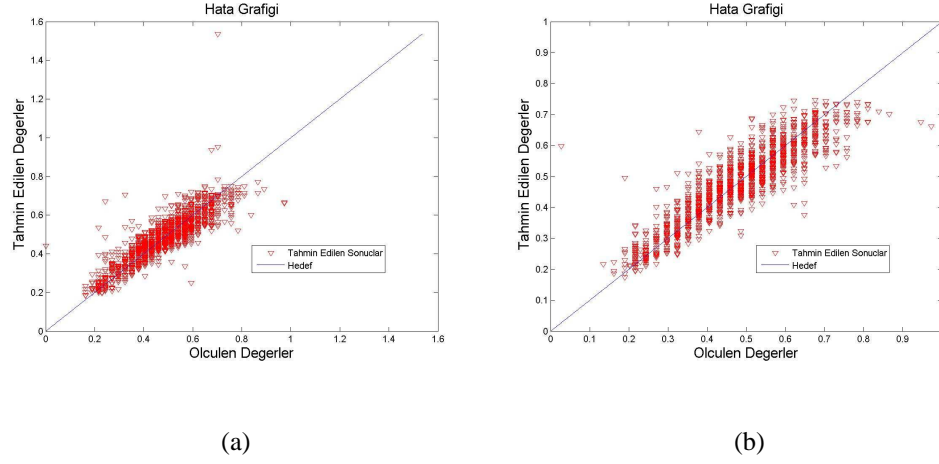
Aort verilerinin ondalık ölçekleme metoduna göre normalize edilmiş ve ANFIS ile işlenmiş haline ait performans kriterlerinin değerleri Tablo 6.12'de verilmiştir. Tabloda eğitime ve teste ait değerlendirmeler aynı anda verilmiştir.

Tablo 6.12 Ondalık ölçekleme metoda ait kalp aort verilerinin ANFIS performans sonuçları

Ondalık Ölçeklemeye ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	10.13157	10.43329
MSE	0.000511	0.000579
R ²	0.986573	0.984789

6.3.1.2.5 Z-skor metoduna ait kalp aort verilerinin ANFIS performans sonuçları

Kalp aort verisinin, z-skor metodunun ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.12’de gösterilmiştir. Şekil 6.12 (a) aort verisinin, eğitime ait hata grafiğini, Şekil 6.12 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.12 Z-skor metoduna ait kalp aort verilerinin (a) ANFIS eğitime hatası grafiği (b) ANFIS test hatası grafiği

Kalp aort verilerinin ondalık ölçekleme metoduna göre ön işlemeden geçirilmiş ve ANFIS ile işlenmiş haline ait performans kriterlerinin değerleri Tablo 6.13’de verilmiştir. Tabloda eğitime ve teste ait değerlendirmeler aynı anda verilmiştir.

Tablo 6.13 Z-skor metoduna ait kalp aort verilerinin ANFIS performans sonuçları

Z-skor metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0.003152	0.004924
R ²	0.987443	0.980393

Kalp aort verisi için tüm metotların eğitime test değerleri, kıyas yapılabilmesi adına tek bir tabloda verilmiştir. Tablo 6.14 çalışmada önerilen ve literatürdeki

yaygın normalizasyon metotlarının aort verisini ANFIS ile değerlendirilmesine ait eğitime hatalarının tümünü ait sonuçları göstermektedir. Tablo 6.15 kalp aort verisinin testine ait ANFIS performans değerlerini göstermektedir.

Kalp aort verisi için en başarılı sonuç gerek eğitime ve gerekse test için ANFIS uygulaması içinde önerilen metotlar çıkmaktadır. Testteki başarıda Öklid uzaklığını normalizasyon olarak kullanılmasını öneren LBNM ve EBNM metotları oldukça başarılı değerlere sahiptir. LBNM metodu kendisine en yakın ondalık ölçeklemeden AD kriterine göre 7.9 kat daha iyidir. R^2 değerinin hem test hemde eğitimde 1 çıkması oldukça başarılı bir sonuçtur.

Tablo 6.14 Tüm normalizasyon metotlarına ait kalp aort verilerinin ANFIS eğitime performans sonuçları

	EBNM	LBNM	Ham	Z-skör	Ondalık	Min.-Mak.
AD	10.42	1.281	10.30585	∞	10.13157	∞
MSE	0	0	5.253771	0.003152	0.000511	0.003754
R^2	0.985	1.0	0.986189	0.987443	0.986573	0.98505

Tablo 6.15 Tüm normalizasyon metotlarına ait kalp aort verisi ANFIS test performans sonuçları

	EBNM	LBNM	Ham	Z-skör	Ondalık	Min.-Mak.
AD	10.48	1.131	10.56803	∞	10.43329	∞
MSE	0	0	6.094904	0.004924	0.000579	0.004343
R^2	0.987	1.0	0.983974	0.980393	0.984789	0.982699

6.3.2 Beton Verisi

Bu bölümde beton verisine ait YSA ve ANFIS sonuçlar verilmiştir. Sonuçlar eğitime ve teste ait hata grafiklerini ve eğitime ve testten elde edilen giriş-çıkış grafiklerini kapsamaktadır. Eğitime ile test verilerine ait grafikler giriş ile çıkış arasındaki uyumluluğa bağlı olarak lineer düz bir çizgi şeklinde olur. Hata grafiğinde ortaya çıkan değerler, arzu edilen değer ile elde edilen çıkışın farkına bağlı olarak

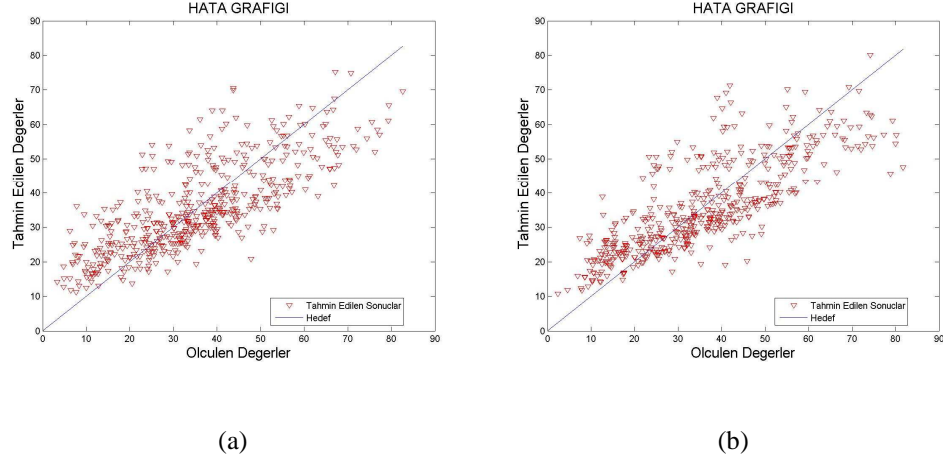
ideal ilişkiyi gösteren çizgiden sapar. Bu yüzden ideali gösteren düz çizgiye olan sadakat sonuca ait kaliteyi de göstermiş olur.

6.3.2.1 Beton verileri YSA sonuçları

Beton verisine ait YSA sonuçları aşağıda sunulmuştur. Bu sonuçlar önerilen ve literatürde kullanılan metotlarla birlikte verinin ham haline ait performans değerlerini kapsamaktadır.

6.3.2.1.1 Ham beton verilerine ait YSA sonuçları

Beton verisinin, ham halinin YSA ile birlikte işlenmesine ait hata grafiği Şekil 6.13’de gösterilmiştir. Şekil 6.13 (a) beton verisinin, eğitime, Şekil 6.13 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.13 Ham beton verilerinin (a) YSA eğitime hata grafiği (b) YSA test hata grafiği

Beton verilerinin normalizasyon uygulanmamış, yalnızca YSA ile işlenmiş haline ait performans kriterlerine ait sayısal veriler Tablo 6.16’da verilmiştir. Ham verilerin diğer metotlarla kıyaslanması yalnızca istatistiksel açıdan yapılabilir. Bunun en önemli nedeni normalize edilen verilerin 0 ile 1 arasında olmasına rağmen ham

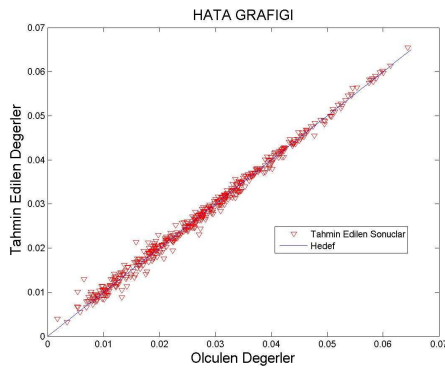
veri değerlerinin daha büyük sayılardan oluşmasıdır. Ham veriler için performans ölçümlerinin eğitime test adına birbirlerine oldukça yakın çıkmaları verinin hiçbir ön işlemeye tabi tutulmamasının da bir göstergesi olarak kabul edilebilir. Bununla birlikte R^2 istatistik değerrinin normalizasyon yapılmış durumdaki değerlere yakın olduğu dikkati çeken bir sonuçtur. Buradan algılanması gereken sonuç elde edilen ham değerler ile normalize edilmiş değerlerin birbirinin benzerleri olduğudur.

Tablo 6.16 Ham beton verilerinin YSA performans sonuçları

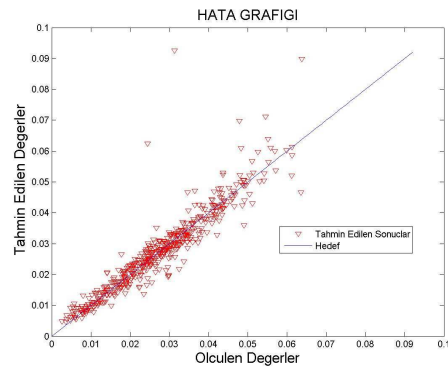
Ham verilere ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	31.20913	31.56236
MSE	106.7343	108.6072
R^2	0.931645	0.930446

6.3.2.1.2 LBNM metoduna ait beton verilerinin YSA performans sonuçları

Beton verisinin, önerilen metot olan LBNM'nin YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.14'de gösterilmiştir. Şekil 6.14 (a) beton verisinin, eğitime ait hata grafiğini, Şekil 6.14 (b) ise teste ait hata grafiğini göstermektedir.



(a)



(b)

Şekil 6.14 LBNM metoduna ait beton verilerinin (a) YSA eğitime hata grafiği (b) YSA test hata grafiği

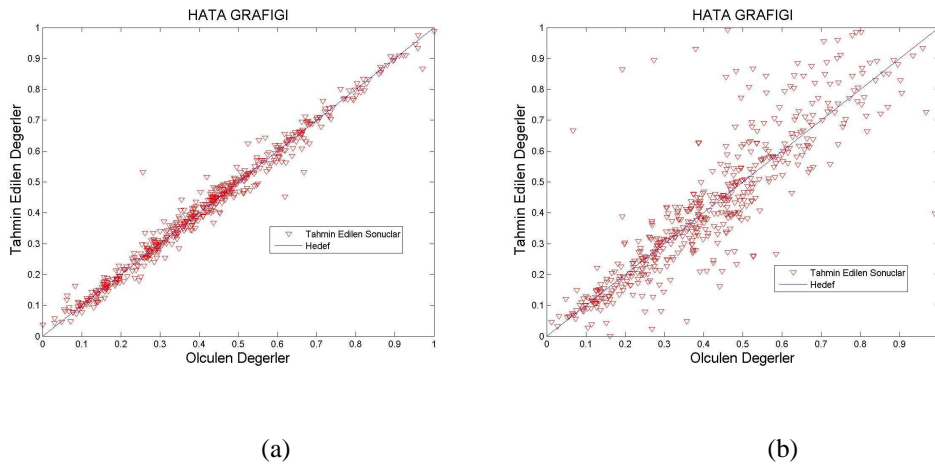
Tablo 6.17 beton verisinin LBNM metodu ile normalize edilmesinden sonra YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir. Eğitim ve test değeri sonuçlarının birbirlerine yakın oluşu ve hata değerlerinin oldukça düşük olması LBNM metodunun bir başarısı olarak kabul edilebilir. R^2 değerinin eğitmede 1'e oldukça yakın oluşu ve hem eğitime hemde test için MSE'in sıfıra yakınlığı bu metodun başarısının bir göstergesidir.

Tablo 6.17 LBNM metoduna ait beton verilerinin YSA performans sonuçları

LBNM Metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	1.671537	4.564226
MSE	0.000014	0.000197
R^2	0.999817	0.997465

6.3.2.1.3 Minimum-Maksimum metoduna ait beton verilerinin YSA performans sonuçları

Beton verisinin, minimum-maksimum metodunun YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.15'de gösterilmiştir.



Şekil 6.15 Min.-Mak. metoduna ait beton verilerinin (a) YSA eğitime hatası grafiği (b) YSA test hatası grafiği

Beton Şekil 6.15 (a) beton verisinin, eğitime ait hata grafiğini, Şekil 6.15 (b) ise teste ait hata grafiğini göstermektedir. Beton verilerinin Min.-Mak. normalizasyonu uygulanmış, YSA ile işlenmiş haline ait performans kriterlerine ait sayısal veriler Tablo 6.18’de verilmiştir. Minimum-maksimum veri setinde hedeflenen değerler içerisinde sıfır olması AD değerinin sonsuz çıkmasına neden olmaktadır.

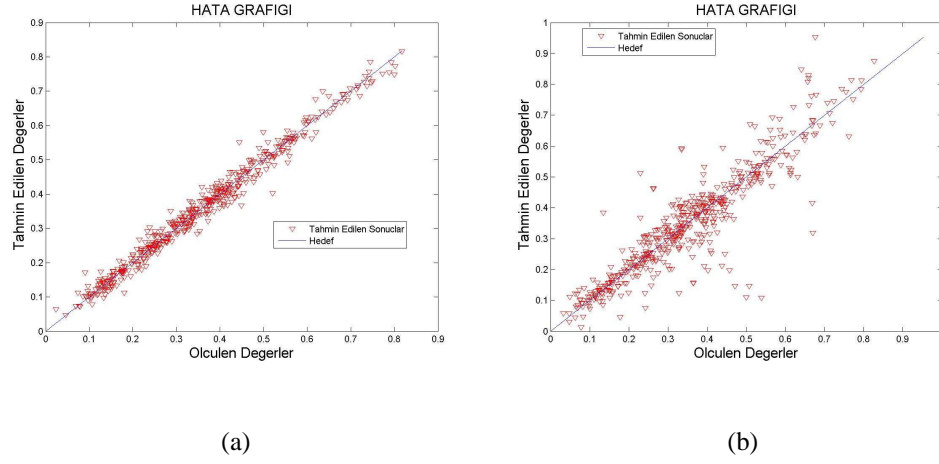
Tablo 6.18 Min.-Mak. Metoduna ait beton verilerinin YSA performans sonuçları

Min.-Mak.metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0.000774	0.0108
R²	0.99644	0.950299

6.3.2.1.4 Ondalık ölçekleme metoduna ait beton verilerinin YSA performans sonuçları

Beton verisinin, ondalık metodunun YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.16’da gösterilmiştir. Şekil 6.16 (a) beton verisinin, eğitime ait hata grafiğini, Şekil 6.16 (b) ise teste ait hata grafiğini göstermektedir.

Beton verilerinin ondalık ölçekleme metoduna göre ön işlemeden geçirilmiş ve YSA ile işlenmiş haline ait performans kriterlerinin değerleri Tablo 6.19’da verilmiştir. Tabloda eğitime ve teste ait değerlendirmeler aynı anda verilmiştir.



Şekil 6.16 Ondalık ölçeklemeye ait beton verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği

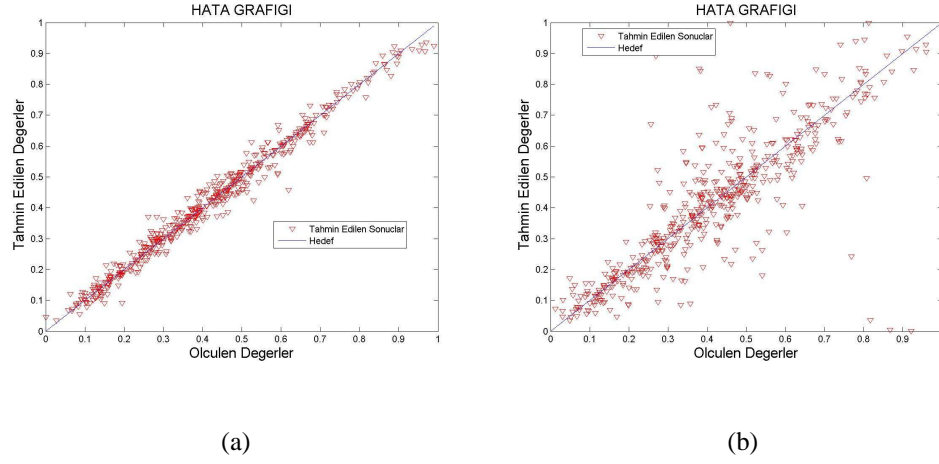
Tablo 6.19 Ondalık ölçekleme metoduna ait beton verilerinin YSA performans sonuçları

Ondalık Metoda ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	6.025085	13.75659
MSE	0.000498	0.004889
R ²	0.996811	0.968674

6.3.2.1.5 Z-skor metoduna ait beton verilerinin YSA performans sonuçları

Beton verisinin, z-skor metodunun YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.17'de gösterilmiştir. Şekil 6.17 (a) beton verisinin, eğitime ait hata grafiğini, Şekil 6.17 (b) ise teste ait hata grafiğini göstermektedir.

Beton verilerinin ondalık ölçekleme metoduna göre ön işlemeden geçirilmiş ve YSA ile işlenmiş haline ait performans kriterlerinin değerleri Tablo 6.20'de verilmiştir. Tabloda eğitime ve teste ait değerlendirmeler aynı anda verilmiştir.



Şekil 6.17 Z-skor metoduna ait beton verilerinin (a) YSA eğitme hatası grafiği (b) YSA test hatası grafiği

Tablo 6.20 Z-skor metoduna ait beton verilerinin performans sonuçları

Z-skor Metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0.000843	0.012305
R ²	0.996119	0.943298

Beton verisi için tüm metotların eğitme test değerleri, kıyas yapılabilmesi adına tek bir tabloda verilmiştir. Tablo 6.21, çalışmada önerilen ve literatürdeki yaygın normalizasyon metotlarının beton verisini YSA ile değerlendirilmesine ait eğitme hatalarının tümüne ait sonuçları göstermektedir. Tablo 6.22 beton verisinin testine ait performans değerlerini göstermektedir.

Beton verisi için en başarılı sonuç gerek eğitme ve gerekse test için LBNM metodunda ölçülmüştür. Testteki başarıda matematiksel olarak LBNM metodu MSE değeri 0.000197 olarak ölçülmüştür. Kendisine en yakın metot olarak, ondalık metottan ölçülen MSE değeri 0.004889'dur ve bu değer LBNM'nin sahip olduğu MSE değerinden 24.8 kat daha kötüdür.

Tablo 6.21 Tüm normalizasyon metotlarına ait beton verisi YSA eğitime performans sonuçları

	EBNM	LBNM	Ham	Z-skör	Ondalık	Min.-Mak.
AD	6.0087	1.671537	31.20913	∞	6.025085	∞
MSE	0.000003	0.000014	106.7343	0.000843	0.000498	0.000774
R²	0.9969061	0.999817	0.931645	0.996119	0.996811	0.99644

Tablo 6.22 Tüm normalizasyon metotlarına ait YSA test performans sonuçları

	EBNM	LBNM	Ham	Z-skör	Ondalık	Min.-Mak.
AD	14.7625	4.564226	31.56236	∞	13.75659	∞
MSE	0.000054	0.000197	108.6072	0.012305	0.004889	0.0108
R²	0.944228	0.997465	0.930446	0.943298	0.968674	0.950299

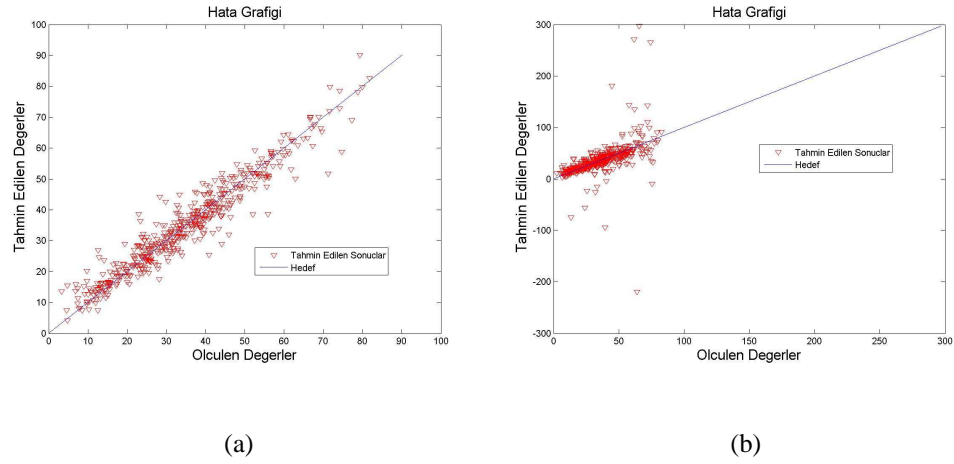
Önerilen metot LBNM için R² değeri oldukça iyidir ve 1'e en yakın bir değere sahiptir.

6.3.2.2 Beton verilerinin ANFIS sonuçları

Beton verisine ait ANFIS sonuçları aşağıda sunulmuştur. Bu sonuçlar önerilen ve literatürde kullanılan metotlarla birlikte verinin ham haline ait performans değerlerini kapsamaktadır.

6.3.2.2.1 Ham beton verilerine ait ANFIS sonuçları

Beton verisinin, ham halinin ANFIS ile birlikte işlenmesine ait hata grafiği Şekil 6.18'de gösterilmiştir. Şekil 6.18 (a) beton verisinin, eğitime, Şekil 6.18 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.18 Ham beton verilerinin (a) ANFIS eğitime hata grafiği (b) ANFIS test hata grafiği

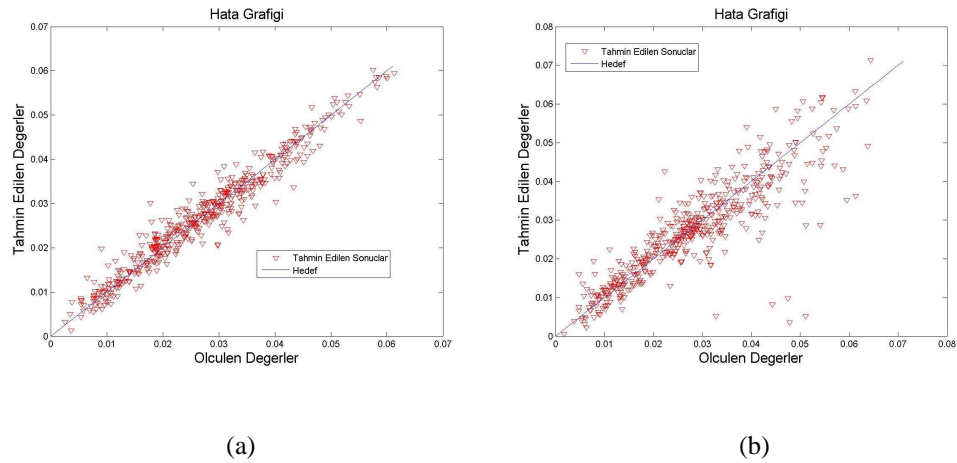
Beton verilerinin normalizasyon uygulanmamış, yalnızca ANFIS ile işlenmiş haline ait performans kriterlerine ait sayısal veriler Tablo 6.23’de verilmiştir. Ham verilerin diğer metotlarla kıyaslanması yalnızca istatistiksel açıdan yapılabilir. Bunun en önemli nedeni normalize edilen verilerin 0 ile 1 arasında olmasına rağmen ham veri değerlerinin daha büyük sayılardan oluşmasıdır. Ham verinin testteki R^2 kriterinin LBNM metodundan düşük kaldığına ve literatürdeki en yaygın metot olarak bilinen Min.-Mak.metodundan daha başarılı olduğuna not etmekte fayda vardır. R^2 değeri LBNM, Ham ve Min.-Mak. normalizasyonu için sırasıyla 0.984, 0.707869 ve 0.681489 ölçülmüştür.

Tablo 6.23 Ham beton verilerinin ANFIS performans sonuçları

Ham verilere ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	11.55066	28.05473
MSE	18.3362	467.6429
R^2	0.988194	0.707869

6.3.2.2.2 LBNM metoduna ait beton verilerinin ANFIS performans sonuçları

Beton verisinin, önerilen metot olan LBNM'nin ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.19'da gösterilmiştir. Şekil 6.19 (a) beton verisinin, eğitime ait hata grafiğini, Şekil 6.19 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.19 LBNM metoduna ait beton verilerinin (a) ANFIS eğitime hata grafiği (b) ANFIS test hata grafiği

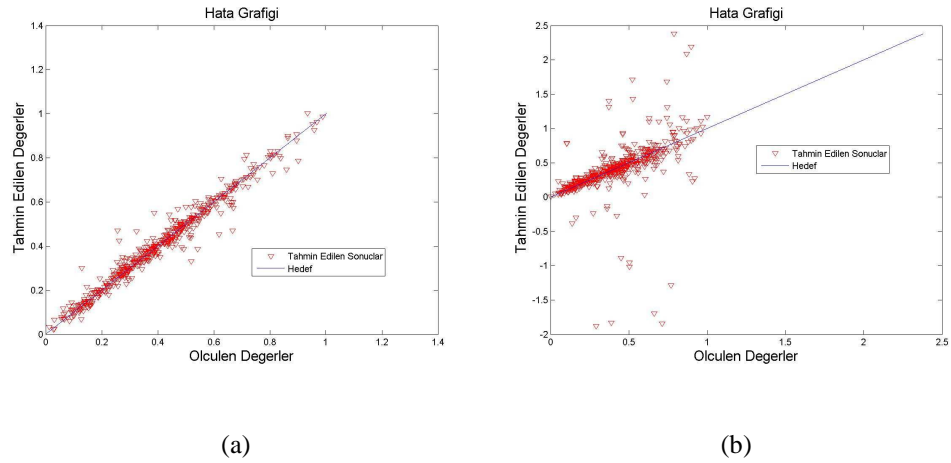
Tablo 6.24 beton verisinin LBNM metodu ile normalize edilmesinden sonra ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.

Tablo 6.24 LBNM metoduna ait beton verilerinin ANFIS performans sonuçları

LBNM Metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	3.327	9.930
MSE	0	0.001
R ²	0.999	0.984

6.3.2.2.3 Minimum-Maksimum metoduna ait beton verilerinin ANFIS performans sonuçları

Beton verisinin, minimum-maksimum metodunun ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.20’de gösterilmiştir. Şekil 6.20 (a) beton verisinin, eğitime ait hata grafiğini, Şekil 6.20 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.20 Min.-Mak. Metoduna ait beton verilerinin (a) ANFIS eğitim hata grafiği (b) ANFIS test hata grafiği

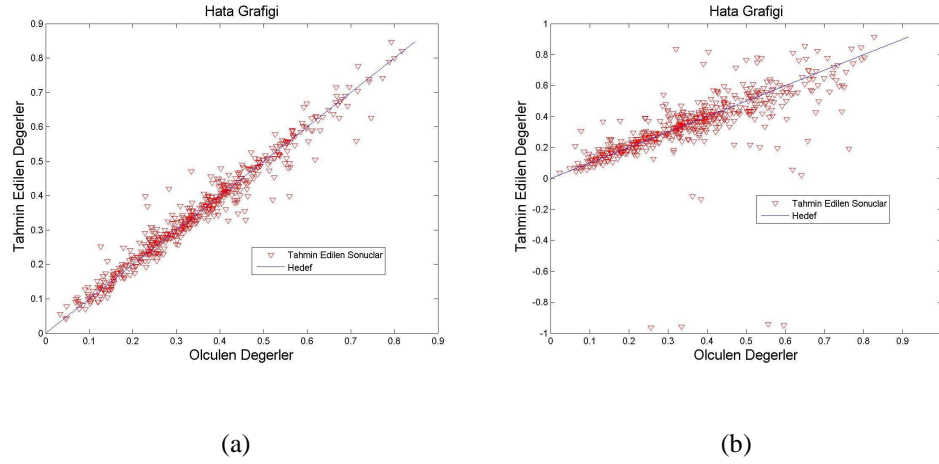
Beton verilerinin Min.-Mak. normalizasyonu uygulanmış, ANFIS ile işlenmiş haline ait performans kriterlerine ait sayısal veriler Tablo 6.25’de verilmiştir.

Tablo 6.25 Min.-Mak. Metoduna ait beton verilerinin ANFIS performans sonuçları

Min.-Mak. Metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0.001132	0.070956
R ²	0.994715	0.681489

6.3.2.2.4 Ondalık ölçekleme metoduna ait beton verilerinin ANFIS performans sonuçları

Beton verisinin, ondalık metodunun ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.21’de gösterilmiştir. Şekil 6.21 (a) beton verisinin, eğitime ait hata grafiğini, Şekil 6.21 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.21 Ondalık ölçekleme metoduna ait beton verilerinin (a) ANFIS eğitime hatası grafiği
(b) ANFIS test hatası grafiği

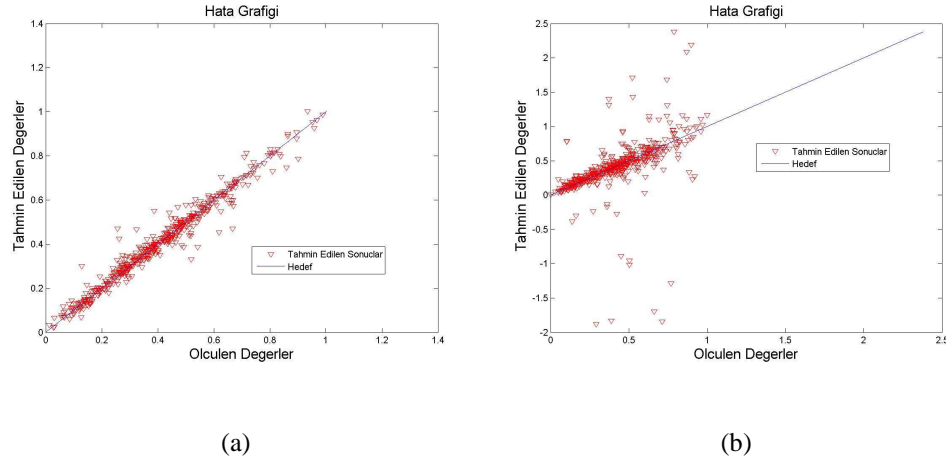
Beton verilerinin ondalık ölçekleme metoduna göre ön işlemeden geçirilmiş ve ANFIS ile işlenmiş haline ait performans kriterlerinin değerleri Tablo 6.26’da verilmiştir. Tabloda eğitime ve teste ait değerlendirmeler aynı anda verilmiştir.

Tablo 6.26 Ondalık ölçekleme metoduna ait beton verilerinin ANFIS performans sonuçları

Ondalık Metoda ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	5.98858	18.54774
MSE	0.000664	0.016865
R ²	0.995591	0.89498

6.3.2.2.5 Z-skor metoduna ait beton verilerinin ANFIS performans sonuçları

Beton verisinin, z-skor metodunun ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.22’de gösterilmiştir. Şekil 6.22 (a) beton verisinin, eğitime ait hata grafiğini, Şekil 6.22 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.22 Z-skor metoduna ait beton verilerinin (a) ANFIS eğitime hatası grafiği (b) ANFIS test hatası grafiği

Beton verilerinin z-skor normalizasyon metoduna göre ön işlemeden geçirilmiş ve ANFIS ile işlenmiş haline ait performans kriterlerinin sonuçları Tablo 6.27’de verilmiştir. Tabloda eğitime ve teste ait değerlendirmeler aynı anda verilmiştir.

Beton verisi için tüm metodların eğitime test değerleri, kıyas yapılabilmesi adına tek bir tabloda verilmiştir. Tablo 6.28, çalışmada önerilen ve literatürdeki yaygın normalizasyon metodlarının beton verisini ANFIS ile değerlendirilmesine ait eğitime hatalarının tümünü ait sonuçları göstermektedir. Tablo 6.29, beton verisinin testine ait performans değerlerini göstermektedir. Beton verisi için en başarılı sonuç gerek eğitime ve gerekse test için LBNM metodunda ölçülmüştür.

Tablo 6.27 Z-skör metoduna ait beton verilerinin ANFIS performans sonuçları

Z-skör Metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0.001132	0.070956
R²	0.994715	0.681489

Testte LBNM metodu MSE değeri 0.001 olarak ölçülmüştür. Kendisine en yakın metot olarak, ondalık metottan ölçülen MSE değeri 0.016865'dur ve bu değerden 16.8 kat daha iyi bir sonuç elde edilmiştir. Önerilen metot LBNM, eğitmede ve testte en başarılı metot olarak görülmektedir. R² değeri 0.984 olarak elde edilmiştir.

Tablo 6.28 Tüm normalizasyon metotlarına ait beton verilerinin ANFIS eğitime performans sonuçları

	EBNM	LBNM	Ham	Z-skör	Ondalık	Min.-Mak.
AD	14.416	3.327	11.55066	∞	5.98858	∞
MSE	0	0	18.3362	0.001132	0.000664	0.001132
R²	0.984	0.999	0.988194	0.994715	0.995591	0.994715

Tablo 6.29 Tüm normalizasyon metotlarına ait beton verilerinin ANFIS test performans sonuçları

	EBNM	LBNM	Ham	Z-skör	Ondalık	Min.-Mak.
AD	23.977	9.930	28.05473	∞	18.54774	∞
MSE	0	0.001	467.6429	0.070956	0.016865	0.070956
R²	0.95	0.984	0.707869	0.681489	0.89498	0.681489

6.3.3 Kalp Sistol Verisi

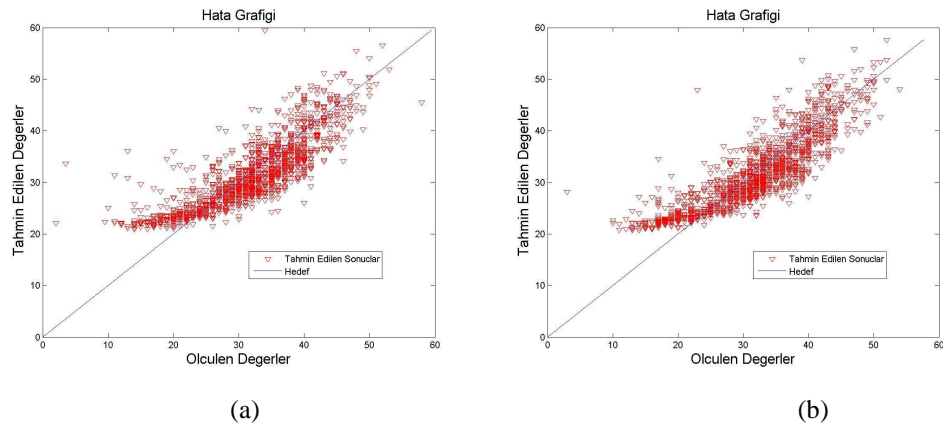
Bu bölümde kalp sistol verisine ait YSA ve ANFIS sonuçlar verilmiştir. Sonuçlar eğitime ve teste ait hata grafiklerini ve eğitime ve testten elde edilen giriş çıkış grafiklerini kapsamaktadır. Eğitime ile test verilerine ait grafikler giriş ile çıkış arasındaki uyumluluğa bağlı olarak diagonal düz bir çizgi üzerine yerleşmelidir. İdeali gösteren düz çizgiye olan yakınsama sonuca ait kaliteyi de göstermiş olur.

6.3.3.1 Kalp sistol verileri YSA sonuçları

Sistol verisine ait YSA sonuçları aşağıda sunulmuştur. Bu sonuçlar önerilen ve literatürde kullanılan metotlarla birlikte verinin ham haline ait performans değerlerini kapsamaktadır.

6.3.3.1.1 Kalp sistol verisi ham verilere ait YSA sonuçları

Sistol verisinin ham haline ait YSA performansına ait hata grafiği Şekil 6.23'de gösterilmiştir. Şekil 6.23 (a) sistol verisinin, eğitime ait hata grafiğini, Şekil 6.23 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.23 Ham kalp sistol verilerinin (a) YSA eğitime hata grafiği (b) YSA test hata grafiği

Tablo 6.30 sistol verisinin YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.

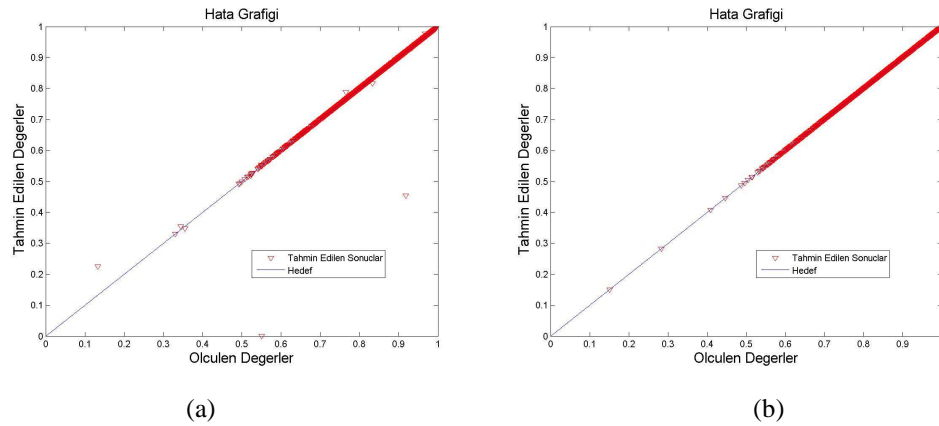
Ham veri değerleri için test ve eğitim R^2 değerlerinin birbirine oldukça yakın değerler elde edilmiştir.

Tablo 6.30 Ham kalp sistol verilerinin YSA performans sonuçları

Ham verilere ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	13.035182	13.050450
MSE	18.694196	18.780512
R^2	0.982790	0.982711

6.3.3.1.2 LBNM metoduna ait kalp sistol verilerinin YSA performans sonuçları

Kalp Sistol verisinin, satır bazında yapılan normalizasyon modelinin YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.24’de gösterilmiştir. Şekil 6.24 (a) sistol verisinin, eğitime ait hata grafiğini, Şekil 6.24 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.24 LBNM metoduna ait kalp sistol verilerinin (a) YSA eğitime hata grafiği (b) YSA test hata grafiği

Tablo 6.31 sistol verisinin YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.

Tablo 6.31 LBNM metoduna ait kalp sistol verilerinin YSA performans sonuçları

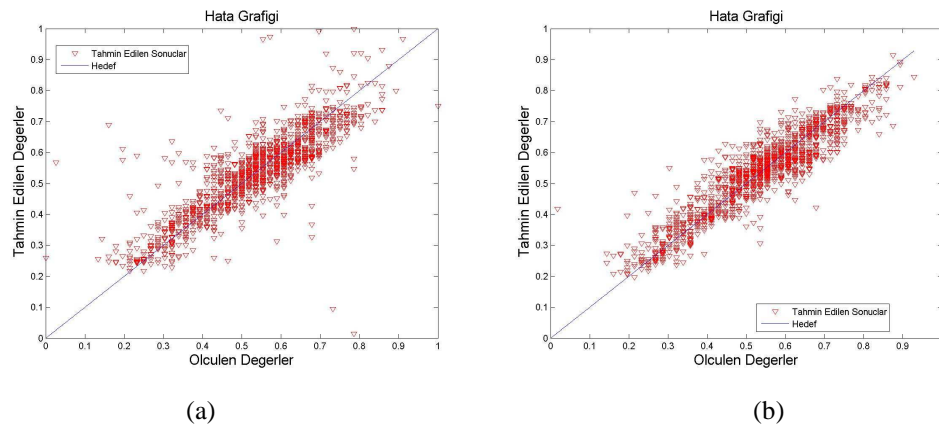
LBNM metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	0.313551	0.520572
MSE	0.000002	0.000072
R^2	0.999991	0.999595

Eğitim ve test değeri sonuçlarının birbirlerine yakın oluşu ve hata değerlerinin oldukça düşük olması satır bazında yapılan normalizasyonun test ve eğitim değerleri oldukça başarılı sonuçlar vermiştir. R^2 değerinin 1'e yakın oluşu bunun bir göstergesidir.

6.3.3.1.3 Minimum-maksimum metoduna ait kalp sistol verilerinin YSA performans sonuçları

Kalp sistol verisinin, minimum-maksimum normalizasyon metodunun YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.25'de gösterilmiştir. Şekil 6.25 (a) sistol verisinin, eğitime ait hata grafiğini, Şekil 6.25 (b) ise teste ait hata grafiğini göstermektedir.

Tablo 6.32 sistol verisinin YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



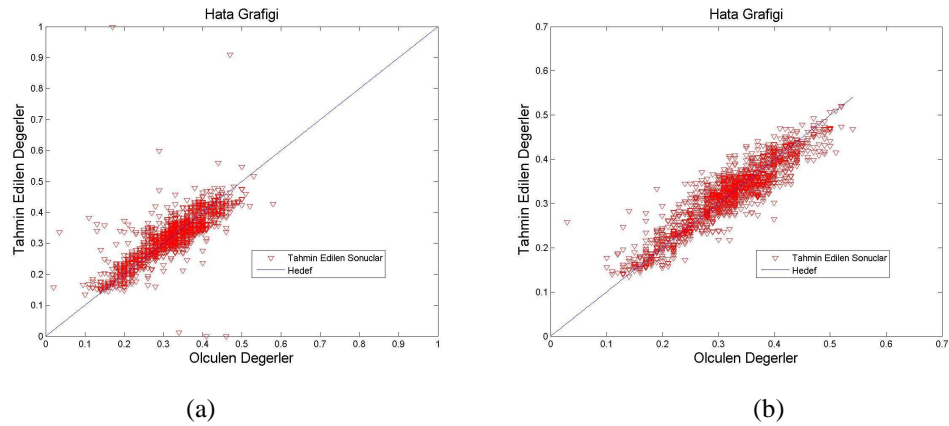
Şekil 6.25 Min.-Mak. Metoduna ait kalp sistol verilerinin (a) YSA eğitime hatası grafiği (b) YSA test hatası grafiği

Tablo 6.32 Min.-Mak. Metoduna ait kalp sistol verilerinin YSA performans sonuçları

Min.-Mak. metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0.003600	0.005858
R ²	0.988270	0.980914

6.3.3.1.4 Kalp sistol verisi ondalık metoda ait YSA performans sonuçları

Kalp sistol verisinin, ondalık normalizasyon metodunun YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.26'da gösterilmiştir. Şekil 6.26 (a) sistol verisinin, eğitime ait hata grafiğini, Şekil 6.26 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.26 Ondalık ölçekleme metoduna ait kalp sistol verilerinin (a) YSA eğitime hatası grafiği (b) YSA test hatası grafiği

Tablo 6.33 sistol verisinin YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.

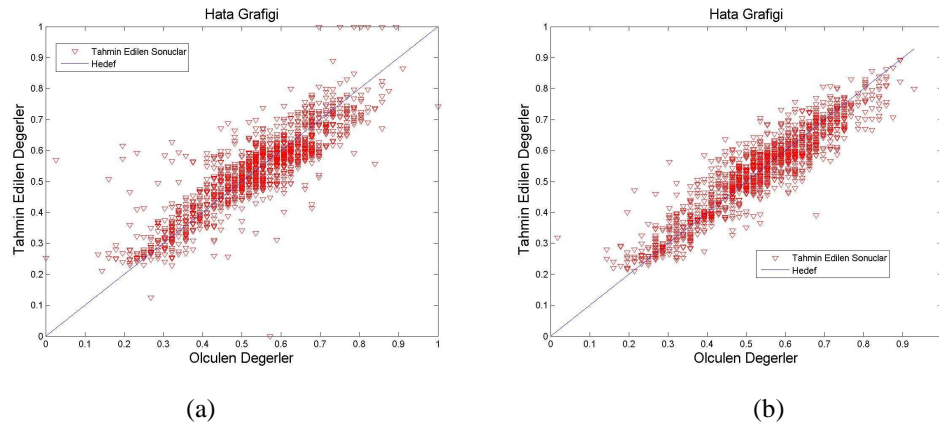
Tablo 6.33 Ondalık ölçekleme metoduna ait kalp sistol verilerinin YSA performans sonuçları

Geliştirilen Metoda İlişkin 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	9.307693	10.79752
MSE	0.001147	0.00213
R ²	0.98944	0.980375

6.3.3.1.5 Z-skor metoduna ait kalp sistol verilerinin YSA performans sonuçları

Sistol verisinin, z-skor ile YSA'nın birlikte kullanılmasına ait hata grafiği Şekil 6.27'de gösterilmiştir. Şekil 6.27 (a) sistol verisinin, eğitime ait hata grafiğini, Şekil 6.27 (b) ise teste ait hata grafiğini göstermektedir.

Tablo 6.34 sistol verisinin YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



Şekil 6.27 Z-skor metoduna ait kalp sistol verilerinin (a) YSA eğitime hatası grafiği (b) YSA test hatası grafiği

Tablo 6.34 Z-skor metoduna ait kalp sistol verilerinin YSA performans sonuçları

Z-skor metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0.003611	0.005101
R ²	0.988232	0.983382

Sistol verisi için tüm metotların eğitime test değerleri, kıyas yapılabilmesi adına tek bir tabloda verilmiştir. Tablo 6.35, çalışmada önerilen ve literatürdeki yaygın normalizasyon metotlarının sistol verisini YSA ile değerlendirilmesine ait eğitime hatalarının tümüne ait sonuçları göstermektedir. Tablo 6.36, sistol verisinin testine ait performans değerlerini göstermektedir. Sistol verisi için en başarılı sonuç gerek eğitime ve gerekse test için LBNM metodunda ölçülmüştür. Testteki başarıda LBNM metodu MSE değeri 0.000072 olarak ölçülmüştür. Kendisine en yakın metot olarak, z-skor metottan ölçülen MSE değeri 0.00213'dur ve bu değerden 29.5 kat daha iyi bir sonuç elde edilmiştir. Önerilen metot LBNM, eğitimde ve testte en başarılı metot olarak görülmektedir. R^2 değerinin 1'e oldukça yakın oluşu bu yöntemin başarılı olduğunun bir göstergesidir.

Tablo 6.35 Tüm normalizasyon metotlarına ait kalp sistol verilerinin YSA eğitime performans sonuçları

	EBNM	LBNM	Ham	Z-skor	Ondalık	Min.-Mak.
AD	14.414596	0.313551	13.035182	∞	9.307693	∞
MSE	0.000011	0.000002	18.694196	0.003611	0.001147	0.003600
R²	0.969955	0.999991	0.982790	0.988232	0.98944	0.988270

Tablo 6.36 Tüm normalizasyon metotlarına ait kalp sistol verilerinin YSA test performans sonuçları

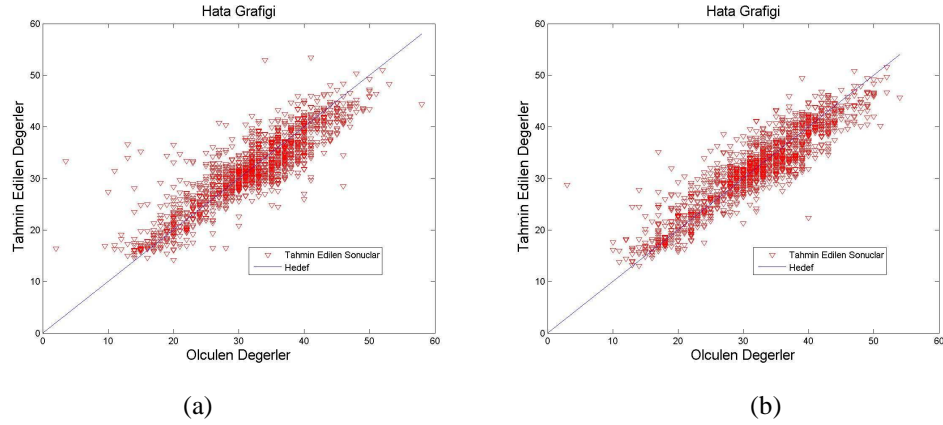
	EBNM	LBNM	Ham	Z-skor	Ondalık	Min.-Mak.
AD	34.408518	0.520572	13.050450	∞	10.79752	∞
MSE	0.002778	0.000072	18.780512	0.005101	0.00213	0.005858
R²	-6.629528	0.999595	0.982711	0.983382	0.980375	0.980914

6.3.3.2 Kalp sistol verisi ANFIS sonuçları

Kalp sistol verisine ait ANFIS sonuçları aşağıda sunulmuştur. Bu sonuçlar önerilen ve literatürde kullanılan metotlarla birlikte verinin ham haline ait performans değerlerini kapsamaktadır.

6.3.3.2.1 Kalp sistol ham verilerine ait ANFIS sonuçları

Kalp sistol verisinin, ham şekliyle ANFIS’de işlenmesine ait hata grafiği Şekil 6.28’de gösterilmiştir. Şekil 6.28 (a) sistol verisinin, eğitime ait hata grafiğini, Şekil 6.28 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.28 Ham kalp sistol verilerinin (a) ANFIS eğitime hata grafiği (b) ANFIS test hata grafiği

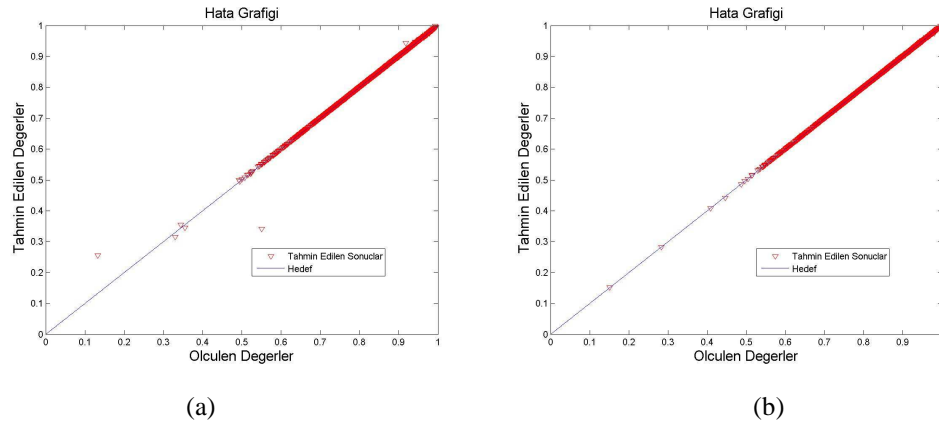
Tablo 6.37 sistol verisinin ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir. Başarı yüzdesi test verileri için %10.43’dür.

Tablo 6.37 Ham kalp sistol verilerinin ANFIS performans sonuçları

Ham verilere ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	10	10.430951
MSE	13.08736	14.413471
R^2	0.98795	0.986729

6.3.3.2.2 LBNM metoduna ait kalp sistol verilerinin ANFIS performans sonuçları

Kalp sistol verisinin, önerilen metot olan LBNM'nin ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.29'da gösterilmiştir. Şekil 6.29 (a) sistol verisinin, eğitime ait hata grafiğini, Şekil 6.29 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.29 LBNM metoduna ait kalp sistol verilerinin (a) ANFIS eğitime hata grafiği (b) ANFIS test hata grafiği

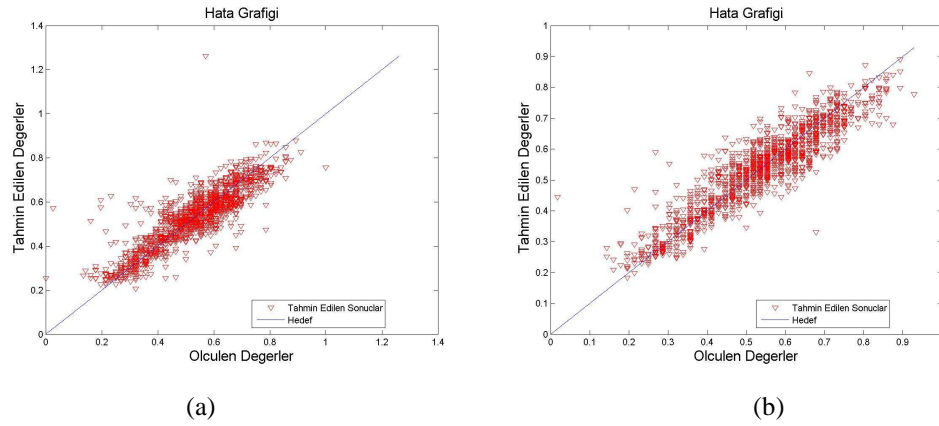
Tablo 6.38 sistol verisinin ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir. Eğitim AD değerininin sıfır oluşu oldukça cesaretlendirici bir sonuçtur ve test değerlerinin de başarılı olması bu yöntemin bir avantajı olarak yorumlanabilir.

Tablo 6.38 LBNM metoduna ait kalp sistol verilerinin ANFIS performans sonuçları

LBNM metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	1.031	0.932
MSE	0	0
R^2	1	1

6.3.3.2.3 Minimum-maksimum metoduna ait kalp sistol verilerinin ANFIS performans sonuçları

Kalp sistol verisinin, minimum-maksimum normalizasyon metodunun ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.30'da gösterilmiştir. Şekil 6.30 (a) sistol verisinin, eğitime ait hata grafiğini, Şekil 6.30 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.30 Min.-Mak. Metoduna ait kalp sistol verilerinin (a) ANFIS eğitime hatası grafiği
(b) ANFIS test hatası grafiği

Tablo 6.39 kalp sistol verisinin ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.

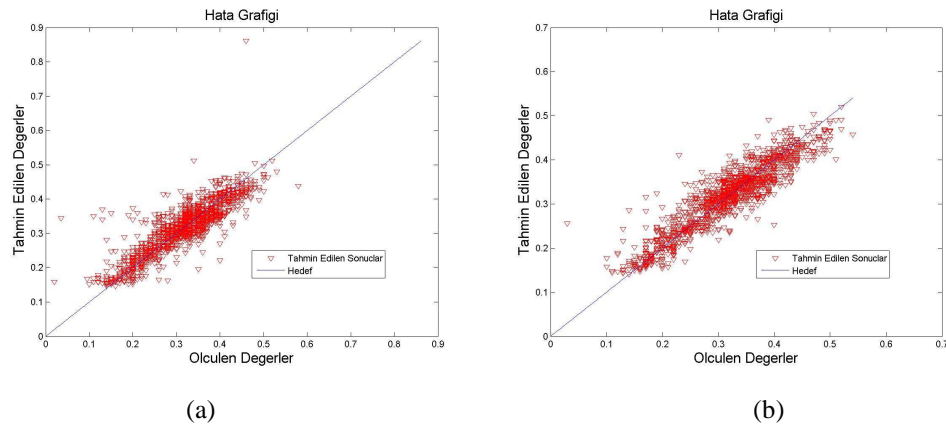
Tablo 6.39 Min.-Mak. Metoduna ait kalp sistol verilerinin ANFIS performans sonuçları

Min.-Mak. metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0.003946	0.004794
R ²	0.98714	0.984380

6.3.3.2.4 Ondalık ölçekleme metoduna ait kalp sistol verilerinin ANFIS performans sonuçları

Sistol verisinin, ondalık normalizasyon ve ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.31’de gösterilmiştir. Şekil 6.31 (a) sistol verisinin, eğitime ait hata grafiğini, Şekil 6.31 (b) ise teste ait hata grafiğini göstermektedir..

Tablo 6.40 sistol verisinin ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



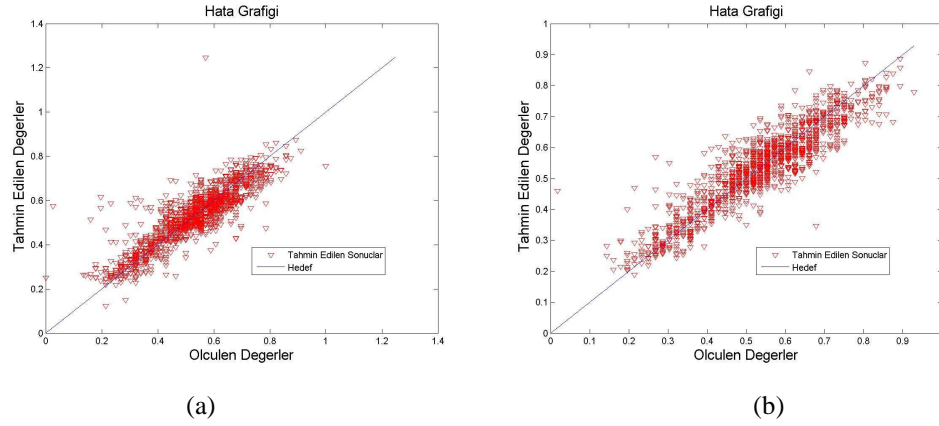
Şekil 6.31 Ondalık ölçekleme metoduna ait kalp sistol verilerinin (a) ANFIS eğitime hatası grafiği (b) ANFIS test hatası grafiği

Tablo 6.40 Ondalık ölçekleme metoduna ait kalp sistol verilerinin ANFIS performans sonuçları

Geliştirilen Metoda İlişkin 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	10	10.273270
MSE	0.001231	0.001490
R ²	0.988665	0.986281

6.3.3.2.5 Z-skor metoduna ait kalp sistol verilerinin ANFIS performans sonuçları

Kalp sistol verisinin, z-skor ve ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.32’de gösterilmiştir.



Şekil 6.32 Z-skor metoduna ait kalp sistol verilerinin (a) ANFIS eğitime hatası grafiği (b) ANFIS test hatası grafiği

Şekil 6.32 (a) sistol verisinin, eğitime ait hata grafiğini, Şekil 6.32 (b) ise teste ait hata grafiğini göstermektedir.

Tablo 6.41 sistol verisinin ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.

Tablo 6.41 Z-skor metoduna ait kalp sistol verilerinin ANFIS performans sonuçları

Z-skor metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0	0.004729
R ²	1	0.984592

Kalp sistol verisi için tüm metotların eğitime test değerleri, kıyas yapılabilmesi adına tek bir tabloda verilmiştir. Tablo 6.42, çalışmada önerilen ve literatürdeki yaygın normalizasyon metotlarının sistol verisini ANFIS ile değerlendirilmesine ait eğitime hatalarının tümünü ait sonuçları göstermektedir. Tablo 6.43, sistol verisinin testine ait performans değerlerini göstermektedir.

Kalp sistol verisi için, satır tabanlı normalizasyon metodunun başarılı sonuçlar elde ettiği Tablo 6.42 ve Tablo 6.43'den görülmektedir. Öklid uzaklığının normalizasyon olarak öneren metod hemsatır da hemde sütunda başarılı sonuçlar vermiştir. AD değeri 0.932 satır tabanlı ve 9.75 sütun tabanlı olmak üzere, R^2 değeri her iki uygulamanın testinde de 0.988 sütun tabanlı ve 1 satır tabanlı olmak üzere diğer normalizasyon metodlarından daha iyi olarak elde edilmiştir.

Tablo 6.42 Tüm normalizasyon metodlarına ait kalp sistol verilerinin ANFIS eğitme performans sonuçları

	EBNM	LBNM	Ham	Z-skor	Ondalık	Min.-Mak.
AD	10.262	1.031	10	∞	10	∞
MSE	0	0	13.08736	0	0.001231	0.003946
R²	0.987	1	0.98795	1	0.988665	0.98714

Tablo 6.43 Tüm normalizasyon metodlarına ait kalp sistol verilerinin ANFIS test performans sonuçları

	EBNM	LBNM	Ham	Z-skor	Ondalık	Min.-Mak.
AD	9.75	0.932	10.430951	∞	10.273270	∞
MSE	0	0	14.413471	0.004729	0.001490	0.004794
R²	0.988	1	0.986729	0.984592	0.986281	0.984380

6.3.4 Diş anterior temporal verisi

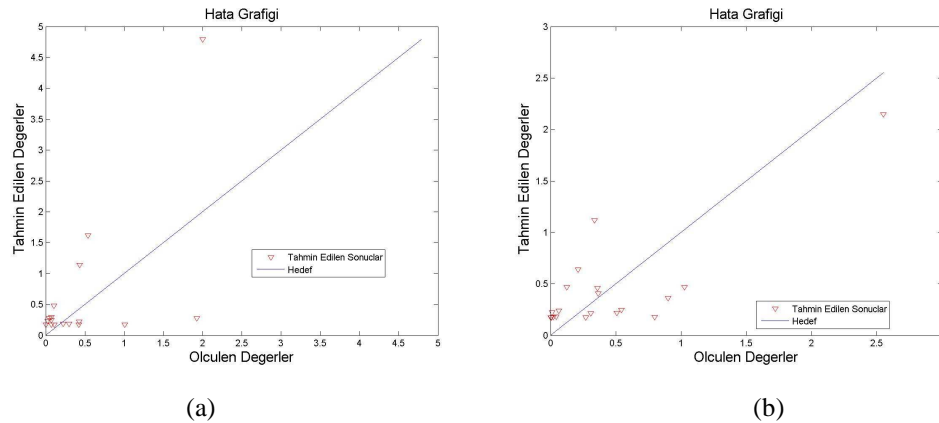
Bu bölümde diş anterior temporal verisine ait YSA ve ANFIS sonuçlar verilmiştir. Sonuçlar eğitme ve teste ait hata grafiklerini ve eğitme ve testten elde edilen giriş çıkış grafiklerini kapsamaktadır. Eğitme ile test verilerine ait grafikler giriş ile çıkış arasındaki uyumluluğa bağlı olarak diagonal düz bir çizgi üzerine yerleşmelidir. İdeali gösteren düz çizgiye olan yakınsama sonuca ait kaliteyi de göstermiş olur.

6.3.4.1 Diş anterior temporal verileri YSA sonuçları

Diş anterior temporal verisine ait YSA sonuçları aşağıda sunulmuştur. Bu sonuçlar önerilen ve literatürde kullanılan metotlarla birlikte verinin ham haline ait performans değerlerini kapsamaktadır.

6.3.4.1.1 Ham Diş anterior temporal verilerine ait YSA sonuçları

Diş anterior temporal verisinin, ham hali ile YSA'nın birlikte kullanılmasına ait hata grafiği Şekil 6.45'de gösterilmiştir. Şekil 6.45 (a) diş anterior temporal verisinin, eğitime ait hata grafiğini, Şekil 6.45 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.33 Ham diş anterior temporal verilerinin (a) YSA eğitime hata grafiği (b) YSA test hata grafiği

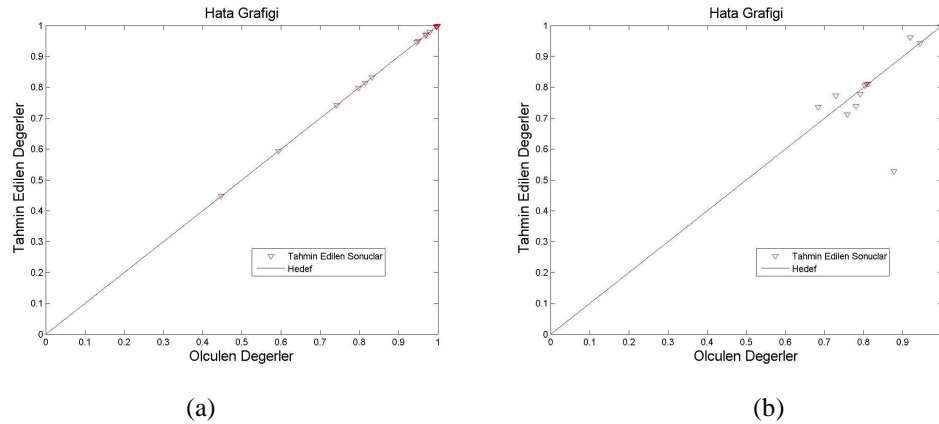
Tablo 6.44 diş anterior temporal verisinin YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.

Tablo 6.44 Ham dış anterior temporal verilerinin YSA performans sonuçları

Ham verilere ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	48124.8	51750.22
MSE	0.161188	0.443984
R^2	0.674346	0.097622

6.3.4.1.2 LBNM metoduna ait dış anterior temporal verilerinin YSA performans sonuçları

Diş anterior temporal verisinin, önerilen metot olan LBNM'nin YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.34'de gösterilmiştir. Şekil 6.34 (a) dış anterior temporal verisinin, eğitime ait hata grafiğini, Şekil 6.34 (b) ise teste ait hata grafiğini göstermektedir



Şekil 6.34 LBNM metoduna ait dış anterior temporal verilerinin (a) YSA eğitime hata grafiği
(b) YSA test hata grafiği

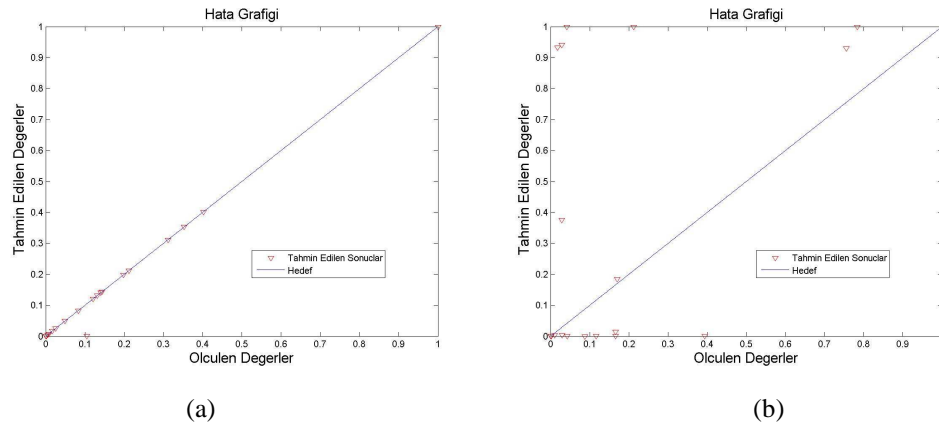
Tablo 6.46 dış anterior temporal verisinin YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.

Tablo 6.45 LBNM metoduna ait dış anterior temporal verilerinin YSA performans sonuçları

LBNM metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	0.08293	16.768094
MSE	0.000008	0.03263
R ²	0.999988	0.953497

6.3.4.1.3 Minimum-maksimum metoduna ait dış anterior temporal verilerinin YSA performans sonuçları

Diş anterior temporal verisinin, minimum-maksimum normalizasyon ve YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.35’de gösterilmiştir. Şekil 6.35 (a) dış anterior temporal verisinin, eğitime ait hata grafiğini, Şekil 6.35 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.35 Min.-Mak. metoduna ait dış anterior temporal verilerinin (a) YSA eğitime hatası grafiği (b) YSA test hatası grafiği

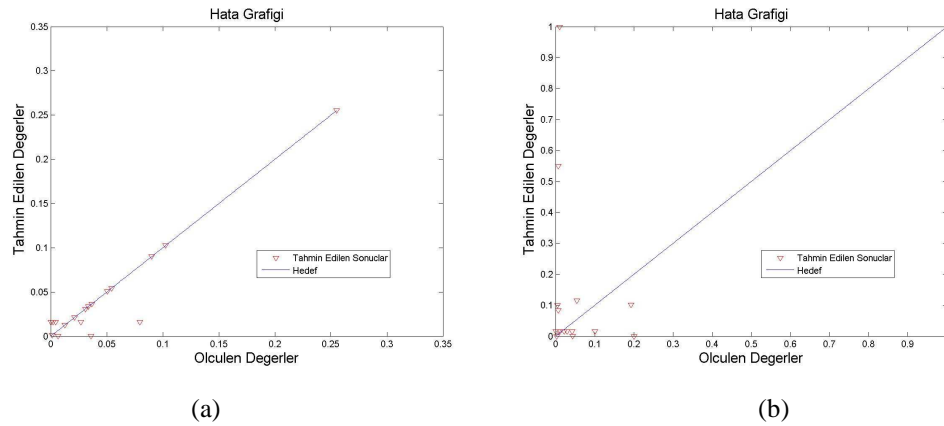
Tablo 6.46 diş anterior temporal verisinin YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.

Tablo 6.46 Dış anterior temporal verisi Min.-Mak. metoduna ait YSA performans sonuçları

Min.-Mak. metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0.021653	0.104561
R ²	0.710651	-0.39032

6.3.4.1.4 Ondalık ölçekleme metoduna ait dış anterior temporal verilerinin YSA performans sonuçları

Dış anterior temporal verisinin, ondalık normalizasyon ve YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.36'da gösterilmiştir. Şekil 6.36 (a) dış anterior temporal verisinin, eğitime ait hata grafiğini, Şekil 6.36 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.36 Ondalık ölçeklemeye ait dış anterior temporal verilerinin (a) YSA eğitime hatası grafiği (b) YSA test hatası grafiği

Tablo 6.47 dış anterior temporal verisinin YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.

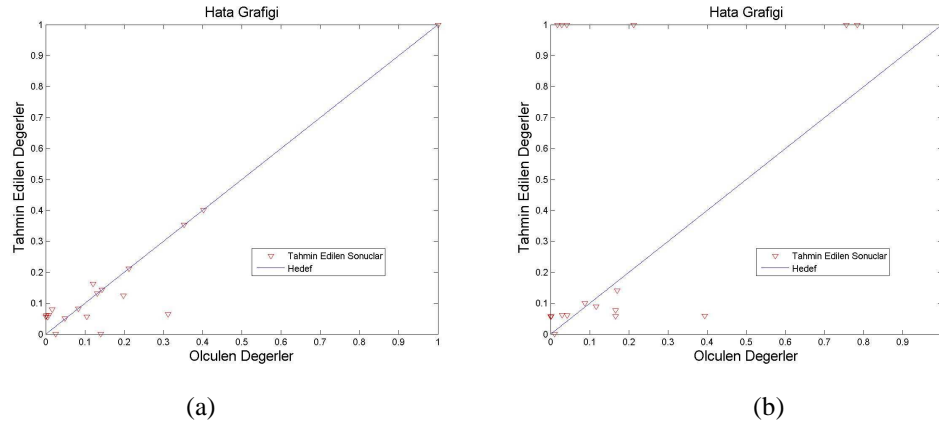
Tablo 6.47 Ondalık ölçekleme metoduna ait dış anterior temporal verilerinin YSA performans sonuçları

Geliştirilen Metoda İlişkin 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	36470.51	35899.43
MSE	0.001382	0.036226
R ²	0.717628	-6.41818

6.3.4.1.5 Z-skor metoduna ait dış anterior temporal verilerinin YSA performans sonuçları

Dış anterior temporal verisinin, z-skor normalizasyonu ve YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.37’de gösterilmiştir. Şekil 6.37 (a) dış anterior temporal verisinin, eğitime ait hata grafiğini, Şekil 6.37 (b) ise teste ait hata grafiğini göstermektedir.

Tablo 6.48 dış anterior temporal verisinin YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



Şekil 6.37 Z-skor metoduna ait dış anterior temporal verilerinin (a) YSA eğitime hatası grafiği
(b) YSA test hatası grafiği

Tablo 6.48 Z-skor metoduna ait dış anterior temporal verilerinin YSA performans sonuçları

Z-skor metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	-1279394	21310.21939
MSE	0.003876	0.163903
R²	0.949737	-1.164496

Diş anterior temporal verisi için tüm metotların eğitime test değerleri, kıyas yapılabilmesi adına tek bir tabloda verilmiştir. Tablo 6.49, çalışmada önerilen ve literatürdeki yaygın normalizasyon metotlarının dış anterior temporal verisini YSA ile değerlendirilmesine ait eğitime hatalarının tümünü ait sonuçları göstermektedir. Tablo 6.50, dış anterior temporal verisinin testine ait performans değerlerini göstermektedir.

Tablo 6.49 Tüm normalizasyon metotlarına ait dış anterior temporal verilerinin YSA eğitime performans sonuçları

	EBNM	LBNM	Ham	Z-skor	Ondalık	Min.-Mak.
AD	118501.56	0.08293	48124.8	-1279394	36470.51	∞
MSE	0.003018	0.000008	0.161188	0.003876	0.001382	0.021653
R²	0.857714	0.999988	0.674346	0.949737	0.717628	0.710651

Tablo 6.50 Tüm normalizasyon metotlarına ait dış anterior temporal verilerinin YSA test performans sonuçları

	EBNM	LBNM	Ham	Z-skor	Ondalık	Min.-Mak.
AD	282283.6	16.768094	51750.22	21310.21939	35899.43	∞
MSE	0.049949	0.032603	0.443984	0.163903	0.036226	0.104561
R²	-0.702878	0.953497	0.097622	-1.164496	-6.41818	-0.39032

Diş anterior temporal verisi için satır tabanlı normalizasyonu ham test hem de eğitime için başarılı sonuçlar verdiği görülmektedir. MSE değeri LBNM metodu için 0.032603 hesaplanmış ve minimum-maksimum modeli için hesaplanan 0.104561 değerinden 3.2 kat daha iyi bir değerdir. Bununla birlikte R² değerinin 1'e en yakın

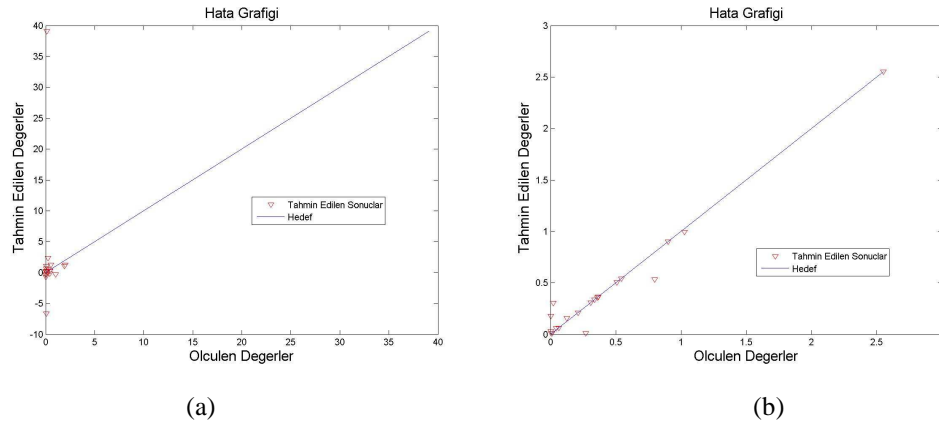
olduğu metodun satır tabanlı normalizasyona ait olduğu da dikkate değer bir sonuçtur.

6.3.4.2 Diş anterior temporal verisi ANFIS sonuçları

Diş anterior temporal verisine ait ANFIS sonuçları aşağıda sunulmuştur. Bu sonuçlar önerilen ve literatürde kullanılan metotlarla birlikte verinin ham haline ait performans değerlerini kapsamaktadır.

6.3.4.2.1 Ham diş anterior temporal verilerine ait ANFIS sonuçları

Diş anterior temporal verisinin, ham halinin ANFIS ile işlenmesine ait hata grafiği Şekil 6.38'de gösterilmiştir. Şekil 6.38 (a) diş anterior temporal verisinin, eğitime ait hata grafiğini, Şekil 6.38 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.38 Ham diş anterior temporal verilerinin (a) ANFIS eğitime hata grafiği (b) ANFIS test hata grafiği

Tablo 6.51 diş anterior temporal verisinin ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.

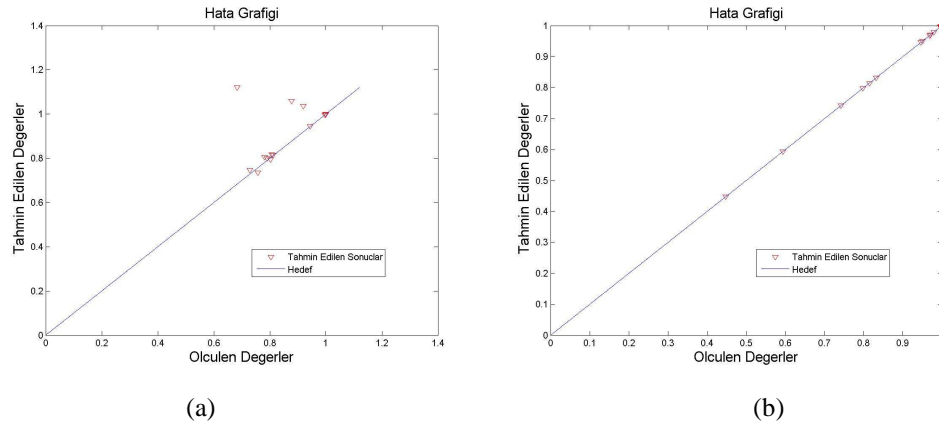
Tablo 6.51 Ham dış anterior temporal verilerinin ANFIS performans sonuçları

Ham verilere ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	30745.05071	32058.39
MSE	0.025937	45.93402
R^2	0.947281	-92.7562

6.3.4.2.2 LBNM metoduna ait dış anterior temporal verilerinin ANFIS performans sonuçları

Diş anterior temporal verisinin, satır tabanlı olarak normalize edilmesi ile ANFIS'in birlikte kullanılmasına ait hata grafiği Şekil 6.39'da gösterilmiştir. Şekil 6.39 (a) diş anterior temporal verisinin, eğitime ait hata grafiğini, Şekil 6.39 (b) ise teste ait hata grafiğini göstermektedir.

Tablo 6.52 diş anterior temporal verisinin ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



Şekil 6.39 LBNM metoduna ait dış anterior temporal verilerinin (a) ANFIS eğitime hata grafiği (b) ANFIS test hata grafiği

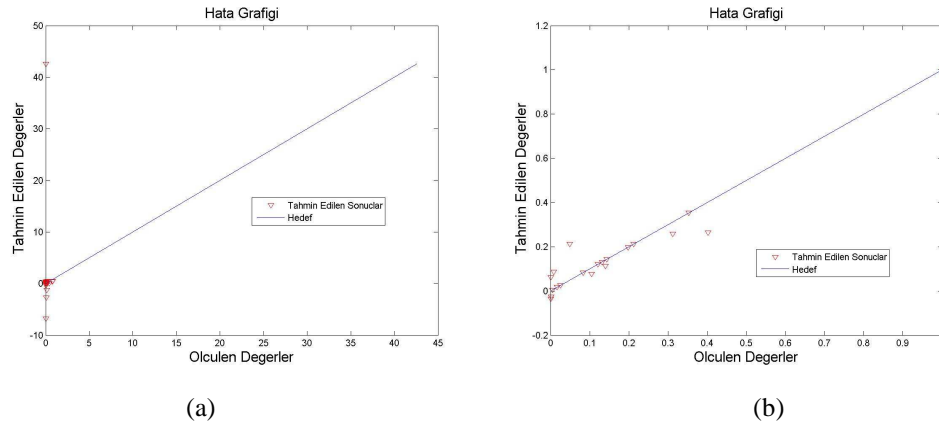
Tablo 6.52 LBNM metoduna ait dış anterior temporal verilerinin ANFIS performans sonuçları

LBNM metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	0.361	0.501
MSE	0	0
R^2	1	1

6.3.4.2.3 Dış anterior temporal verisi minimum-maksimum metoduna ait ANFIS performans sonuçları

Dış anterior temporal verisinin, minimum-maksimum normalizasyon metodu ile ANFIS'in birlikte kullanılmasına ait hata grafiği Şekil 6.40'da gösterilmiştir. Şekil 6.40 (a) dış anterior temporal verisinin, eğitime ait hata grafiğini, Şekil 6.40 (b) ise teste ait hata grafiğini göstermektedir.

Tablo 6.54 dış anterior temporal verisinin ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



Şekil 6.40 Min.-Mak. Metoduna ait dış anterior temporal verilerinin (a) ANFIS eğitime hatası grafiği (b) ANFIS test hatası grafiği

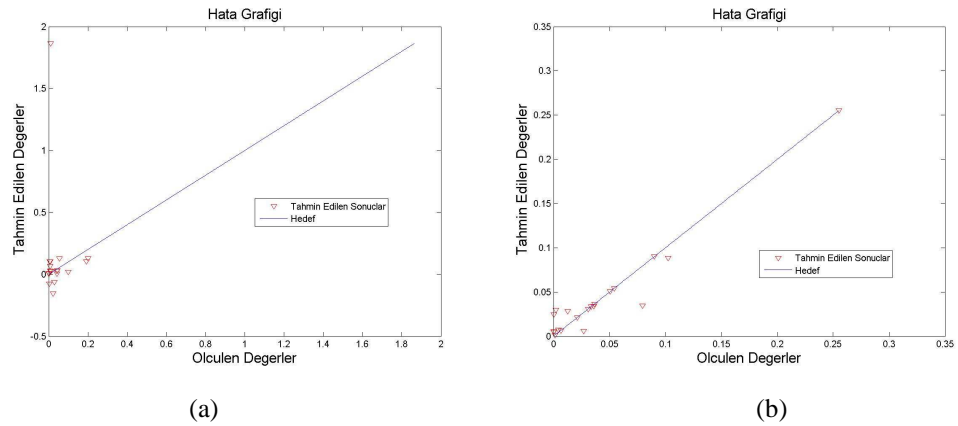
Tablo 6.53 Min.-Mak. metoduna ait dış anterior temporal verilerinin ANFIS performans sonuçları

Min.-Mak. metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0.006531	49.752765
R^2	0.913566	-662.491060

6.3.4.2.4 Ondalık ölçekleme metoduna ait dış anterior temporal verilerinin ANFIS performans sonuçları

Diş anterior temporal verisinin, ondalık normalizasyon ve ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.41’de gösterilmiştir. Şekil 6.41 (a) diş anterior temporal verisinin, eğitime ait hata grafiğini, Şekil 6.41 (b) ise teste ait hata grafiğini göstermektedir.

Tablo 6.54 diş anterior temporal verisinin ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



Şekil 6.41 Ondalık ölçeklemeye ait dış anterior temporal verilerinin (a) ANFIS eğitime hatası grafiği (b) ANFIS test hatası grafiği

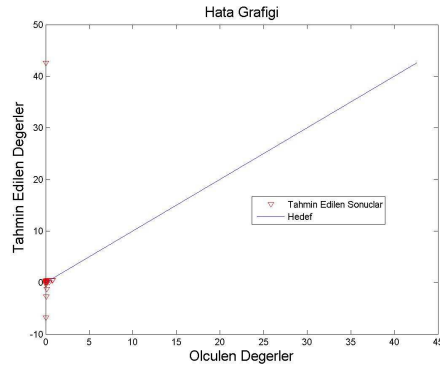
Tablo 6.54 Ondalık ölçekleme metoduna ait dış anterior temporal verilerinin ANFIS performans sonuçları

Geliştirilen Metoda İlişkin 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	11707.07	20263.62
MSE	0.000365	145.0873
R ²	0.925916	-28535.7

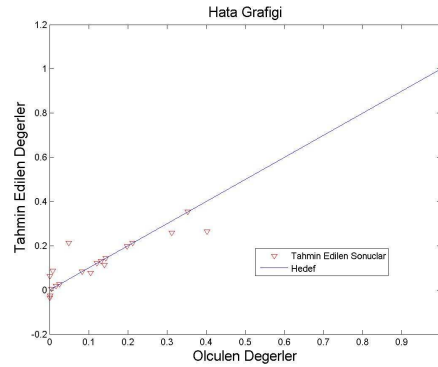
6.3.4.2.5 Z-skor metoduna ait dış anterior temporal verilerinin ANFIS performans sonuçları

Diş anterior temporal verisinin, z-skor normalizasyon metodunun ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.42’de gösterilmiştir. Şekil 6.42 (a) diş anterior temporal verisinin, eğitime ait hata grafiğini, Şekil 6.42 (b) ise teste ait hata grafiğini göstermektedir.

Tablo 6.55 diş anterior temporal verisinin ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



(a)



(b)

Şekil 6.42 Z-skor metoduna ait dış anterior temporal verilerinin (a) ANFIS eğitime hatası grafiği (b) ANFIS test hatası grafiği

Tablo 6.55 Z-skor metoduna ait dış anterior temporal verilerinin ANFIS performans sonuçları

Z-skor metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	-8.0895E+14	-1.49793E+15
MSE	0.004861	59.812063
R²	0.935887	-791.306006

Diş anterior temporal verisi için tüm metotların eğitime test değerleri, kıyas yapılabilmesi adına tek bir tabloda verilmiştir. Tablo 6.56, çalışmada önerilen ve literatürdeki yaygın normalizasyon metotlarının dış anterior temporal verisini ANFIS ile değerlendirilmesine ait eğitime hatalarının tümünü ait sonuçları göstermektedir. Tablo 6.57, dış anterior temporal verisinin testine ait performans değerlerini göstermektedir.

Diş anterior temporal verisi için en başarılı sonuç gerek eğitime ve gerekse test için satır tabanlı normalizasyon metodunda ölçülmüştür. Testteki başarıda LBNM'nin MSE değeri sıfır ve R² değeri 1 olarak hesaplanmıştır ki bu limit değerlerdir.

Tablo 6.56 Tüm normalizasyon metotlarına ait dış anterior temporal verilerinin ANFIS eğitime performans sonuçları

	EBNM	LBNM	Ham	Z-skor	Ondalık	Min.-Mak.
AD	8575266	0.361	30745.05	-8.08E+14	11707.07	∞
MSE	0.021	0	0.025937	0.004861	0.000365	0.00653
R²	0.408	1	0.947281	0.935887	0.925916	0.9135

Tablo 6.57 Tüm normalizasyon metotlarına ait dış anterior temporal verilerinin ANFIS test performans sonuçları

ANFIS	EBNM	LBNM	Ham	Z-skor	Ondalık	Min.-Mak.
AD	381636.2	0.501	32058.39	-1.497E+15	20263.62	∞
MSE	0.013	0	45.93402	59.812063	145.0873	49.752765
R²	0.087	1	-92.7562	-791.3060	-28535.7	-662.4910

6.3.5 Diş üst dudak verisi

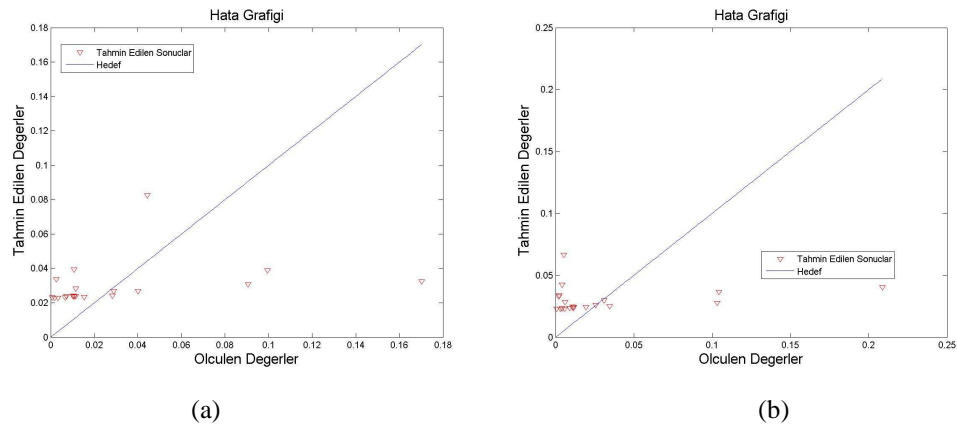
Bu bölümde diş üst dudak verisine ait YSA ve ANFIS sonuçlar verilmiştir. Sonuçlar eğitime ve teste ait hata grafiklerini ve eğitime ve testten elde edilen giriş çıkış grafiklerini kapsamaktadır. Eğitime ile test verilerine ait grafikler giriş ile çıkış arasındaki uyumluluğa bağlı olarak diagonal düz bir çizgi üzerine yerleşmelidir. İdeali gösteren düz çizgiye olan yakınsama sonuca ait kaliteyi de göstermiş olur.

6.3.5.1 Diş üst dudak verisi YSA sonuçları

Diş üst dudak verisine ait YSA sonuçları aşağıda sunulmuştur. Bu sonuçlar önerilen ve literatürde kullanılan metotlarla birlikte verinin ham haline ait performans değerlerini kapsamaktadır.

6.3.5.1.1 Ham diş üst dudak verilerine ait YSA sonuçları

Diş üst dudak verisinin, ham değerlerinin YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.43'de gösterilmiştir. Şekil 6.43 (a) diş üst dudak verisinin, eğitime ait hata grafiğini, Şekil 6.43 (b) ise teste ait hata grafiğini göstermektedir.



Şekil 6.43 Ham diş üst dudak verilerinin (a) YSA eğitime hata grafiği (b) YSA test hata grafiği

Tablo 6.58 dış üst dudak verisinin YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.

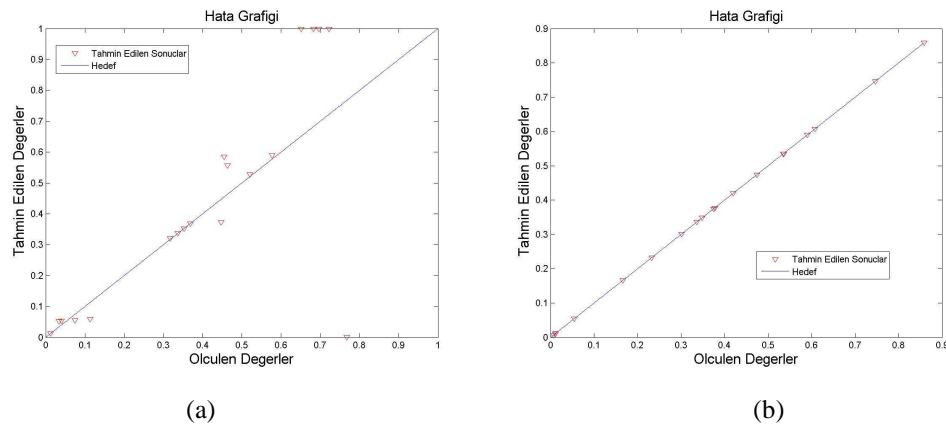
Tablo 6.58 Ham dış üst dudak verilerinin YSA performans sonuçları

Ham verilere ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	486.5085	481.7434
MSE	0.002030	0.002033
R^2	0.342096	0.341132

6.3.5.1.2 LBNM metoduna ait dış üst dudak verilerinin YSA performans sonuçları

Dış üst dudak verisinin, önerilen metot olan satır tabanlı metot ve YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.44’de gösterilmiştir. Şekil 6.44 (a) dış üst dudak verisinin, eğitime ait hata grafiğini, Şekil 6.44 (b) ise teste ait hata grafiğini göstermektedir.

Tablo 6.59 dış üst dudak verisinin YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



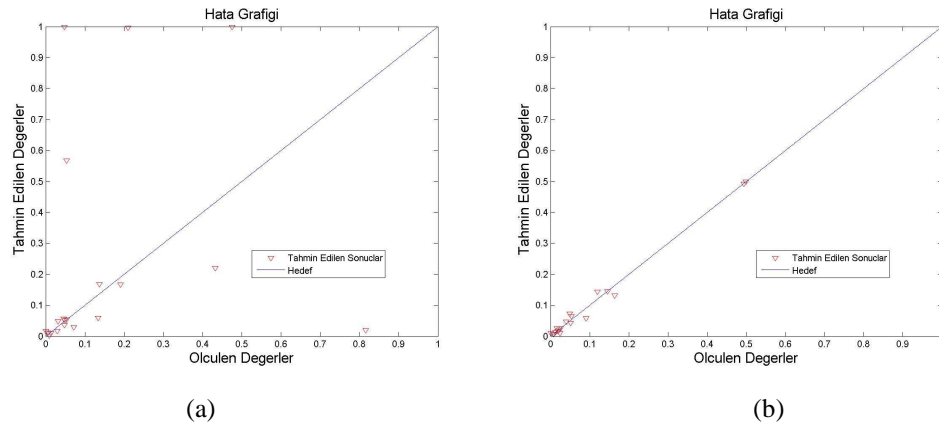
Şekil 6.44 LBNM metoduna ait dış üst dudak verilerinin (a) YSA eğitime hata grafiği (b) YSA test hata grafiği

Tablo 6.59 LBNM metoduna ait diş üst dudak verilerinin YSA performans sonuçları

LBNM metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	0.002775	9.586834
MSE	0	0.01888
R ²	1	0.969922

6.3.5.1.3 Minimum-maksimum metoduna ait diş üst dudak verilerinin YSA performans sonuçları

Diş üst dudak verisinin, minimum-maksimum normalizasyon metodu ve YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.45’de gösterilmiştir. Şekil 6.45 (a) diş üst dudak verisinin, eğitime ait hata grafiğini, Şekil 6.45 (b) ise teste ait hata grafiğini göstermektedir. Tablo 6.60 diş üst dudak verisinin YSA ile işlenmesine ait eğitim ve teste ait performans sonuçlarını göstermektedir.



Şekil 6.45 Min.-Mak. Metoduna ait diş üst dudak verilerinin (a) YSA eğitme hatası grafiği

(b) YSA test hatası grafiği

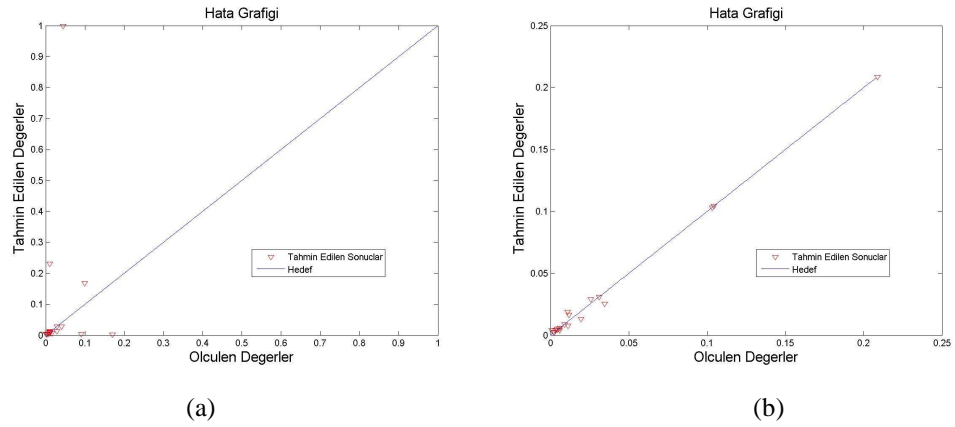
Tablo 6.60 Min.-Mak. Metoduna ait diş üst dudak verilerinin YSA performans sonuçları

Min.-Mak. metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0.000161	0.121518
R ²	0.997721	-0.801131

6.3.5.1.4 Ondalık ölçekleme metoduna ait dış üst dudak verilerinin YSA performans sonuçları

Diş üst dudak verisinin, ondalık normalizasyon ve YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.46'da gösterilmiştir. Şekil 6.46 (a) diş üst dudak verisinin, eğitime ait hata grafiğini, Şekil 6.46 (b) ise teste ait hata grafiğini göstermektedir.

Tablo 6.61 diş üst dudak verisinin YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



Şekil 6.46 Ondalık ölçekleme metoduna ait dış üst dudak verilerinin (a) YSA eğitime hatası grafiği (b) YSA test hatası grafiği

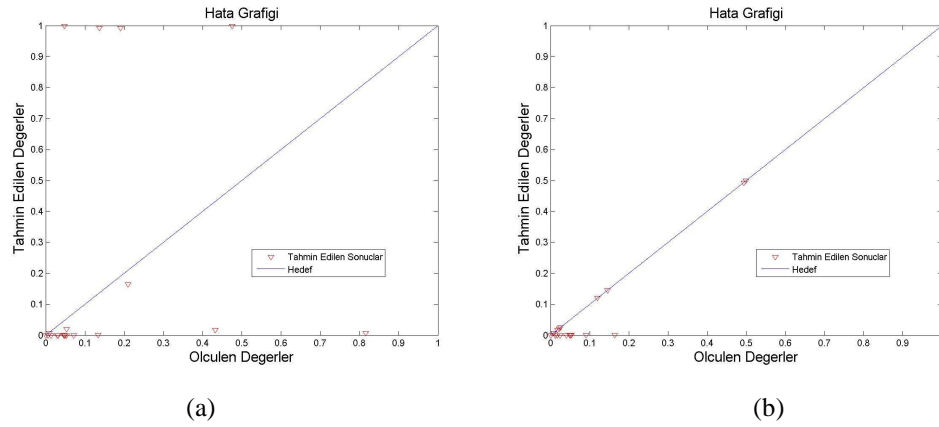
Tablo 6.61 Ondalık ölçekleme metoduna ait dış üst dudak verilerinin YSA performans sonuçları

Geliştirilen Metoda İlişkin 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	43.87655	828.2770
MSE	0.000008	0.119619
R ²	0.997455	-35.83238

6.3.5.1.5 Z-skor metoduna ait diş üst dudak verilerinin YSA performans sonuçları

Diş üst dudak verisinin, z-skor ile normalizasyonu ve YSA ile birlikte kullanılmasına ait hata grafiği Şekil 6.47’de gösterilmiştir. Şekil 6.47 (a) diş üst dudak verisinin, eğitime ait hata grafiğini, Şekil 6.47 (b) ise teste ait hata grafiğini göstermektedir

Tablo 6.62 diş üst dudak verisinin YSA ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



Şekil 6.47 Z-skor metoduna ait diş üst dudak verilerinin (a) YSA eğitime hatası grafiği (b) YSA test hatası grafiği

Tablo 6.62 Z-skor metoduna ait diş üst dudak verilerinin YSA performans sonuçları

Z-skor metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	59239243	625.8914
MSE	0.001161	0.251735
R ²	0.985147	-2.517934

Diş üst dudak verisi için tüm metotların eğitime test değerleri, kıyas yapılabilmesi adına tek bir tabloda verilmiştir. Tablo 6.63, çalışmada önerilen ve literatürdeki yaygın normalizasyon metotlarının diş üst dudak verisini YSA ile

değerlendirilmesine ait eğitme hatalarının tümünü ait sonuçları göstermektedir. Tablo 6.64, dış üst dudak verisinin testine ait performans değerlerini göstermektedir.

Tablo 6.63 Tüm normalizasyon metotlarına ait dış üst dudak verilerinin YSA eğitme performansı sonuçları

	EBNM	LBNM	Ham	Z-skör	Ondalık	Min.-Mak.
AD	60.324946	0.002775	486.5085	59239243	43.87655	∞
MSE	0.019345	0	0.002030	0.001161	0.000008	0.000161
R²	0.499579	1	0.342096	0.985147	0.997455	0.997721

Tablo 6.64 Tüm normalizasyon metotlarına ait dış üst dudak verilerinin YSA test performansı sonuçları

	EBNM	LBNM	Ham	Z-skör	Ondalık	Min.-Mak.
AD	517.1249	9.586834	481.7434	625.8914	828.2770	∞
MSE	0.101177	0.01888	0.002033	0.251735	0.119619	0.121518
R²	-1.969359	0.96992	0.341132	-2.517934	-35.8323	-0.801131

Dış üst dudak verisi için en başarılı sonuç satır tabanlı normalizasyon uygulandığı zaman hesaplanmıştır. Testteki başarıda LBNM metodunun MSE değeri 0.0188 olarak hesaplanmış ve bu değer kendisine en yakın metot olarak, ondalık metottan ölçülen MSE değeri 0.119619 değerinden 6.3 kat daha iyi bir değerdir.

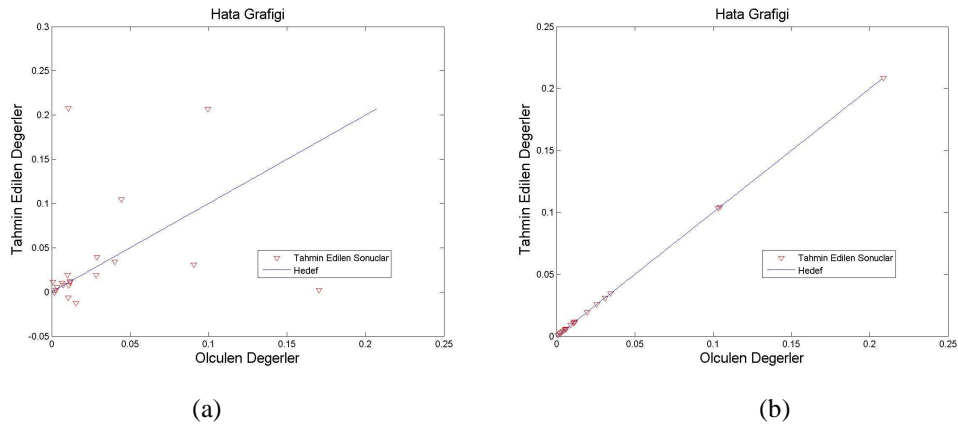
6.3.5.2 Dış üst dudak verisi ANFIS sonuçları

Dış üst dudak verisine ait ANFIS sonuçları aşağıda sunulmuştur. Bu sonuçlar önerilen ve literatürde kullanılan metotlarla birlikte verinin ham haline ait performans değerlerini kapsamaktadır.

6.3.5.2.1 Ham diş üst dudak ham verilerine ait ANFIS sonuçları

Diş üst dudak verisinin, ham değerlerinin ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.48’de gösterilmiştir. Şekil 6.48 (a) diş üst dudak verisinin, eğitime ait hata grafiğini, Şekil 6.48 (b) ise teste ait hata grafiğini göstermektedir.

Tablo 6.65 diş üst dudak verisinin ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



Şekil 6.48 Ham diş üst dudak verilerinin (a) ANFIS eğitime hata grafiği (b) ANFIS test hata grafiği

Tablo 6.65 Ham diş üst dudak verilerinin ANFIS performans sonuçları

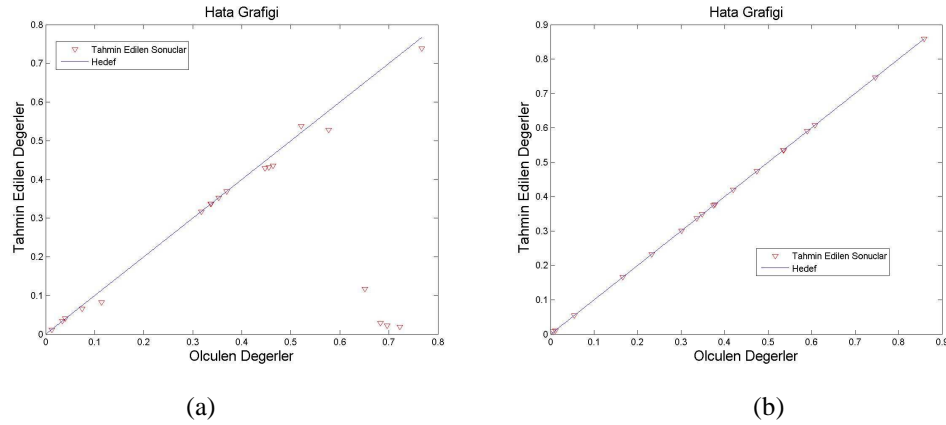
Ham verilere ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	19.366631	480.945240
MSE	0.000001	0.006240
R^2	0.999531	-0.995212

6.3.5.2.2 LBNM metoduna ait diş üst dudak verilerinin ANFIS performans sonuçları

Diş üst dudak verisinin, önerilen metot olan satır tabanlı normalizasyon ve ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.49’da gösterilmiştir. Şekil

6.49 (a) diş üst dudak verisinin, eğitime ait hata grafiğini, Şekil 6.49 (b) ise teste ait hata grafiğini göstermektedir.

Tablo 6.66 diş üst dudak verisinin ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



Şekil 6.49 LBNM metoduna ait diş üst dudak verilerinin (a) ANFIS eğitime hata grafiği (b) ANFIS test hata grafiği

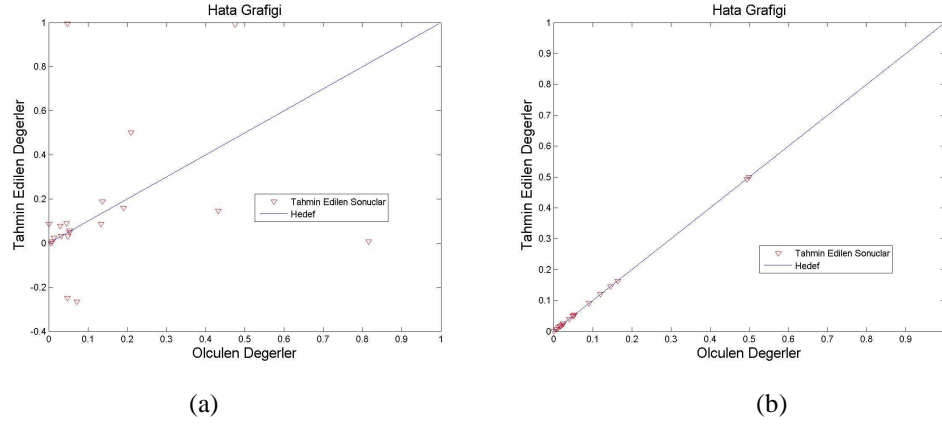
Tablo 6.66 LBNM metoduna ait diş üst dudak verilerinin ANFIS performans sonuçları

LBNM metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	12.365	11.655
MSE	0.001	0.004
R^2	0.999	0.994

6.3.5.2.3 Minimum-maksimum metoduna ait diş üst dudak verilerinin ANFIS performans sonuçları

Diş üst dudak verisinin, minimum-maksimum normalizasyon metodu ve ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.50’de gösterilmiştir. Şekil 6.50 (a) diş üst dudak verisinin, eğitime ait hata grafiğini, Şekil 6.50 (b) ise teste ait hata grafiğini göstermektedir.

Tablo 6.67 diş üst dudak verisinin ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



Şekil 6.50 Min.-Mak. Metoduna ait diş üst dudak verilerinin (a) ANFIS eğitime hatası grafiği
(b) ANFIS test hatası grafiği

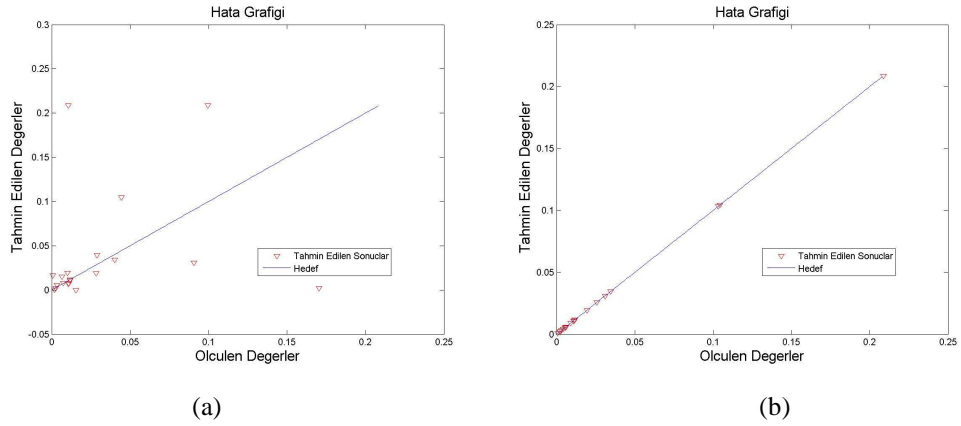
Tablo 6.67 Min.-Mak. Metoduna ait diş üst dudak verilerinin ANFIS performans sonuçları

Min.-Mak. metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	∞	∞
MSE	0.000015	0.246590
R^2	0.999751	-2.338624

6.3.5.2.4 Ondalık ölçekleme metoduna ait diş üst dudak verilerinin ANFIS performans sonuçları

Diş üst dudak verisinin, ondalık normalizasyon ve ANFIS'in birlikte kullanılmasına ait hata grafiği Şekil 6.51'de gösterilmiştir. Şekil 6.51 (a) diş üst dudak verisinin, eğitime ait hata grafiğini, Şekil 6.51 (b) ise teste ait hata grafiğini göstermektedir.

Tablo 6.68 diş üst dudak verisinin ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



Şekil 6.51 Ondalık ölçekleme metoduna ait diş üst dudak verilerinin (a) ANFIS eğitime hatası grafiği (b) ANFIS test hatası grafiği

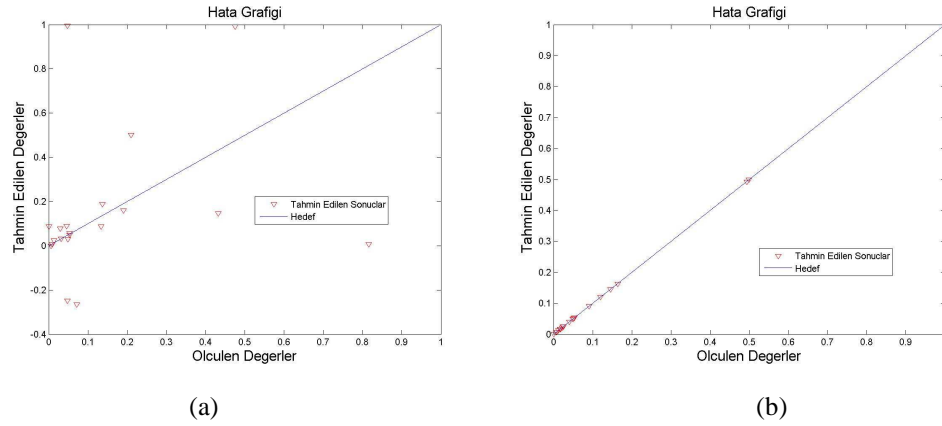
Tablo 6.68 Ondalık ölçekleme metoduna ait diş üst dudak verilerinin ANFIS performans sonuçları

Geliştirilen Metoda İlişkin 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	22.93043	3023.395
MSE	0.000002	0.285376
R^2	0.999275	-82.01201

6.3.5.2.5 Z-skor metoduna ait diş üst dudak verilerinin ANFIS performans sonuçları

Diş üst dudak verisinin, z-skor normalizasyon metodunun ANFIS ile birlikte kullanılmasına ait hata grafiği Şekil 6.52’de gösterilmiştir. Şekil 6.52 (a) diş üst dudak verisinin, eğitime ait hata grafiğini, Şekil 6.52 (b) ise teste ait hata grafiğini göstermektedir

Tablo 6.69 diş üst dudak verisinin ANFIS ile işlenmesine ait eğitim ve teste ait performans kriterlerinin listelenmesine ait sayısal değerleri içermektedir.



Şekil 6.52 Z-skör metoduna ait dış üst dudak verilerinin (a) ANFIS eğitme hatası grafiği (b) ANFIS test hatası grafiği

Tablo 6.69 Z-skör metoduna ait dış üst dudak verilerinin ANFIS performans sonuçları

Z-skör metoduna ait 2-Çapraz Geçerlilik Sonucu Değerleri		
Sonuç	Eğitim	Test
AD	71894244	39096933
MSE	0.000015	0.296045
R²	0.999759	-2.966870

Diş üst dudakverisi için tüm metotların eğitme test değerleri, kıyas yapılabilmesi adına tek bir tabloda verilmiştir. Tablo 6.70, çalışmada önerilen ve literatürdeki yaygın normalizasyon metotlarının dış üst dudak verisini ANFIS ile değerlendirilmesine ait eğitme hatalarının tümünü ait sonuçları göstermektedir. Tablo 6.71, dış üst dudak verisinin testine ait performans değerlerini göstermektedir. Veri setinde elde edilen başarılı sonuçlar

Tablo 6.70 Tüm normalizasyon metotlarına ait dış üst dudak verilerinin ANFIS eğitme performans sonuçları

	EBNM	LBNM	Ham	Z-skör	Ondahk	Min.-Mak.
AD	519.099	12.365	19.366631	71894244	22.93043	∞
MSE	0.026	0.001	0.000001	0.000015	0.000002	0.000015
R²	0.340	0.999	0.999531	0.999759	0.999275	0.999751

Tablo 6.71 Tüm normalizasyon metotlarına ait dış üst dudak verilerinin ANFIS test performans sonuçları

	EBNM	LBNM	Ham	Z-skor	Ondalık	Min.-Mak.
AD	683.158	11.655	480.945240	39096933	3023.395	∞
MSE	0.009	0.004	0.006240	0.296045	0.285376	0.246590
R ²	0.178	0.994	-0.995212	-2.966870	-82.01201	-2.338624

Testteki başarıda LBNM metodu için elde edilen MSE değeri 0.004 ve minimum-maksimum normalizasyon için elde edilen MSE değeri ise 0.2465'tir. Sonuçlardan satır tabanlı normalizasyonun performans değerlendirmeleri diğer yöntemlere göre oldukça iyidir. Yalnız ham veriler için R² değerinin diğer tüm metotlardan daha iyi olduğunu bir gerçektir.

6.4 Sınıflama Verilerinin Değerlendirilmesi

Sınıflama verileri önerilen normalizasyon metodunun performansının denendiği bir diğer veri seti grubudur. Değerlendirmeler toplu halde yapılmış olup performans sonuçları SPECT ve kalp hastalığı için, sınıflama doğruluğu, duyarlılık (sensitivity), seçicilik (specificity) ve AUC (Under Area of Curve) değeri olarak verilmiştir. Tablo 6.72'de SPECT veri setine ait performans sonuçları göstermektedir.

Tablo 6.72 SPECT veri setine ait YSA ve ANFIS performans sonuçları

Normalizasyon Metodu	Performans Ölçüm Kriteri			
	Sınıflama Doğruluğu	Duyarlılık	Seçicilik	AUC Değeri
Minimum-Maksimum	YSA: 77.4	YSA: 87.25	YSA: 45.16	YSA: 0.679
	ANFIS: 77.4	ANFIS: 88	ANFIS: 45.4	ANFIS: 0.692
Desimal Ölçekleme	YSA: 76.69	YSA: 86.4	YSA: 43.34	YSA: 0.66
	ANFIS: 66.16	ANFIS: 86.74	ANFIS: 32	ANFIS: 0.635
Z-Skor Normalizasyon	YSA: 79.69	YSA: 87.61	YSA: 50	YSA: 0.693
	ANFIS: 73.68	ANFIS: 87.36	ANFIS: 39.4	ANFIS: 0.669
Önerilen Metod (LBNM)	YSA: 79.69	YSA: 88.34	YSA: 50	YSA: 0.707
	ANFIS: 64.66	ANFIS: 86.41	ANFIS: 30.7	ANFIS: 0.626

Doppler, VEP ve Hepatit verisi için yalnızca sınıflama doğrulukları verilmiştir.

Birimsizleştirilmiş satır tabanlı normalizasyonun YSA performansının diğer tüm modellerden daha başarılı olduğu görülebilir. Duyarlılık, seçicilik, sınıflama doğruluğu ve AUC değerleri sırasıyla 79.69, 88.34, 50 ve 0.707'dir. AUC değeri kendisine en yakın z-skor normalizasyon değeri 0.693'den %20 daha büyüktür.

Kalp hastalığı veri setinde satır normalizasyonunda YSA performansının %78.51 değerine çıktığı ve AUC değerinin de 0.781 olarak elde edildiği Tablo 6.73'den görülmektedir. Sadece normalizasyon modelinin değişmesi ile %0.73'lük bir başarı artışı oluşmuştur ki bu değer AUC değerine de yansımıştır.

Tablo 6.73 Kalp hastalığı verisine ait YSA ve ANFIS sınıflama performansı sonuçları

Normalizasyon Metodu	Performans Ölçüm Kriteri			
	Sınıflama Doğruluğu	Duyarlılık	Seçicilik	AUC Değeri
Minimum-Maksimum	YSA: 71.11	YSA: 67.11	YSA: 74.32	YSA: 0.708
	ANFIS: 54.81	ANFIS: 49.59	ANFIS: 85.00	ANFIS: 0.588
Desimal Ölçekleme	YSA: 77.78	YSA: 76.78	YSA: 78.48	YSA: 0.771
	ANFIS: 73.34	ANFIS: 67.64	ANFIS: 79.10	ANFIS: 0.736
Z-Skor Normalizasyon	YSA: 72.59	YSA: 67.69	YSA: 77.14	YSA: 0.720
	ANFIS: 71.85	ANFIS: 67.74	ANFIS: 75.34	ANFIS: 0.709
Önerilen Metod (LBNM)	YSA: 78.51	YSA: 76.27	YSA: 80.26	YSA: 0.781
	ANFIS: 72.59	ANFIS: 66.67	ANFIS: 78.79	ANFIS: 0.726

Tablo 6.74'de Doppler veri setine ait sınıflama performansına ait sonuçlar verilmiştir. Doppler veri seti için önerilen metodun hem satır hem de sütun bazında değerlendirmeleri tabloda verilmiştir. Satır bazındaki normalizasyonun %98.25 olan başarı değerini %100'e taşıdığı ve az bir değer de olsa bu başarı artışının ANFIS uygulaması içinde olduğu görülmüştür.

Tablo 6.74 Doppler veri setine ait YSA ve ANFIS performans sonuçları

Normalizasyon Metodu	Sınıflama Doğruluğu	
	YSA	ANFIS
Minimum-Maksimum	98.25	98.24
Önerilen Metod (Sütun bazında)	82.45	96.49
Önerilen Metod (LBNM)	100	98.25

Tablo 6.75 VEP veri setinin normalizasyon değerlerine ait performans sonuçlarını göstermektedir. Klasik model olan minimum-maksimum modelinden sütun bazındaki uygulama YSA’da başarılı iken satır tabanlı uygulama ise ANFIS uygulamasında başarılı sonuçlar elde etmiştir. Özellikle satır tabanlı ANFIS sonucu minimum-maksimum uygulamayı 20.32 puan arttırmıştır.

Tablo 6.75 VEP veri setine ait YSA ve ANFIS performans sonuçları

Normalizasyon Metodu	Sınıflama Doğruluğu	
	YSA	ANFIS
Minimum-Maksimum	75	73.43
Önerilen Metod (Sütun bazında)	78.12	92.18
Önerilen Metod (LBNM)	70.31	93.75

Tablo 6.76 Hepatit veri setinin minimum-maksimum normalizasyonu ve önerilen normalizasyon metodunun satır ve sütundaki uygulanan performans değerlerini içermektedir. Burada da önerilen normalizasyon ya sütun bazında yada satır bazında klasik model olan minimum-maksimum modelden daha başarılı sonuçlar elde etmiştir. Sütun bazındaki YSA sonucu minimum-maksimum modelindeki sonucun 2.29 puan daha üstündedir. Satır tabanlı uygulama ise ANFIS sonucuna göre 7.79 puan daha iyidir.

Tablo 6.76 Hepatit veri setine ait YSA ve ANFIS performans sonuçları

Normalizasyon Metodu	Sınıflama Doğruluğu	
	YSA	ANFIS
Minimum-Maksimum	68.83	68.83
Önerilen Metod (Sütun bazında)	70.12	68.83
Önerilen Metod (LBNM)	68.83	76.62

7. ÖZ SONUÇLAR VE ÖNERİLER

Bu bölüm içinde YSA ve ANFIS yapay zeka uygulamalarına ait Bölüm-6'da verilmiş olan sınıflama ve tahmin veri setlerinin sonuçları kıyaslanmış, her bir tahmin verisi ayrı ayrı değerlendirilerek ileriye yönelik olası çalışmalar üzerinde durulmuştur.

7.1 Özsonuçlar

Bu çalışmada dikkati çeken ham verilerin yapay zeka uygulamalarındaki başarılarının normalizasyon uygulanmış veriye göre düşük çıkması ve bu bağlamda bir ön işlemenin gerekliliğine de işaret etmektedir. Özellikle YSA uygulamalarında transfer fonksiyonunun seçiminin hızlı yakınsama ve daha az hata elde etme adına büyük bir önemi vardır. Bununla birlikte hızlı yakınsama sağlayan transfer fonksiyonlarının veri aralığı maalesef -1,1 ya da 0,1 gibi dar bir aralıktır ve verinin normalizasyonunu gerektirir.

Önerilen normalizasyon formülasyonu satır ve sütunda uygulanabilir şekildedir. Satır bazında uygulama sözkonusu ise birimlerden arınmak adına normalizasyon minimum-maksimum ile öklid uzaklığını beraber kullanan bir yapıya dönüşmektedir. Satır bazındaki tüm uygulamalar Eşitlik 2.10'e uygun olarak yapılmıştır. Sütun bazında yapılan uygulamalar ise Eşitlik 2.9'a uygun olarak yapılmıştır.

Yapılan çalışmalarda satır bazında normalizasyonun çıkış değeri ile giriş parametreleri arasında korelasyonun değiştirdiği Bölüm-3'de anlatılmıştı (Tablo 3.4 ve Tablo 3.5). artan korelasyon ilişkisinin ve *p-olasılık* değerlerinin yapay zeka performansını arttırması umulmuştur ve artan performans sonuçlarını bu iyileşen korelasyon ve *p-olasılık* değerlerine bağlanmıştır. Artan ilişki yapay zekanın daha az hata ile çözüm sunabilmesini sağlamakta ve böylece daha az MSE değeri ve daha yüksek R^2 performansı ortaya çıkmaktadır. Şunu gözden kaçırmamak gerekir ki; tüm

satır normalizasyonu sonunda korelasyon ve *p-olasılık* değerlerinin artmasını beklemek doğru değildir. Nitekim yapılan çalışmalar içerisinde, satır normalizasyonu sonunda korelasyon ve *p-olasılık* değerlerinin olduğundan daha kötü çıktığı uygulamalar olmuştur ki, bunun sonucu yapay zeka sonuçları da kötüleşmiştir.

7.1.1 Aort verilerinin değerlendirilmesi

Aort verilerinin YSA ilgili test çıktıları Bölüm 6.3.1.1’de Tablo 6.8 ve ANFIS test sonuçları Bölüm 6.3.1.2.5’de Tablo 6.15’de verilmiştir. Bu tabloların incelenmesi sonucu önerilen normalizasyon metodunun satır bazında işletilmesine ait tüm performans sonuçları klasik modellerden daha öndedir. Sütun bazında yapılan normalizasyon sonucu ise YSA’ya ait R^2 performansı hariç tüm şartlarda yine iyidir. Ancak önerilen formülasyonun sütun bazında kullanılması veri adedinin çokluğuna bağlı olarak sayıların daha da küçülmesine sebep olacağından MSE değerinin küçük çıkması başarılı olduğu anlamına gelmeyebilir.

Bölüm 3.1’den Tablo 3.4 ve Tablo 3.5’den kalp aort verisine ait normalizasyon öncesi ve sonrası oluşan ortalama korelasyon ve *p-olasılık* değerlerine baktığımızda bu değerlerin satır normalizasyonu sonunda daha da iyileştiği görülmektedir. Satır normalizasyonu (LBNM) ile YSA ve ANFIS sonuçlarının artması bu nedene bağlanabilir. Ortalama korelasyon değeri (liner ilişki katsayısı) 0.546733’den 0.590467 değerine artmış ve *p-olasılık* katsayısı ise 0.001367’den 0 değerine inmiştir. P-olasılık kriteri 0.05’den küçük olması durumunda dikkate alınır. Her iki durumda da 0.05 değerinden küçük olmasına rağmen satır normalizasyonu sonunda daha da küçülmüştür.

7.1.2 Beton verilerinin değerlendirilmesi

Beton verilerinin YSA ilgili test çıktıları Bölüm 6.3.2.1’de Tablo 6.22 ve ANFIS test sonuçları Bölüm 6.3.2.2.5’de Tablo 6.29’da verilmiştir. Bu tabloların incelenmesi sonucu önerilen normalizasyon metodunun satır bazında işletilmesine ait tüm performans sonuçları klasik modellerden daha öndedir. Sütun bazında yapılan normalizasyon sonucu ise ANFIS değerleri satır normalizasyonundan hemen sonra gelmesine karşın YSA sonuçları R^2 kriterinde göre z-skor ile minimum-maksimum arasında bir değere yerleşmektedir. Z-skor değeri 0.9432 ve minimum-maksimum metodu ise 0.9502 olarak elde edilmiştir. Sütun tabanlı çalışılan model de ise R^2 değeri 0.9442 olarak elde edilmiştir.

Beton verileri için Bölüm 3.1’den Tablo 3.4 ve Tablo 3.5’den normalizasyon öncesi ve sonrası oluşan ortalama korelasyon ve *p-olasılık* değerlerine baktığımızda ortalama korelasyon değerinin artarken *p-olasılık* değerinin 0.0000778 değerinden 0.000111 değerine çıkarak 1.42 kat kötüleştiği görülmektedir. Yine de bu değer 0.05 değerinden küçük oluşu istatistiksel bağı bozmadığı yorumu yapılabilir. Ortalama korelasyon değeri ise 0.228344 değerinden 0.303167 değerine çıkarak 1.32 kat iyileşmiştir.

p değerindeki sınırı aşmayan kötüleşme göz ardı edildiğinde satır normalizasyonunda iyileşen değerlerin sebebi olarak korelasyondaki iyileşme gösterilebilir.

7.1.3 Sistol verilerinin değerlendirilmesi

Bir diğer kalp verisi olan sistol verilerine ait YSA ilgili test çıktıları Bölüm 6.3.3.1’de Tablo 6.36’da ve ANFIS test sonuçları Bölüm 6.3.3.2.5’de Tablo 6.43’de verilmiştir. Sistol verilerinin de aort benzeri karakteristiğe sahip olması aort sonuçları ile benzer performans sonuçlarının elde edilmesine yol açmıştır. Satır tabanlı normalizasyonunun ANFIS ve YSA sonuçları tüm kriterler için başarılı olmasına

rağmen sütun bazında yapılan normalizasyonun YSA performansı klasik modellerin gerisindedir.

Bölüm 3, Tablo 3.4 ve Tablo 3.5'den normalizasyon öncesi ve sonrası ortalama korelasyon değerlerine bakıldığında korelasyon değerinin 0.427867 den 1.082 kat artarak 0.462967 değerine çıktığı görülmektedir. Aynı şekilde *p-olasılık* değerinde de bir artış söz konusudur.

Sistol verisindeki YSA ve ANFIS performanslarının artmasını ortalama korelasyon değerinin artmasının bir sonucu olduğu yorumu yapılabilir.

7.1.4 Dış anteior temporal veri setinin değerlendirilmesi

Bu veri seti iki özellikli bir veri olup tek giriş ve tek çıkıştan oluşmaktadır. Anterior verisine ait YSA ilgili test çıktıları Bölüm 6.3.4.1'de Tablo 6.50 ve ANFIS test sonuçları Bölüm 6.3.4.2.5'de Tablo 6.57'de verilmiştir. Önerilen normalizasyon modelinin satır ve sütunda kullanımı olan ve LBNM ve EBNM ile gösterilen her iki model de ANFIS performansları tüm kriterlerde ilk iki sıradadır. YSA sonuçları ise sütun bazında kullanılan model EBNM geride kalsa da satır tabanlı model uygulaması tüm kriterlerde öndedir.

Bunun sebebi olarak Bölüm 3.1'den Tablo 3.4 ve Tablo 3.5'deki ortalama korelasyon ve *p-olasılık* değeri gösterilebilir. Bu iki tabloda özellikle ortalama korelasyondaki artış dikkate değerdir. Korelasyon değeri incelendiği 0.478'den 1.99 kat artarak 0.9538 değerine çıkmıştır. Doğrusal İlişkinin bu kadar arttığı bir model için YSA ve ANFIS performansının artması doğaldır. Korelasyon değerine ilaveten *p-olasılık* değerindeki artışta iyileşme yönündedir ve 0.0018 değerinden 0 değerine inmiştir.

7.1.5 Diş üst dudak veri setinin değerlendirilmesi

Bu diş verisi de anterior temporal veri seti gibi tek giriş ve tek çıkıştan oluşan bir veri seti olup YSA ilgili test çıktıları Bölüm 6.3.5.1’de Tablo 6.64 ve ANFIS test sonuçları Bölüm 6.3.5.2.5’de Tablo 6.71’de verilmiştir. Önerilen normalizasyon modelinin satır ve sütunda kullanımı olan ve LBNM ve EBNM ile gösterilen her iki model de ANFIS performansları tüm kriterlerde ilk iki sıradadır. YSA sonuçları ise sütun bazında kullanılan model EBNM başarısız olsa da satır tabanlı model uygulaması YSA’da da tüm kriterlerde başarılı sonuçlar elde etmiştir.

Başarılı performans değerlerinin sebebi olarak anterior temporal verisindeki gibi korelasyon ve *p-olasılık* değerindeki iyileşme gösterilebilir. Korelasyon değeri ve *p-olasılık* değeri Bölüm 3.1’den Tablo 3.4 ve Tablo 3.5’den sırasıyla 0.9521 ve 0 değerlerine çıkmıştır. Korelasyon 4.12 kat artarken özellikle *p-olasılık* değerindeki iyileşme gözden kaçmamalıdır. 0.1522 değeri ile *p* olasılık değeri 0.05’den küçük olmadığından istatistiksel bir bağıntının varlığına işaret etmezken bu değer 0 değerine düşerek bağıntılı bir hale gelmiştir. Bu iki neden başarılı sonucun bir sebebi olarak kabul edilebilir.

7.2 Sınıflama verilerinin değerlendirilmesi

Sınıflama verilerinde de başarının arttığı uygulamalar dikkate değerdir. Özellikle hepatit, doppler ve VEP verilerindeki başarı dikkate değerdir. Önerilen metodun sütun ya da satırdaki kullanımından herhangi biri minimum-maksimum normalizasyonundan daha başarılıdır. SPECT veri setinde YSA’nın başarısı tüm duyarlılık ve seçicilik ve AUC boyunca belirgindir. En az başarılı olduğu kalp hastalığı verisinde de, sınıflama doğruluğu ve AUC kriterlerine göre başarılıdır. Önerilen normalizasyon metodu LBNM veya EBNM sınıflama verilerinde de başarılı sayılabilecek sonuçlar elde etmiştir. Özellikle SPECT ve kalp hastalığı verisine ait ayrıntı Bölüm 6.4, Tablo 6.72 ve 6.73’de verilmiştir.

7.3 Öneriler

Yapılan bu tez çalışmasında literatürde kullanılan yaygın normalizasyon metotlarına alternatif olabilecek güçlü bir değerlendirme performansına sahip yeni bir metot önerilmiştir. Yapılan çalışma veri özelliğinden çıkarılan bir uzaklık kiteri olan Öklid bağıntısının normalizasyonda kullanılmasını önermekle kalmayıp aynı zamanda uygulama metodunu da büyük çapta değiştirmektedir. Öngörülen yeni metod özellikle tahmin amaçlı veriler üzerine kurgulanmış bir ön işleme metodudur. Öngörülen metod gerek satır gerekse sütun bazında kullanılabilir yapıdadır.

Satır bazındaki uygulama sonunda veri setinin çıkışı ile girişi arasındaki korelasyon ilişkisi ve *p-olasılık* değeri değişmektedir. Bu korelasyon ve *p-olasılık* değerlerinin değişmesi verinin pozitif bir anlam ifade ediyorsa bunun sonucunda YSA ve ANFIS performanslarının artması beklenebilir. Yapılan çalışmalarda bu korelasyon değerlerinin artması YSA ve ANFIS performans sonuçlarını da arttırmıştır.

Satır tabanlı önerilen normalizasyon metoduna LBNM, sütun bazında yapılan çalışmaya ise EBNM denilmiştir. Her iki yöntem de temelde aynı formülasyonu kullanan yalnızca uygulama yönüne ait olarak değişiklik göstermektedir. Yapılan çalışmada önerilen metotlar yalnızca normalizasyon sonuçlarına göre değerlendirilmiş olsalar da bu uygulamayı farklı alanlara taşımak mümkün olabilir. Örneğin veri setinde özellik seçme ayrı bir uygulama sahası olup literatürde oldukça önemli bir konudur. LBNM ve EBNM metotlarının özellik seçiminde kullanılıp kullanılmayacağı yönünde çalışmalar yapılabilir ve özellik seçmeye uygulanabilirliği araştırılabilir.

Bununla birlikte önerilen normalizasyon modeline ait formül içerisine özelliklere has olarak farklı katsayılar eklenerek normalizasyon aşamasında verilerin daha ayrışabilir ya da korelasyonu artacak şekilde normalizasyonu sağlayacak çalışmalar yapılabilir.

Normalizasyonun çok deęişkenli verilerde kullanımı sırasında sayıların aşırı küçölmesine mani olmak için normalizasyon işleminde küçölmeyi engelleyici yapısal tedbirler alacak çalışmalar yapılabilir.

8. KAYNAKLAR:

- Akdemir B., Güneş S., Yosunkaya Ş., New Data Pre-Processing on Assessing of Obstructive Sleep Apnea Syndrome: Line Based Normalization Method (LBNM), International Conference on Intelligent Computing, CCIS (15), ICIC-2008, China, pp: 185-191, 2008a.
- Akdemir B., Güneş S., Oran B., Prediction of Aortic Diameter Values in Healthy Turkish Infants, Children and Adolescents via Adaptive Network Based Fuzzy Inference System, International Conference on Intelligent Computing, LNCS (5226), ICIC-2008 China, pp: 498-505, 2008b
- Akdemir B., Oran B., Güneş S., Karaaslan S., Prediction of Aortic Diameter Values in Healthy Turkish Infants, Children, and Adolescents by Using Artificial Neural Network, Journal of Medical Systems, Accepted: 31 July 2008c (Basımda).
- Beasley D., Bull D.R., ve Martin R.R., An Overview of Genetic Algorithms: Part 1, Fundamentals. University Computing, Vol.15(2), pp. 58-69, U.K. 1993.
- Beiu V., Taylor J. G., On the Circuit Complexity of Sigmoid Feedforward Neural Networks Neural Networks, Vol.9(7), pp:1155-1171, October 1996.
- Buckley J. J., Hayashi Y., Fuzzy Neural Networks: A Survey, Fuzzy Sets and Systems, Vol. 66, pp:1-13, 1994.
- Cios, K.J. Goodenday, L.S. Shah, K.K. Serpen, G., "A novel algorithm for classification of SPECT images of a human heart", Proceedings Ninth IEEE Symposium on Computer-Based Medical Systems, 1– 5, 1996.
- Dirgenali, F., Kara, S., 2006, "Recognition of early phase of atherosclerosis using principles component analysis and artificial neural networks from carotid artery Doppler signals", Expert Systems with Applications, 31(3): 643-651.
- Dorigo M., Maniezzo V., ve Colorni A., Ant System: Optimization by a Colony of Cooperating Agents", IEEE Transactions on Systems, Man, and Cybernetics–Part B, 26 (1): pp:29–41. 1996.
- Evans, D.H., McDicken W.N., Skidmore R., Woodcock J.P., "Doppler Ultrasound", 1989
- Fawcett, F., 2005, "An Introduction to ROC analysis", Pattern Recognition Letters, 27(8): 861-874.

- Gloria A. De., Ferrari P., Grosso D., Olivieri M., Puglisi L., Implementation Techniques for Fuzzy Theory Systems and Their Applications Fuzzy Theory Systems, pp: 1439-1490, 1999.
- Gongde G., Neagu D., Similarity-Based Classifier Combination for Decision Making, Systems, Man and Cybernetics, IEEE International Conference on Vol.1(10), pp: 176-181, 12 Oct. 2005.
- Gonzlez M., Perna A. V. E., Elas F. A., Forcada A. G., A Neural Network-Based Approach For Optimising Rubber Extrusion Lines, International Journal of Computer Integrated Manufacturing, Vol.20(8), pp:828-837, January 2007.
- Govindaraj V., Ramasamy J. V., Optimum Detailed Design of Reinforced Concrete Frames Using Genetic Algorithms Engineering Optimization, Vol. 39(4), pp: 471-494, 2007.
- Gugerty L., Newell and Simon's Logic Theorist: Historical Background and Impact on Cognitive Modeling, Source: Human Factors and Ergonomics Society Annual Meeting Proceedings, General Sessions, Publisher: Human Factors and Ergonomics Society, pp. 880-884(5), 2007.
- Güven, A., 2005, "Göze Ait Elektrofizyolojik Sinyaller Kullanılarak Yapay Sinir Ağları Destekli Bazı Göz Hastalıklarının Teşhisi", Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi.
- Güven, A., Kara, S., 2006, "Diagnosis of the Macular diseases from pattern electroretinography signals using artificial neural networks", Expert Systems with Applications, 30(2): 361-366.
- Güven, A., Polat, K., Kara, S., Güneş, S., 2008, "The effect of generalized discriminate analysis (GDA) to the classification of optic nerve disease from VEP signals", Computers in Biology and Medicine, 38(1):62-68.
- Hebb D. O., The Organization of Behavior, Wiley: New York; 1949 Author: Morris R.G.M.1 Source: Brain Research Bulletin, Publisher: Elsevier Vol.50 (5), pp: 417-437, November 1999.
- Jancic, S., J., Kostic, V., Carelli, V., 2007, "Visual evoked potentials (VEP) and brainstem evoked potentials (BAEP) as diagnostic procedures in Leber's hereditary optic neuropathy", Clinical Neurophysiology, 118(4):119.
- Jang J. S. R., ANFIS: Adaptive-Network-Based Fuzzy Inference Systems, IEEE Trans. Systems, Man & Cybernetics Vol. 23 pp:665-685, 1993.

- Jia L.M., Zhang X.D., Distributed Intelligent Railway Traffic Control: A Fuzzy-Decisionmaking-Based Approach Engineering Applications of Artificial Intelligence, Vol.7(3), pp: 311-319, June 1994.
- Jiang-Ming J. W., Z-Score, Encyclopedia of Biopharmaceutical Statistics: Second Edition, pp:1-4, Published on 23 April 2007.
- Kaveh N. S., Ashrafizadeh S. N., Mohammadi F., Development of An Artificial Neural Network Model For Prediction of Cell Voltage And Current Efficiency in a Chlor-Alkali Membrane Cell, Chemical Engineering Research and Design, Vol. (86-5A), pp: 461-472, 2008.
- Lawlor D. A., Davey G. S., Clark H., Leon D. A., Associations of Birthweight, Gestational Age and Childhood BMI With Type 2 Diabetes: Findings From the Aberdeen Children of the 1950s Cohort, Diabetologia, Vol.49(11), November, 2006.
- Lee E.T., Zadeh L.A., Note on Fuzzy Languages Information Sciences, Vol.1(4), pp:421-424, October 1969.
- Levenberg K., A Method For The Solution of Certain Non-Linear Problems in Least Squares, The Quarterly of Applied Mathematics Vol. 2, pp:164-168, 1944.
- Lin, C.-T., Lee C. S. G., Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems. Upper Saddle River, NJ: Prentice Hall. 1996.
- Loh E., Heat Transfer of Fine-Wire Fuse, Components Hybrids and Manufacturing Technology, IEEE Transactions, Vol. 7(3), pp: 264-267, Sep 1984.
- Mamdani E.H., Assilian S., An Experiment in Linguistic Synthesis With a Fuzzy Logic Controller, International Journal of Man-Machine Studies, Vol. 7, No. 1, pp. 1-13, 1975.
- McCarthy J., Hayes P. J., Some Philosophical Problems From the Standpoint of Artificial Intelligence, *In Meltzer, B., and Michie, D., eds., Machine Intelligence Vol 4*, Edinburgh: Edinburgh University Press, pp:463-502, 1969.
- McCulloch W., Pitts W., A Logical Calculus of The Ideas Immanent in Nervous Activity, Bulletin of Mathematical Biophysics, Vol. 7, pp:115-133, 1943.
- McCulloch W., Pitts W., On How We Know Universals: The Perception of Auditory and Visual Forms, Bulletin of Mathematical Biophysics, Vol.9, pp:127-147, 1947.
- Mamdani E. H., Application of Fuzzy Algorithms for Control of Simple Dynamic Plant, proc. IEEE, Vol.121, pp: 1585 – 1588, 1974.

- Minai A. A., Williams R. D., On the Derivatives of the Sigmoid Neural Networks, Vol.6(6), pp: 845-853 1993.
- Minsky M., Papert S., Perceptrons, MIT Press, Cambridge, MA, 1969.
- Munson, B. R., Young, D. F., Okiishi, T. H., Fundamentals of Fluid Mechanics, 4. Edition, Wiley, USA, 2002.
- Nabih N. A., A Simplex Algorithm For Minimum Fuel Problems of Linear Discrete Control Systems, International Journal of Control, Vol.26(4), pp: 635-642, 1977.
- Nabiyev V. V., Yapay Zeka, Seçkin Yayınları, 2. baskı, 2005.
- Nelson D. A., Darga L. L., Simon M. S., Severson R. K., Radial Bone Density and Breast Cancer Risk in White and African-American Women, Osteoporosis International, Vol.15(7), July, 2004.
- Onbaşılı, O. A., Tekten, T., Ceyhan, C., Ercan, E., Dilate Kardiyomiyopati Hastalarında Ultrasonik Integrated Backscatter Analizi: Sağlıklı Bireylerle Karşılaştırma, Anadolu Kardiyoloji Dergisi, 2, 4-11, 2002.
- Ökkesim Ş., Kara S., Uysal T., Yağcı, A., Pre-Ortodontik Trainer Aparesi Kullanılan Hastalarda Çene Kaslarının Elektromyogram ve Ayrık Dalgacık Dönüşümü ile Analizi, Biyomedikal Mühendisliği Ulusal Toplantısı (BİYOMUT'06), İstanbul, May. 2006,
- Özşen S., Kara S., Latifoğlu F., Güneş S., A New Supervised Classification Algorithm in Artificial Immune Systems With its Application to Carotid Artery Doppler Signals to Diagnose Atherosclerosis, Computer Methods and Programs in Biomedicine, Vol.88(3), pp: 246-255, December 2007.
- Pan Y., Jiang J., Wang R., Cao H., Zhao J., Prediction of Auto-Ignition Temperatures of Hydrocarbons by Neural Network Based on Atom-Type Electrotological-State Indices Journal of Hazardous Materials, Vol.157(2-3), pp: 510-517, 15 September 2008.
- Parichatprecha R., Nimityongskul P., Analysis of Durability of High Performance Concrete Using Artificial Neural Networks, Construction and Building Materials, In Press, Corrected Proof, Available online 24 June 2008.
- Razmi-Rad E., Ghanbarzadeh B., Mousavi S.M., Emam-Djomeh Z., Khazaei J., Prediction of Rheological Properties of Iranian Bread Dough From Chemical Composition of Wheat Flour by Using Artificial Neural Networks, J Food Eng, Vol.81(4), pp:728-734, 2007.

- Roger Jang J.-S., ANFIS: Adaptive-Network-Based Fuzzy Inference Systems, IEEE Transactions on Systems, Man, and Cybernetics, Vol.23(03), pp 665-685, May 1993.
- Rumelhart D. E., McClelland J. L., Parallel Distributed Processing, MIT Press, Cambridge, MA, 1986a.
- Rumelhart D. E., Hinton G. E., Williams R. J., Learning Internal Representations by Error Propagation, in [Rumelhart and McClelland 86], pp:318-364, 1986b.
- Saltık S., Işık, K., Saltık, İ. L., Çiftçi, H., Talasemi Majörlü Hastalarda Kardiyak Fonksiyonların Ekokardiyografik Olarak Değerlendirilmesi, İstanbul Çocuk Kliniği Dergisi, 32 (1), 23-27, 1997
- Sağıroğlu Ş., Beşdok E., Erler M., Mühendislikte Yapay Zeka Uygulamaları-1, Ufuk Yayıncılık, 2003.
- Shalabi L. A., "Coding and Normalization: The Effect of Accuracy, Simplicity, and Training Time", RCED'05, Al-Hussain Bin Talal University, 2006a.
- Shalabi L. A., Shaaban Z., Kasasbeh B., Data mining: a preprocessing engine, Journal of Computer Science, September, 2006b.
- Shalabi L. A., Shaaban Z., Normalization as a Preprocessing Engine for Data Mining and the Approach of Preference Matrix, Depcos-Relcomex, pp. 207-214, International Conference on Dependability of Computer Systems (DEPCOS-RELCOMEX'06), 2006c.
- Sugeno M., Industrial Applications of Fuzzy Control, Elsevier Science Inc., Newyork, NY, 1985.
- Sungmook L., LPAKO: A Simplex-Based Linear Programming Program, Soondal Park Optimization Methods and Software, Vol. 17(4), pp:717-745, 2002.
- Tasman, W., 1992, Duane's Foundations of Clinical Ophthalmology, Lippincott Williams and Wilkins, United States of America.
- The UCI KDD Archive Information and Computer Science University of California, Irvine <http://kdd.ics.uci.edu/databases/>
- Van-Tsai L., Chun-Liang L., Gean-Pao L., Neural-Network-Based Identification and Control of A Thin Plate Using Piezoelectric Actuators and Sensors International Journal of Systems Science, Vol.35(6), pp:355-373, 2004.

- Vong C. M., Wong P. K., Li Y. P., Ho C. M., Modelling of Modern Automotive Petrol Engine Performance Using Support Vector Machines, Journal of Zhejiang University SCIENCE, J Zhejiang Univ, Vol.6A(1), pp:1-8, 2005.
- Wang L.X., A Course in Fuzzy Systems and Control, Prentice Hall, Paperback, Published August 1996.
- Widrow, B., Lehr M. A., 30 years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation, Proc. IEEE, Vol.78(9), pp:1415-1442, 1990.
- William, P., Honan, James R. Heron, David H. Foster, Graham K. Edgar and Mark O. Scase, "Abnormalities of visual function in hereditary motor and sensory neuropathy", Journal of the Neurological Sciences, 114(2):188-192, 1993
- Yan, H., Jiang Y., Zheng J., Peng C., Li Q., "A multilayer perceptron-based medical decision support system for heart disease diagnosis", Expert Systems with Applications, 30(2), 272-281, 2006.
- Yasuhara, A., Recursive Function Theory and Logic, Academic Press, New York, 1971.