T.C.
ERCİYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

BENZETİLMİŞ TAVLAMA ALGORİTMASI İLE ADAPTİF AĞ TABANLI BULANIK MANTIK ÇIKARIM SİSTEMİNİN (ANFIS) EĞİTİLMESİ

Hazırlayan
Bülent HAZNEDAR

Danışman
Prof. Dr. Adem KALINLI

Doktora Tezi

Mayıs 2017
KAYSERİ

T.C.
ERCİYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

BENZETİLMİŞ TAVLAMA ALGORİTMASI İLE ADAPTİF AĞ TABANLI BULANIK MANTIK ÇIKARIM SİSTEMİNİN (ANFIS) EĞİTİLMESİ

(Doktora Tezi)

Hazırlayan
Bülent HAZNEDAR

Danışman
Prof. Dr. Adem KALINLI

Bu çalışma; Erciyes Üniversitesi Bilimsel Araştırma Projeleri Birimi tarafından FDK-2016-6371 kodlu proje ile desteklenmiştir.

Mayıs 2017
KAYSERİ

BİLİMSEL ETİĞE UYGUNLUK

Bu çalışmadaki tüm bilgilerin, akademik ve etik kurallara uygun bir şekilde elde edildiğini beyan ederim. Aynı zamanda bu kural ve davranışların gerektirdiği gibi, bu çalışmanın özünde olmayan tüm materyal ve sonuçları tam olarak aktardığımı ve referans gösterdiğimi belirtirim.

Bülent HAZNEDAR

İmza :

YÖNERGEYE UYGUNLUK

Benzetilmiş Tavlama Algoritması ile Adaptif Ağ Tabanlı Bulanık Mantık Çıkarım Sisteminin (ANFIS) Eğitilmesi adlı Doktora tezi, Erciyes Üniversitesi Lisansüstü Tez Önerisi ve Tez Yazma Yönergesi'ne uygun olarak hazırlanmıştır.

Tezi Hazırlayan	Danışman
Bülent Haznedar	Prof. Dr. Adem Kalınlı

Bilgisayar Mühendisliği ABD Başkanı

Prof. Dr. Derviş Karaboğa

Prof. Dr. Adem KALINLI danışmanlığında **Bülent HAZNEDAR** tarafından hazırlanan "**Benzetilmiş Tavlama Algoritması ile Adaptif Ağ Tabanlı Bulanık Mantık Çıkarım Sisteminin (ANFIS) Eğitilmesi**" adlı bu çalışma, jürimiz tarafından Erciyes Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalında **Doktora** tezi olarak kabul edilmiştir.

…….. /…… / ……..

JÜRİ:

Danışman	:…….	…….
Üye	:…….	…….
Üye	:…….	…….
Üye	:…….	…….
Üye	:…….	…….

ONAY:
Bu tezin kabulü Enstitü Yönetim Kurulunun ……...… tarih ve ……..…… sayılı kararı ile onaylanmıştır.

……. /……../ ……

Prof. Dr. Mehmet AKKURT

Enstitü Müdürü

ÖNSÖZ / TEŞEKKÜR

Çalışmalarım boyunca beni yönlendiren, hiç bir konuda bilgi, tecrübe, yardım ve yakın ilgisini esirgemeyen, üzerimde çok fazla emeği olan kıymetli danışmanım Prof. Dr. Adem KALINLI hocama saygılarımı sunar, teşekkür ederim. Tez izleme komitemde bulunarak çalışmalarımı yönlendiren hocalarım Prof. Dr. Derviş Karaboğa'ya ve Prof. Dr. Kenan Danışman'a da ayrıca teşekkür ederim.

Doktora tez çalışmamı Yurt İçi Lisansüstü Burs Programı kapsamında yürütülen 2211-A Genel Yurt İçi Doktora Bursu ile destekleyen TÜBİTAK Bilim İnsanı Destekleme Daire Başkanlığı birimine teşekkür ederim. Ayrıca, çalışmama maddi destek veren Erciyes Üniversitesi Bilimsel Araştırma Projeleri Birimi'ne (Proje No: FDK-2016-6371) teşekkürü bir borç bilirim.

Ayrıca; çalışmalarım süresince destekleri ile her zaman yanımda olan, sevgi, saygı ve minnet sözcüklerinin yetersiz kaldığı sevgili aileme ve tüm dostlarıma teşekkürlerimi sunuyorum.

Bülent HAZNEDAR

Kayseri, Mayıs 2017

BENZETİLMİŞ TAVLAMA ALGORİTMASI İLE ADAPTİF AĞ TABANLI BULANIK MANTIK ÇIKARIM SİSTEMİNİN (ANFIS) EĞİTİLMESİ

Bülent HAZNEDAR

Erciyes Üniversitesi, Fen Bilimleri Enstitüsü

Doktora Tezi, Mayıs 2017

Danışman: Prof. Dr. Adem KALINLI

ÖZET

Zor problemlerin çözümünde yaygın olarak kullanılan Adaptif Ağ Tabanlı Bulanık Mantık Çıkarım Sistemi (ANFIS), örneklerden öğrenebilme ve öğrenmiş olduğu bilgiyi belirli kurallara bağlı olarak ağ üzerinde dağıtabilme yeteneği olan esnek ve güçlü bir yapay zekâ tekniğidir. Başlangıç ve sonuç parametrelerinden oluşan ANFIS parametrelerinin eğitimi, doğrusal olmayan oldukça zor bir optimizasyon problemidir. Bu kapsamda, ANFIS'in eğitimi için geleneksel algoritmalar yaygın olarak kullanılmaktadır. Ancak, geleneksel algoritmalar ile ANFIS parametrelerinin eğitimi çok uzun süreler gerektirebilmekte ve özellikle zor problemlerdeki performansı çoğunlukla düşük olmaktadır. Bu tez çalışmasında, ANFIS modelinin başarılı bir şekilde eğitilmesi için yapay zekâ optimizasyon algoritmalarından iteratif tabanlı benzetilmiş tavlama (SA) algoritmasının kullanılması önerilmiştir. Önerilen yaklaşımın performansı farklı problemler üzerinde incelenmiş ve elde edilen sonuçlar farklı yöntemlere ait sonuçlarla karşılaştırılmıştır.

Tez kapsamında, dinamik sistemlerin kimliklendirilmesi amacıyla ANFIS modeli SA algoritması kullanılarak eğitilmiş ve elde edilen sonuçlar farklı yöntemlere ait sonuçlar ile karşılaştırılmıştır. Ayrıca, mikrodizi gen ifade verilerini sınıflandırmak için de çalışmalar yapılmıştır. Sınıflandırmada kullanılan ANFIS modelinin eğitilmesi için SA algoritması kullanılarak, önerilen yöntemin performansı farklı algoritmaların performanslarıyla karşılaştırılmıştır. Elde edilen sonuçlardan SA algoritmasının, ANFIS parametrelerini eğitmede başarılı bir şekilde kullanılabileceği kanaatine varılmıştır.

SA algoritmasının küresel optimuma erişebilmesi için gereken süre başlangıç çözümüne bağlı olarak değişim göstermektedir. Bu durum temel SA algoritmasının dinçlik bakımından performansını çoğunlukla düşürmektedir. Popülasyon tabanlı algoritmaların

ise çözümlerin bir seti ile çalışması, araştırmanın küresel optimumun bulunduğu bölgeye hızla erişmesini sağlamaktadır. Tez çalışması kapsamında, SA algoritmasının yukarıda bahsedilen dezavantajını ortadan kaldırıp performansını artırmak için, popülasyon tabanlı algoritmaların paralel yapı avantajını kullanarak yeni bir paralel benzetilmiş tavlama algoritması (PSA) modeli önerilmiştir. Önerilen modelin performansı doğrusal ve doğrusal olmayan dinamik sistemlerin kimliklendirilmesi amacıyla ANFIS parametrelerinin eğitilmesi için kullanılmıştır. Yapılan analizler sonucunda temel SA algoritmasının performansında önemli düzeyde iyileşme sağlandığı görülmüştür.

Anahtar Kelimeler: ANFIS; benzetilmiş tavlama algoritması; ANFIS ağının eğitilmesi; sistem kimliklendirme; sınıflandırma; mikrodizi gen ifade.

TRAINING ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM (ANFIS) USING SIMULATED ANNEALING ALGORITHM

Bülent HAZNEDAR

Erciyes University, Graduate School of Natural and Applied Sciences

Ph.D. Thesis, May 2017

Supervisor: Prof. Dr. Adem KALINLI

ABSTRACT

Adaptive Neuro-Fuzzy Inference System (ANFIS), widely used in the solution of difficult problems, is a flexible and powerful artificial intelligence technique that has the ability to learn from examples and to distribute the knowledge that has been learned on the network depending on certain rules. Training of the ANFIS parameters consisting of the antecedent and conclusion parameters is a quite difficult nonlinear optimization problem. In this context, traditional algorithms for the training of ANFIS are widely used. However, training of ANFIS parameters with traditional algorithms may require very long times and the performance on especially difficult problems is often low. In this thesis study, it is proposed to successfully train the ANFIS model using iterative-based simulated annealing (SA) algorithm which is one of the artificial intelligence algorithm. The performance of the proposed approach is examined on different problems and the obtained results are compared with the results of different methods.

In the scope of the thesis, the ANFIS model is trained using the SA algorithm for the dynamic systems identification and the obtained results are compared with the results of different methods. Also, studies are conducted to classify microarray gene expression data. The SA algorithm is used to train the ANFIS model for classification and the performance of the proposed method is compared with the performance of different algorithms. From the obtained results, it is concluded that SA algorithm can be successfully used for training ANFIS parameters.

The time required for the SA algorithm to reach the global optimum may vary depending on the initial solution. This reduces the performance of the basic SA algorithm in terms of robustness. On the other hand, since population-based algorithms work with a set of candidate solutions, it is also known that they are able to access the global optimum region

quickly. In the scope of the thesis, a new parallel simulated annealing algorithm (PSA) model is proposed by using the parallel structure advantage of population-based algorithms eliminating the disadvantage mentioned above for the SA algorithm to improve its performance. The performance of the proposed model is used to train ANFIS parameters to identify linear and nonlinear dynamical systems. As a result of the analyzes made, it has been seen that the performance of the basic SA algorithm has improved significantly.

Keywords: ANFIS; simulated annealing algorithm; training ANFIS network; system identification; classification; microarray gene expression.

İÇİNDEKİLER

BENZETİLMİŞ TAVLAMA ALGORİTMASI İLE ADAPTİF AĞ TABANLI BULANIK MANTIK ÇIKARIM SİSTEMİNİN (ANFIS) EĞİTİLMESİ

BİLİMSEL ETİĞE UYGUNLUK SAYFASI ii

YÖNERGEYE UYGUNLUK SAYFASI iii

KABUL VE ONAY SAYFASI iv

ÖNSÖZ / TEŞEKKÜR v

İÇİNDEKİLER x

KISALTMALAR VE SİMGELER xv

TABLOLAR LİSTESİ xvi

ŞEKİLLER LİSTESİ xvii

1. BÖLÜM

GİRİŞ

1.1. Genel Bilgiler 1

1.2. Literatür Özeti 3

1.3. Tezin Amacı ve Yöntemi 7

1.4. Tezin Bölümleri 10

2. BÖLÜM

ADAPTİF AĞ TABANLI BULANIK MANTIK ÇIKARIM SİSTEMİ

2.1. Bulanık Mantık ve Bulanık Model 12

2.1.1. Mamdani Bulanık Modeli 16

2.1.2. Takagi-Sugeno Bulanık Modeli 18

2.2. Adaptif Ağ Tabanlı Bulanık Mantık Çıkarım Sistemi (ANFIS)..... 20

2.2.1. ANFIS Modelini Kimliklendirme Yöntemleri 24

2.2.1.1. Izgara Bölümleme (Grid Partitioning)..... 25

2.2.1.2. Eksiltici Kümeleme (Subtractive Clustering)..... 26

2.2.1.3. Bulanık C-Ortalamalar Kümeleme (Fuzzy C-Means Clustering-FCM) 28

2.2.2. ANFIS Modelinin Eğitilmesi..... 30

2.2.2.1. Klasikleşmiş Algoritmalar..... 31

2.2.2.1.1. Geri Yayılım Algoritması..... 31

2.2.2.1.2. Hibrid Öğrenme Algoritması..... 33

2.2.2.2. Sezgisel Algoritmalar 34

2.2.2.2.1. Genetik Algoritma 35

2.2.2.2.2. Parçacık Sürü Optimizasyon Algoritması..... 38

3. BÖLÜM

BENZETİLMİŞ TAVLAMA ALGORİTMASI İLE ANFIS AĞININ EĞİTİLMESİ

3.1. Benzetilmiş Tavlama Algoritması..... 41

3.1.1. Fiziksel Tavlama Süreci 42

3.1.2. SA Algoritmasının Çalışması..... 44

3.1.3. SA Algoritmasına ait Parametreler 46

3.1.3.1. Teorik Soğutma Stratejileri..... 48

3.1.3.2. Basit Soğutma Stratejileri 49

3.2. Trombofili Hastalığı ile Genetik Bozukluklar Arasındaki İlişkinin Tespit Edilmesi 51

3.2.1. Tıp Bilişimi 51

3.2.1.1. Tıp Bilişimi Uygulama Alanları..... 53

3.2.1.2. Tıbbi Karar Destek Sistemleri (Medical Decision Support Systems) 53

3.2.2. Trombofili Hastalığının Genel Yapısı 56

3.2.2.1. Kalıtsal Trombofili 59

3.2.3. Deneysel Çalışmalar ve Analiz 64

3.2.3.1. Makine Öğrenmesi Algoritmaları ile Başarım Analizi 64

3.2.3.2. ANFIS ile Başarım Analizi 69

3.2.4. Genel Değerlendirme ve Sonuçlar 69

3.3. Dinamik Sistemlerin Kimliklendirilmesi 71

3.3.1. Sistem Kimliklendirmenin Tanımı ve Amacı 71

3.3.2. Dinamik Sistemlerin Kimliklendirilmesi için SA Algoritması ile ANFIS Ağının Eğitilmesi 72

3.3.3. Önerilen Yöntemin Genel Yapısı 74

3.3.4. Simülasyon Sonuçları 76

3.3.4.1. Dinamik Sistem 1 79

3.3.4.2. Dinamik Sistem 2 80

3.3.4.3. Dinamik Sistem 3 82

3.3.5. Genel Değerlendirme ve Sonuçlar 83

4. BÖLÜM

BULANIK C-ORTALAMALAR TABANLI ANFIS AĞININ BENZETİLMİŞ TAVLAMA ALGORİTMASI İLE EĞİTİLMESİ

4.1. Mikrodizi Gen İfade Kanser Verilerinin Sınıflandırılması 85

4.1.1. Gen Ekspresyonu (İfade) 85

4.1.2. Mikrodizi Teknolojisi 86

4.1.3. Mikrodizi Gen İfade Kanser Verilerini Sınıflandırmak için SA Algoritması Kullanılarak FCM Tabanlı ANFIS Ağının Eğitilmesi 88

4.1.4. Deneysel Veri Setleri **90**

4.1.4.1. Akciğer Kanseri *(Gordon-2002)* **90**

4.1.4.2. Beyin Kanseri *(Nutt-2003)* **91**

4.1.4.3. Merkezi Sinir Sistemi Kanseri *(Pomeroy-2002)* **91**

4.1.4.4. Rahim Kanseri *(Risinger-2003)* **91**

4.1.4.5. Prostat Kanseri *(Singh-2003)* **91**

4.1.5. Deneysel Çalışmalar ve Analiz **92**

4.1.6. Genel Değerlendirme ve Sonuçlar **106**

4.2. Dinamik Sistemlerin Kimliklendirilmesi için FCM Tabanlı ANFIS Ağının SA Algoritması ile Eğitilmesi **107**

4.2.1. Önerilen Yöntem **108**

4.2.2. Simülasyon Sonuçları **110**

4.2.2.1. Dinamik Sistem 1 ..... **114**

4.2.2.2. Dinamik Sistem 2 ..... **115**

4.2.2.3. Dinamik Sistem 3 ..... **117**

4.2.3. Genel Değerlendirme ve Sonuçlar **118**

5. BÖLÜM

BENZETİLMİŞ TAVLAMA ALGORİTMASI İÇİN YENİ BİR MODEL

5.1. Giriş **120**

5.2. Önerilen Modelin Genel Yapısı **121**

5.3. Sistem Kimliklendirme için Önerilen PSA Modeli Kullanılarak ANFIS Ağının Eğitilmesi **123**

5.3.1. Dinamik Sistem 1 ..... **124**

5.3.2. Dinamik Sistem 2 ..... **126**

5.3.3. Dinamik Sistem 3 ..... **127**

5.4. Genel Değerlendirme ve Sonuçlar **128**

6. BÖLÜM

TARTIŞMA, SONUÇ VE ÖNERİLER

6.1. Tartışma ve Sonuç 130

6.2. Gelecekte Yapılacak Çalışmalar 134

KAYNAKÇA 135

ÖZGEÇMİŞ 152

KISALTMALAR VE SİMGELER

ANFIS	Adaptif Ağ Tabanlı Bulanık Mantık Çıkarım Sistemi
GSP	Gezgin Satıcı Problemi
BP	Geriye Yayılım (Backpropagation)
LSE	En Küçük Kareler Tahmini (Least Squares Estimation)
GD	En Dik İniş (Gradient Descent)
LM	Levenberg Marquardt
KF	Kalman Filtresi (Kalman Filter)
GA	Genetik Algoritma
PSO	Parçacık Sürüsü Optimizasyon Algoritması
ABC	Yapay Arı Kolonisi Algoritması
ACO	Karınca Koloni Optimizasyon Algoritması
DGA	Diferansiyel Gelişim Algoritması
RKLM	Runge-Kutta Öğrenme Algoritması
ICA	Emperyalist Rekabetçi Algoritması
GP	Izgara Bölümleme (Grid Partitioning)
SC	Eksiltici Kümeleme (Subtractive Clustering)
FCM	Bulanık C-Ortalamalar Kümeleme (Fuzzy C-Means Clustering)
SA	Benzetilmiş Tavlama Algoritması (Simulated Annealing Algorithm)
MFs	Üyelik Fonksiyonları
HB	Hibrid Öğrenme Algoritması
TB	Tıp Bilişimi
RMSE	Ortalama Karekök Hata (Root Mean Square Error)
TS	Tabu Arama Algoritması
PTS	Paralel Tabu Arama Algoritması
SVM	Destek Vektör Makineleri (Support Vector Machine)
YSA	Yapay Sinir Ağları
RTYSA	Radyal Tabanlı Yapay Sinir Ağları
J48	J48 Karar Ağacı
PSA	Paralel Benzetilmiş Tavlama Algoritması
BAS	Baskın Aday Seçimi

TABLOLAR LİSTESİ

Tablo 2.1. Hibrid öğrenme algoritmanın iki geçişli öğrenme süreci 33
Tablo 3.1. Tavlama süreci ve optimizasyon problemi arasındaki ilişki 44
Tablo 3.2. Edinsel trombofili nedenleri 57
Tablo 3.3. Kalıtsal trombofili nedenleri 58
Tablo 3.4. Uygulanması önerilen testler 62
Tablo 3.5. Parametreler ve bu parametre değerine sahip hasta sayıları 66
Tablo 3.6. Farklı yöntemlere ait sınıflandırma sonuçları. 68
Tablo 3.7. ANFIS modeli ile başarım sonuçları 69
Tablo 3.8. Doğrusal sistem için sonuçların karşılaştırılması 80
Tablo 3.9. Sistem 2 için sonuçların karşılaştırılması 81
Tablo 3.10. Sistem 3 için sonuçların karşılaştırılması 83
Tablo 4.1. Çalışmada kullanılan veri setleri 92
Tablo 4.2. Veri setleri ve gen sayıları 93
Tablo 4.3. Optimize edilen ANFIS modellerine ait bilgiler 94
Tablo 4.4. Sınıflandırma algoritmalarının doğruluk yüzdeleri (%) 101
Tablo 4.5. Metotların ortalama başarımları ve gelişim yüzdeleri 102
Tablo 4.6. Akciğer kanseri veri setine ait model başarım ölçütleri 103
Tablo 4.7. Beyin kanseri veri setine ait model başarım ölçütleri 103
Tablo 4.8. Merkezi sinir sistemi kanser veri setine ait model başarım ölçütleri 104
Tablo 4.9. Rahim kanseri veri setine ait model başarım ölçütleri 104
Tablo 4.10. Prostat kanseri veri setine ait model başarım ölçütleri 104
Tablo 4.11. Yöntemlere ait genel model başarım ölçütleri 105
Tablo 4.12. Doğrusal sistem için sonuçların karşılaştırılması 115
Tablo 4.13. Sistem 2 için sonuçların karşılaştırılması 116
Tablo 4.14. Sistem 3 için sonuçların karşılaştırılması 117
Tablo 5.1. Doğrusal sistem için sonuçların karşılaştırılması 125
Tablo 5.2. Sistem 2 için sonuçların karşılaştırılması 126
Tablo 5.3. Sistem 3 için sonuçların karşılaştırılması 128

ŞEKİLLER LİSTESİ

Şekil 2.1. Yaygın olarak kullanılan üyelik fonksiyonları a) Üçgen b) Yamuk c) Normal Dağılım d) Çan Şekilli 13
Şekil 2.2. Bulanıklaştırma – durulaştırma birimli bulanık model sistemi 14
Şekil 2.3. Mamdani bulanık mantık modeli 17
Şekil 2.4. Sugeno bulanık mantık modeli 19
Şekil 2.5. Adaptif ağ tabanlı bulanık mantık çıkarım sistemi 21
Şekil 2.6. Gaussian üyelik fonksiyonlarındaki parametrelerin hesaplanmış hali 25
Şekil 2.7. GP yöntemi ile oluşturulmuş dokuz kural bölgesi 26
Şekil 2.8. Kümeleme temelli bölümleme 28
Şekil 2.9. Geriye yayılım algoritmasının temel adımları 32
Şekil 2.10. Genetik algoritmanın temel adımları 36
Şekil 2.11. Parçacık sürü optimizasyon algoritmasının temel akış diyagramı 39
Şekil 2.12. Parçacık sürü optimizasyon algoritmasının temel adımları 40
Şekil 3.1. Isıl işlem kademeleri 42
Şekil 3.2. SA algoritmasının temel adımları 45
Şekil 3.3. SA algoritmasının akış diyagramı 46
Şekil 3.4. Basamak tipte sıcaklık düşürülmesi 49
Şekil 3.5. Sürekli olarak sıcaklığın düşürülmesi 50
Şekil 3.6. Monoton olmayan şekilde sıcaklığın düşürülmesi 50
Şekil 3.7. Türkiye' de tıp bilişimi uygulamalarının gelişimi 52
Şekil 3.8. Günümüz karar destek sistemlerinin ilişkili olduğu bilim dalları 55
Şekil 3.9. Pubmed' de yıllara göre gerçekleştirilen yapay zekâ ile ilgili yayın sayısı 56
Şekil 3.10. Öznitelik Seçimi yöntemiyle bulunan ki-kare değerleri 67
Şekil 3.11. Bir dinamik sistem şeması 71
Şekil 3.12. Sistem kimliklendirme için ANFIS modelinin SA algoritması ile eğitilmesi planı 75
Şekil 3.13. Sistem 1 için 6 girişli ANFIS modeli 77
Şekil 3.14. Sistem 2 için 3 girişli ANFIS modeli 78
Şekil 3.15. Sistem 3 için 4 girişli ANFIS modeli 78
Şekil 3.16. Sistem 1 için, sistemin ve SA algoritması ile eğitilen ANFIS modelinin çıkışları (*RMSE* hata değeri = 1,0378E-09) 80

Şekil 3.17. Sistem 2 için, sistemin ve SA algoritması ile eğitilen ANFIS modelinin çıkışları (*RMSE* hata değeri = 1,3216E-03) 82
Şekil 3.18. Sistem 3 için, sistemin ve SA algoritması ile eğitilen ANFIS modelinin çıkışları (*RMSE* hata değeri = 2,1518E-03) 83
Şekil 4.1. FCM ile modellenmiş temel ANFIS mimarisi 93
Şekil 4.2. Akciğer kanseri verisinin sınıflandırılması ile elde edilen eğitim (*RMSE*) hata değerleri 96
Şekil 4.3. Beyin kanseri verisinin sınıflandırılması ile elde edilen eğitim (*RMSE*) hata değerleri 96
Şekil 4.4. Merkezi Sinir Sistemi kanseri verisinin sınıflandırılması ile elde edilen eğitim (*RMSE*) hata değerleri 97
Şekil 4.5. Rahim kanseri verisinin sınıflandırılması ile elde edilen eğitim (*RMSE*) hata değerleri 97
Şekil 4.6. Prostat kanseri verisinin sınıflandırılmasında elde edilen eğitim (*RMSE*) hata değerleri 98
Şekil 4.7. Akciğer kanseri verisinin sınıflandırılmasında elde edilen test (*RMSE*) hata değerleri 98
Şekil 4.8. Beyin kanseri verisinin sınıflandırılmasında elde edilen test (*RMSE*) hata değerleri 99
Şekil 4.9. Merkezi Sinir Sistemi kanseri verisinin sınıflandırılmasında elde edilen test (*RMSE*) hata değerleri 99
Şekil 4.10. Rahim kanseri verisinin sınıflandırılmasında elde edilen test (*RMSE*) hata değerleri 100
Şekil 4.11. Prostat kanseri verisinin sınıflandırılmasında elde edilen test (*RMSE*) hata değerleri 100
Şekil 4.12. Tüm mikrodizi gen ifadesi profilleri üzerindeki sınıflandırma algoritmalarının ortalama doğruluk yüzdeleri (%) 102
Şekil 4.13. Tüm mikrodizi gen ifadesi profilleri üzerindeki sınıflandırma algoritmalarının ortalama Kesinlik, Duyarlılık ve F-ölçütü değerleri 105
Şekil 4.14. Sistem kimliklendirme için FCM-ANFIS modelinin SA algoritması ile eğitilmesi planı. 109
Şekil 4.15. Sistem 1 için 6 girişli ANFIS modeli 111
Şekil 4.16. Sistem 2 için 3 girişli ANFIS modeli 111

Şekil 4.17. Sistem 3 için 4 girişli ANFIS modeli 112
Şekil 4.18. FCM ile modellenmiş temel ANFIS mimarisi 113
Şekil 4.19. Sistem 1 için, sistemin ve SA algoritması ile eğitilen ANFIS modelinin cevapları (*RMSE* hata değeri = 9,2087E-12) 114
Şekil 4.20. Sistem 2 için, sistemin ve SA algoritması ile eğitilen ANFIS modelinin cevapları (*RMSE* hata değeri = 8,7368E-04) 116
Şekil 4.21. Sistem 3 için, sistemin ve SA algoritması ile eğitilen ANFIS modelinin cevapları (*RMSE* hata değeri = 8,6247E-04) 118
Şekil 5.1. Önerilen paralel benzetilmiş tavlama algoritmasının yapısı 122
Şekil 5.2. Sistem kimliklendirme için ANFIS modelinin PSA algoritması ile eğitilmesi planı. 123
Şekil 5.3. Sistem 1 için, sistemin ve PSA algoritması ile eğitilen ANFIS modelinin cevapları (*RMSE* hata değeri = 3,1971E-04) 125
Şekil 5.4. Sistem 2 için, sistemin ve PSA algoritması ile eğitilen ANFIS modelinin cevapları (*RMSE* hata değeri = 8,7335E-04) 127
Şekil 5.5. Sistem 3 için, sistemin ve PSA algoritması ile eğitilen ANFIS modelinin cevapları (*RMSE* hata değeri = 1,7015E-03) 128

1. BÖLÜM

GİRİŞ

1.1. Genel Bilgiler

Mühendislik problemleri için geliştirilmiş muhtelif çözüm yöntemleri bulunmaktadır. Geleneksek yöntemler olarak adlandırılan bu yöntemler genellikle basit problemlerin çözümünde etkili sonuçlar sağlayabilmektedir. Oysa büyük boyutlu gerçek dünya problemleri genellikle zor problemler sınıfındadır. Bu problemlerin parametreleri arasında genellikle doğrusal olmayan ilişkiler söz konusudur ve yine bu problemlerin bazılarının matematiksel olarak modellenmesi dahi mümkün olamamaktadır. Genel olarak, kabul edilebilir bir zaman dilimi içerisinde en iyi çözümün bulunmasının garanti edilemediği problemler zor problemler olarak tanımlanır. Geleneksel yöntemler ile bu tür problemlerin çözümü ya çok uzun süreler gerektirebilmekte ya da mümkün olamamaktadır.

Bu kapsamda, tasarım, üretim, kontrol, haberleşme ve taşıma gibi pratik alanlardaki gerçek dünya problemlerinin büyük zorluğu, araştırmacıların ilgisini yapay zekâ ve tabii bilimler üzerine yoğunlaştırmıştır. Çoğunlukla insanda ve doğada var olan mekanizmaları ve sistematikleri taklit ederek geliştirilen yapay zekâ yöntemleri zor problemlerin çözümünde etkili olarak kullanılabilen araçlardır. Bu tür yöntemler zor problemler için her zaman en iyi çözümü bulmayı garanti etmese de kabul edilebilir sürelerde kabul edilebilir çözümler sağlamayı garanti edebilirler.

20. yüzyılın ikinci yarısında temelleri atılan ve günümüzde çok sayıda araştırmacının üzerinde çalıştığı yapay zekâ teknikleri, pek çok problemin çözümünde başarı ile kullanılmış ve geleneksel çözüm yöntemlerine alternatif olmuştur. Başlangıçta, tıp alanında hastalık teşhisi, uydu fotoğraflarının okunması, askeri strateji belirleme, kontrol problemleri ve endüstriyel uygulamalar gibi çok farklı amaçlar için geliştirilen bu

teknikler günümüzde hemen her disiplinde kullanılır olmuştur. Yapay sinir ağları, bulanık küme teorisi, genetik algoritmalar, uzman sistemler araştırmacılar tarafından yaygın olarak kullanılan yapay zekâ yöntemlerinden bazılarıdır [1].

Yapay zekânın bir dalı olan Adaptif Ağ Tabanlı Bulanık Mantık Çıkarım Sistemi (Adaptive Neuro-Fuzzy Inference System, ANFIS), örneklerden öğrenebilme ve öğrenmiş olduğu bilgiyi belirli kurallara bağlı olarak ağ üzerinde dağıtabilme yeteneği gibi esnek ve güçlü yapısı sayesinde makinalı öğrenme, sınıflandırma ve optimizasyon gibi alanlarda başarı ile kullanılmaktadır. ANFIS, bulanık sistemlerin ve sinirsel ağların bilinen en iyi özelliklerini birleştiren ağ tabanlı bir yapıdır. Bulanık sistemlerin, giriş bilgisini ağ üzerine dağıtarak araştırma uzayını azaltması, sinirsel ağların ise geri yayılım davranışlarını kullanarak kontrol parametrelerinin doğrusal olmayan problemler için en uygun değerlerinin bulunması ANFIS modeline önemli üstünlükler kazandırmaktadır.

Başlangıç ve sonuç parametrelerinden oluşan ANFIS parametrelerinin başarılı bir şekilde güncellenmesi ve eğitimi, problemlerin çözümü için oldukça önemlidir. ANFIS mimarisinin bir problem için eğitim başarımı büyük ölçüde yapının oluşturulmasına yönelik başlangıç koşullarına, eğitim algoritmalarına ve eğitimde kullanılan özelliklerin seçimine bağlıdır. Başlangıç koşulları ve özelliklerin seçimi ise genel olarak problemin yapısına göre değişmektedir. Bunun yanı sıra problemin amaç fonksiyonu ve kısıtlamaları çoklu veya keskin zirvelere sahip ise öğrenme maliyetini azaltmak ve yüksek doğruluk elde etmek genellikle zordur. ANFIS modelinin zor problemler için eğitilmesi, temelde modelin giriş ve çıkış parametrelerine ait en uygun değerlerin belirlenmesi sürecidir. Bu kapsamda, ANFIS'in eğitimi için Geriye Yayılım (Backpropagation, BP), En Küçük Kareler Tahmini (Least Squares Estimation, LSE), En Dik İniş (Gradient Descent, GD), Levenberg Marquardt (LM) ve Kalman Filtresi (Kalman Filter, KF) gibi türeve dayalı algoritmalar yaygın olarak kullanılmaktadır. Ancak, türeve dayalı algoritmalarda her adımda eğim hesaplanması nispeten zordur ve kullanılması gereken zincir kuralı çoğunlukla yerel minimuma takılma problemine yol açmaktadır. Ayrıca türeve dayalı algoritmalarda parametrelerin yakınsaması oldukça yavaş ve çözümün başarımı büyük oranda başlangıç değerlerine bağlı olmaktadır. Bu sebeplerden dolayı ANFIS modelinin eğitilmesi temel sorunlardan biridir.

1.2. Literatür Özeti

ANFIS parametrelerinin başarılı bir şekilde eğitilmesi için literatürde önerilmiş muhtelif yaklaşımlar ve yöntemler mevcuttur. Bu yöntemlerden bazıları türev tabanlı olmayan ve yapay zekâ optimizasyon algoritmalarından olan Genetik Algoritma (GA), Parçacık Sürüsü Optimizasyon Algoritması (PSO), Yapay Arı Kolonisi Algoritması (ABC), Karınca Koloni Optimizasyon Algoritması (ACO) ve Diferansiyel Gelişim Algoritması (DGA) gibi sezgisel algoritmalardır [2].

Simon, doğrusal olmayan dinamik sistemlerin kimliklendirilmesi problemi için türev tabanlı genişletilmiş kalman filtresi yöntemi ile bulanık sistemleri eğitmiştir. Önerilen yöntemin sonuçlarını GD algoritmasından elde edilen sonuçlar ile kıyaslamışlardır. Yapılan karşılaştırma çalışmaları sonucunda önerilen yönteminin daha hızlı yakınsadığı ve daha iyi sonuçlar verdiği, ancak daha çok hesaplama maliyeti gerektirdiği görülmüştür [2]. Surmann, pillerin şarj edilme süreleri ile ilgili bir çalışmada, kural tabanlı bulanık sistemin optimize edilmesi için GA'nın kullanıldığı bir yaklaşım önermiş ve önerilen yaklaşım ile şarj sürelerini azaltmıştır [3]. Ghomsheh ve arkadaşları ANFIS'in parametrelerinin optimizasyonu için modifiye edilmiş PSO algoritmasını önermişlerdir. Bu yeni yaklaşım ile popülasyondaki en kötü parçacığı çıkarıp yerine yeni bir parçacık eklemişlerdir. Böylelikle temel PSO algoritması ile kıyaslandığında yakınsama daha hızlı gerçekleşmiştir [4].

Shoorehdeli ve arkadaşları, doğrusal olmayan sistemlerin modellenmesinde kullandıkları ANFIS'in eğitimi için PSO tabanlı hibrid bir algoritma önermişlerdir. Hibrid yöntem PSO ve GD'nin birlikte kullanılması ile oluşturulmuştur. Bu çalışmada araştırmacılar ANFIS'in başlangıç parametrelerini eğitmek için PSO ve sonuç parametrelerini eğitmek için ise GD algoritmalarını kullanmışlardır [5].

Turki ve arkadaşları, doğrusal olmayan sistemlerin kimliklendirilmesi için ANFIS'in eğitiminde GA ve PSO algoritmalarının hibrid olarak kullanıldığı metasezgisel bir yöntem sunmuşlardır. Önerdikleri yöntemde PSO ile başlangıç ve GA ile sonuç parametrelerini optimize etmişlerdir [6]. Carrano ve arkadaşları çok amaçlı optimizasyon problemlerinin çözümünde ANFIS'in eğitimi için GA'yı kullanmışlardır. Bu yaklaşım ile

GA'nin doğal yapısının özellikle çok amaçlı optimizasyon problemleri için daha iyi olduğunu göstermişlerdir [7].

Cus ve arkadaşları, ANFIS'in parametrelerini eğitmek için ACO algoritmasını kullanmışlardır. Geliştirdikleri yöntemin performans analizini lineer olmayan sistemlerde deneyerek, önerilen yöntemin karmaşık optimizasyon problemlerinin çözümünde verimli bir şekilde kullanılabildiğini göstermişlerdir [8]. Karaboğa ve Kaya, dinamik sistemlerin kimliklendirilmesi için ABC algoritmasını kullanarak ANFIS'i optimize etmişlerdir. ANFIS'in başlangıç ve sonuç parametrelerinin eğitilmesinde ABC algoritması kullanılarak türev tabanlı algoritmalara göre daha başarılı bir eğitim gerçekleştirilmiştir [9].

Juang, bulanık sistemlerin tasarımında PSO ve ACO algoritmalarının kombinasyonundan oluşan hibrid bir yöntem sunmuştur. Önerdiği yöntem ile bulanık sistemin başlangıçtaki parametrelerin değerlerini ve kuralların sayısını ACO algoritması ile belirledikten sonra PSO algoritması ile eğitime devam ederek başlangıç ve sonuç parametrelerini optimize etmiştir. Dolayısıyla ACO algoritması sadece PSO algoritmasının ilk parçacıklarını belirlemek için kullanılmıştır. Böylelikle hızlı yakınsama sağlanarak daha kısa sürede daha iyi sonuçlar elde edilmiştir [10].

Liu ve arkadaşları geliştirdikleri Quantum-Behaved Parçacık Sürüsü Optimizasyon algoritmasını (QPSO) doğrusal olmayan dinamik sistemlerin kimliklendirilmesi için oluşturulan ANFIS modelinin eğitiminde kullanmışlardır. Önerilen yaklaşım ile temel PSO algoritmasının yerel ve global arama dengesini sağlayan σ katsayısını dinamik hale getirerek daha hızlı yakınsama ve daha iyi sonuçlar elde etmişlerdir [11]. Loganathan ve Girijia, doğrusal olmayan sistemlerin kimliklendirilmesi için GD ve Runge-Kutta öğrenme algoritmasından (RKLM) oluşan hibrid bir yöntem kullanarak ANFIS modelini optimize etmişlerdir. Kullandıkları hibrid yöntemde GD algoritmasıyla başlangıç parametrelerini, RKLM algoritmasıyla da sonuç parametrelerini eğitmişlerdir [12].

Behmanesh ve arkadaşları kaotik zaman serisi tahmini için ANFIS'in eğitiminde bölgesel arama stratejileri geliştirmiş, Emperyalist Rekabetçi Algoritması (ICA) ve LSE yönteminin birlikte kullanıldığı hibrid bir metot kullanmışlardır. Yaptıkları çalışma ile başlangıç parametrelerini ICA algoritması ile eğiterek türev tabanlı algoritmaların yerel

minimum problemini aşmışlardır. [13]. Salleh ve Hussain, KEEL veri tabanından alınan 4 farklı veri kümesini sınıflandırmak için Mine Blast Algoritmasını (MBA) kullanarak ANFIS modelini eğitmişlerdir. Sınıflandırma problemlerinin doğruluğunu artırmak için ANFIS'in tüm parametreleri MBA ile optimize edilmiştir [14].

Marzi ve arkadaşları, doğrusal olmayan fonksiyonların çözümü ve zaman serisi tahminleri için Arı Algoritması (BA) ile LSE yönteminin birlikte kullanıldığı hibrid bir yöntem (BA-LSE) ile ANFIS modelini optimize etmişlerdir. Önerilen BA-LSE yönteminde BA ile başlangıç parametrelerini ve LSE ile sonuç parametrelerini eğiterek daha başarılı sonuçlar elde edilmiştir [15]. Cardenas ve arkadaşları, otomatik bir fabrikada enerji yükü tahmini için GA ile eğitilen ANFIS ağ yapısını kullanmışlardır. Bu çalışmada enerji yükü tahmininin yapılabilmesi için çok fazla rastgele değişken kullanılmasına rağmen GA ile ANFIS modeli başarılı bir şekilde optimize edilerek yüksek oranda doğruluk ile tahmin gerçekleştirmiştir [16].

Giriş değişkenlerinin seçilmesi, giriş uzayının bölümlenmesi, girişler için üyelik fonksiyonlarının türünün ve sayısının seçilmesi, kural katmanının oluşturularak kural sayısının belirlenmesi ve üyelik fonksiyonları parametrelerinin başlangıç değerlerinin belirlenmesi gibi tüm işlemler ANFIS modelinin oluşturulması ve yapılandırılması sürecini kapsamaktadır.

ANFIS modelinin yapılandırılması için genellikle tecrübeye dayalı uzman bilgisi kullanılmaktadır. Ancak birçok durumda uzman bilgisi en doğru tercihleri yaparak ANFIS için en uygun kriterlerin belirlenmesinde yetersiz kalabilmektedir. Farklı denemeler yaparak en uygun özelliklerin seçilmesi süreci zaman almakla birlikte uygunluğu tartışılabilmektedir. Ayrıca, bu yaklaşımın dezavantajı giriş sayısı arttıkça üyelik fonksiyonu sayısı ve kural sayısının da hızlı bir şekilde artmasıdır. Bu problem ANFIS ağının eğitilmesini zorlaştırdığı gibi, hesaplama maliyetini de artırarak uzun hesaplama süreci gerektirmektedir. Dolayısıyla hesaplama maliyetinin artması kullanılan bilgisayar işlemcisi yükünü artırmakta ve daha fazla hafıza gerektirmektedir. Bu nedenlerden dolayı tecrübeye dayalı yöntemleri en uygun ANFIS modelinin oluşturulmasında kullanmak oldukça zordur. Bu yöntem giriş sayısı az olan ve / veya her bir giriş için tahsis edilen üyelik fonksiyonu sayısı az olan problemler için uygundur.

Dolayısıyla, ANFIS modelinin oluşturulması ve yapılandırılması süreci de çözülmesi gereken temel problemlerdendir.

Bu süreci otomatikleştiren ve bulanık modelin performansını ve yorumlanabilirliğini geliştirirken çalışma sürecini önemli ölçüde azaltan çeşitli yöntemler bulunmaktadır. Izgara Bölümleme (Grid Partitioning, GP), Eksiltici Kümeleme (Subtractive Clustering, SC) ve Bulanık C-Ortalamalar Kümeleme (Fuzzy C-Means Clustering, FCM) yöntemleri yaygın olarak kullanılan yöntemlerden bazılarıdır. Bu yöntemlerin arkasındaki temel fikir ise, giriş-çıkış örnek verisi arasındaki ilişkiyi öğrenerek bulanık kuralları tahmin etmektir. Giriş uzayını etkili bir şekilde bölümlemek, ANFIS'in kural sayısını ve üyelik fonksiyonlarının sayısını azaltarak, modelin hem öğrenme hem de uygulanma aşamasında hızını ve performansını artırmaktadır. Bu kapsamda son yıllarda ANFIS'in yapılandırma sürecini otomatikleştiren ve modeli daha performanslı hale getiren muhtelif yöntemler de önerilmiştir.

Abdulshahed ve arkadaşları, makine aletlerinin termal hatalarını azaltmak için termal görüntüleme kamerasından elden edilen verileri kullanan FCM ile oluşturulmuş yeni bir ANFIS modeli (FCM-ANFIS) önermişlerdir. FCM-ANFIS ile standart ANFIS ağından daha az kurala sahip, dolayısıyla daha az maliyetli bir model oluşturarak, daha başarılı sonuçlar elde edilmiştir [17]. Al-Himyari ve arkadaşları, motor koşullarını izlemede hava-yakıt oranı ve kontrol parametreleri arasındaki doğrusal olmayan bağlantıyı yakalayabilmek için FCM yöntemi ile modellenmiş ANFIS ağını kullanmışlardır. Önerilen yöntem ile daha performanslı bir ANFIS modeli geliştirilmiştir [18].

Quintero ve arkadaşları, Kolombiya enerji piyasasında optimal portföy kullanımı için FCM ile yapılandırılmış bir ANFIS modeli önermişlerdir. Bu çalışmada, ANFIS modeli GP, SC ve FCM yöntemleriyle yapılandırılarak, modellerin performansları karşılaştırılmıştır. Yapılan simülasyon çalışmaları sonucunda en iyi performansın FCM-ANFIS modeline ait olduğu görülmüştür [19]. Mehrabi ve arkadaşları, alüminyum-su nanosıvı ısı iletkenlik oranını tahmin etmek amacıyla FCM tabanlı ANFIS modelini önermişlerdir. Çalışmada, yapay sinir ağlarını GA ile optimize edip, oluşan modeli (GA-NN) yine aynı probleme uygulayarak, yöntemin performansını FCM-ANFIS ile kıyaslamışlardır. Önerilen yöntemin GA-NN yöntemine göre daha başarılı olduğu görülmüştür [20].

Karahoca, süt kabı yönetiminde hassas bir tahmin modeli geliştirmek için FCM tabanlı ANFIS modeli uygulamıştır. FCM-ANFIS yöntemi ile sistemin karmaşıklığını azaltarak ANFIS ağının daha hızlı eğitilmesini sağlamıştır. Böylelikle hem daha az maliyetli ve daha yüksek performanslı bir sistem elde edilmiştir [21]. Pinpin ve arkadaşları, ANFIS ağını daha hızlı ve verimli modellemek için bulanık kümeleme tabanlı topoloji isimli yeni bir danışmansız kümeleme (TFC) algoritması kullanmışlardır [22]. Rantala ve Koivisto, ANFIS ağının modellenmesinde daha verimli kümeleme yöntemi geliştirmek için SC yöntemini GA ile optimize ederek optimize edilmiş SC kümeleme yöntemini (OSC) önermiştir. Önerilen yöntemi doğrusal olmayan regresyon ve zaman serisi veri setleri üzerinde test ederek, daha hızlı ve performansı daha yüksek bir model elde etmişlerdir [23].

Tawafan ve arkadaşları, yüksek empedans arıza tespiti için SC yöntemi ile modellenmiş ANFIS ağını kullanmışlardır. Yapılan bu çalışma ile standart GP ile oluşturulmuş ANFIS modelinden daha başarılı sonuçlar elde etmişlerdir [24]. Berneti, rezervuar geçirgenliğini tahmin etmek için SC ve PSO tabanlı hibrid bir yaklaşım sunmuştur. Önerilen yaklaşımda ANFIS modelini oluşturan SC yönteminin küme merkezi katsayısını (r_a) PSO algoritması ile optimize etmiştir. Böylelikle, SC yönteminin r_a katsayısı probleme uygun olarak optimize edildiğinden performans artışı sağlanmıştır [25]. Gogoi ve Sarma, uydu görüntülerinden çıkarılan bilgileri kullanmak için K-ortalamalar kümeleme (K-means clustering, KMC) yöntemi ile oluşturulmuş ANFIS modelini önermişlerdir. Ancak, önerilen yöntemin diğer yöntemlere göre daha fazla süreler gerektirmesi nedeniyle gerçek zamanlı bir süreçte bu modeli kullanmanın uygun olmadığı görülmüştür [26].

ANFIS modelinin oluşturulması süreci ve parametrelerinin eğitilmesi ile ilgili temel problemlere çözüm üreten tüm bu çalışmalara rağmen, farklı türden problemlere aynı oranda başarılı sonuçlar üreten, daha performanslı ve daha hızlı ANFIS mimarileri geliştirmek için yeni yöntemlere ve çalışmalara ihtiyaç vardır.

1.3. Tezin Amacı ve Yöntemi

ANFIS'in kural sayısı, parametre sayısı ve kullanılan eğitim algoritması gibi kriterleri ağın başarımı üzerinde son derece önemli olup, modelin oluşturulması için kullanılan yöntemler bu kriterlerin belirlenmesinde doğrudan etkilidir. Kural sayısı ve parametre

sayısı dikkate alınan probleme göre değişiklik göstermektedir. Problemin giriş sayısı arttıkça kural sayısı ve optimize edilmesi gereken parametre sayısı da üstel olarak hızlı bir şekilde artmaktadır. Bu artış, ANFIS ağının eğitilmesini zorlaştırdığı gibi, hesaplama maliyetini de artırarak uzun ve karmaşık hesaplama süreci gerektirmektedir. Seçilen eğitim algoritması da ağın eğitim performansı üzerinde etkilidir. Bu konuda muhtelif çalışmalar bulunmasına rağmen halen istenilen düzeyde çözümler geliştirilememiştir.

Tez çalışmasında temel olarak, adaptif ağ tabanlı bulanık mantık çıkarım sisteminin eğitilmesinde yeni bir yaklaşım olarak benzetilmiş tavlama algoritmasının (Simulated Annealing, SA) kullanılması amaçlanmıştır. Bu kapsamda öncelikle temel bir ANFIS ağı SA algoritması ile farklı problemler için eğitilmiş ve elde edilen sonuçlar literatürdeki farklı yöntemlerden elde edilen sonuçlarla karşılaştırılmıştır. Ayrıca, temel ANFIS yapısının performansını artırmak için önerilen ve birçok problemin çözümünde daha başarılı sonuçlar sağladığı görülen FCM-ANFIS modeli de aynı türden problemler için SA algoritması ile eğitilerek önerilen yaklaşımın performansı incelenmiştir. Temelde iteratif olarak çalışan SA algoritması için yeni bir paralel algoritma modeli de önerilmiştir ve önerilen algoritmanın ANFIS eğitim performansı, dikkate alınan farklı problemler üzerinde test edilmiştir.

Sistem kimliklendirme işlemi ile dinamik bir sistemin matematiksel modelinde yer alan bilinmeyen parametreler, deneysel verilerden faydalanılarak tahmin edilir. Sistemin çıkış değeri hakkında her zaman bilgi sahibi olunabilmesi açısından, parametre değerlerinin mümkün olan en az hata ile tahmin edilmesi önemlidir. Ancak, sistemin giriş ve çıkış değerleri arasında doğrusal olmayan bir ilişki mevcutsa ve istenmeyen bozucu girişler sistemi etkiliyorsa, parametre değerlerinin minimum hata ile belirlenmesi güçleşir. Dolayısıyla, hem zaman hem de elde edilen performans açısından kimliklendirme işlemi verimsizleşmektedir. Bu sebeple, dinamik sistemlerin kimliklendirilmesi için başarıyla kullanılabilecek yöntemlere ihtiyaç vardır. Bulanık sinir ağları ve yapay sinir ağlarına dayalı yaklaşımlar dinamik sistemlerin kimliklendirilmesinde yaygın olarak kullanılan yöntemlerdendir. Ancak, sahip olduğu bazı üstün özellikler nedeniyle dinamik sistemlerin kimliklendirilmesi amacıyla ANFIS ağının kullanılmasına yönelik çalışmalar son yıllarda giderek artmaya başlamıştır. Bu tez kapsamında, dinamik sistemlerin kimliklendirilmesi amacıyla ANFIS modelinin SA algoritması kullanılarak eğitilmesi ve sonuçların farklı yöntemlerden elde edilen sonuçlar ile karşılaştırılması sunulmuştur.

DNA mikroarray teknolojisi, içerisindeki spotlar sayesinde yüzlerce geni bünyesinde barındıran bir gen çipidir. Bu teknoloji sayesinde aynı dokudan alınan sağlıklı ve hastalıklı hücrelerin gen aktiviteleri birbirleriyle karşılaştırılabilmektedir. Böylece kanser gibi hastalıklara neden olan genlerin bulunmasında DNA mikroarray teknolojisi araştırmacılara çok büyük kolaylıklar sunmaktadır. Bu yüzden mikroarray gen ifade (ekspresyon) problemlerini çözmek için yüksek performansa sahip sınıflandırma yöntemleri büyük önem teşkil etmektedir. Mikroarray gen ifade (ekspresyon) verilerini sınıflandırmak için çeşitli istatistiksel yöntemler yaygın olarak kullanılmaktadır. Ancak, bu verilerdeki gen sayılarının çokluğu ve bu veriler arasında çoğunlukla doğrusal olmayan ilişkiler bulunması gibi nedenlerden dolayı geleneksel sınıflandırma algoritmalarının başarımları çoğunlukla sınırlı kalabilmektedir. Bu kapsamda, Rutgers Üniversitesinin Biyoinformatik Laboratuvarlarında bulunan veri tabanlarından elde edilen mikroarray gen ifade verilerini sınıflandırmak için ANFIS modeli uygulanmıştır. Sınıflandırmada kullanılan ANFIS'in oluşturulmasında FCM yöntemi kullanılmıştır. Ayrıca oluşturulan ANFIS modelinin eğitilmesi için SA algoritması kullanılarak sunulan yöntemin performansı farklı algoritmaların performanslarıyla karşılaştırılmıştır.

SA iteratif tabanlı bir algoritmadır ve genellikle rastgele seçilen tek bir çözüm adayı üzerinde çalışır. Algoritma bu seri yapısı ile birçok problem için çok başarılı sonuçlar vermesine rağmen, küresel optimuma erişebilmesi için gereken süre başlangıç çözümüne bağlı olarak değişim göstermektedir. Bu durum temel SA algoritmasının dinçlik bakımından performansını düşürmektedir. Ancak, bu tür iteratif tabanlı algoritmalar genel olarak bulundukları bölgenin optimumunu popülasyon tabanlı algoritmalara göre daha kısa sürelerde bulabilmektedirler. Bununla beraber, popülasyon tabanlı algoritmaların çözümlerin bir seti ile çalışması ve rastgele araştırma algoritmalarının bir türü olması, araştırmanın küresel optimumun bulunduğu bölgeye hızla erişmesini sağlar. Ancak, bu algoritmaların da yüksek oranda olasılık tabanlı arama stratejilerine sahip olması nedeniyle bulundukları bölgedeki optimum çözümü bulabilmeleri genellikle uzun süreler gerektirebilmektedir. Tez çalışması kapsamında, SA algoritmasına popülasyon tabanlı algoritmaların paralel yapı avantajını kazandırarak yukarıda bahsedilen dezavantajını yok edip, performansını artırmak amacıyla yeni bir paralel benzetilmiş tavlama algoritması (PSA) modeli önerilmiş ve önerilen modelin performansı doğrusal ve doğrusal olmayan dinamik sistemlerin kimliklendirilmesi amacıyla ANFIS

parametrelerinin eğitilmesi için kullanılmıştır. Ayrıca, önerilen yöntemin performansı, temel SA algoritmasının ve farklı yöntemlerin performanslarıyla karşılaştırılmıştır.

SA ve önerilen PSA algoritmalarının ANFIS modelinin eğitimindeki başarıları araştırılırken, algoritmaların performansları literatürde sık kullanılan yöntemler ile karşılaştırılmıştır. Bu kapsamda, türeve dayalı algoritmalardan geri yayılım algoritması; evrimsel algoritmalardan ise genetik algoritma ve parçacık sürü optimizasyon algoritması sonuçların karşılaştırılmasında kullanılmıştır.

1.4. Tezin Bölümleri

Tez çalışmasının ikinci bölümünde çalışmanın çıkış noktası olan adaptif ağ tabanlı bulanık mantık çıkarım sistemi (ANFIS) tanıtılmıştır. Bu bölümde, bulanık mantık ve bulanık modeller hakkında genel bilgi verilmiş ve ANFIS ağının modellenmesi ve eğitilmesinde kullanılan teknikler anlatılmıştır. Ayrıca, bu tez kapsamında simülasyon çalışmaları için ANFIS ağının eğitiminde kullanılan algoritmalar incelenmiştir.

Üçüncü bölümde, ANFIS ağının SA algoritması ile eğitilmesi kapsamında, ilk olarak SA algoritması tanıtılıp, tıp bilişimi hakkında bilgi verilmiştir, Trombofili hastalığının genel yapısı detaylı bir şekilde anlatılmıştır. Sonrasında ise Trombofili hastalığı ile genetik bozukluklar arasındaki ilişkinin ANFIS ile tespit edilmesi çalışmaları anlatılmıştır. Bu kapsamda ANFIS modelinin performansı istatistiksel sınıflandırma algoritmalarının başarımları ile karşılaştırılmıştır. Ayrıca, önerilen yöntemin performansını farklı bir problem üzerinde test etmek için, sistem kimliklendirme ile ilgili genel bilgi verilip, dinamik sistemlerin kimliklendirilmesi için SA algoritması kullanılarak ANFIS ağının eğitilmesi anlatılmıştır. Çalışma kapsamında kullanılan örnek sistemlerden elde edilen sonuçlar, aynı sistemler için literatürden alınan farklı yöntemlere ait sonuçlar ile kıyaslanarak önerilen yaklaşımın başarımı hakkında bilgi verilmiştir.

Dördüncü bölümde gen ekspresyonu ve mikrodizi teknolojisi hakkında bilgi verilerek, mikrodizi gen ifade kanser verilerini sınıflandırmak için SA algoritması ile bulanık c-ortalamalar tabanlı ANFIS ağının eğitilmesi çalışmaları sunulmuştur. Bu bölümde, beş farklı mikrodizi kanser verisi için ANFIS ağı ile sınıflandırma çalışmaları yapılmış ve önerilen SA algoritmasına dayalı yaklaşımın performansı diğer yöntemlere ait sonuçlarla

karşılaştırılmıştır. Ayrıca, dinamik sistem kimliklendirilmesi için SA algoritması kullanılarak bulanık c-ortalamalar tabanlı ANFIS ağının eğitilmesi çalışmaları sunulmuştur. Bulanık c-ortalamalar kümeleme yöntemi ile modellenen ANFIS ağının sistem kimliklendirmedeki başarımı diğer yöntemlerle ve literatürden alınan farklı yöntemlere ait sonuçlar ile karşılaştırılmıştır.

Beşinci bölümde SA algoritmasına, performansını artırmak amacıyla popülasyon tabanlı algoritmaların paralel yapı avantajını kazandırarak yeni bir paralel SA algoritması modeli önerilmiştir. İteratif ve popülasyon tabanlı algoritmaların avantajlarını birleştiren bu algoritmanın performansı doğrusal ve doğrusal olmayan sistem kimliklendirme problemleri üzerinde test edilmiş ve elde edilen sonuçlar temel SA algoritmasına ve farklı yöntemlere ait sonuçlarla karşılaştırılmıştır.

Sonuç ve öneriler bölümünde ise tez çalışmasından elde edilen sonuçlar ve öneriler tartışılmıştır.

2. BÖLÜM

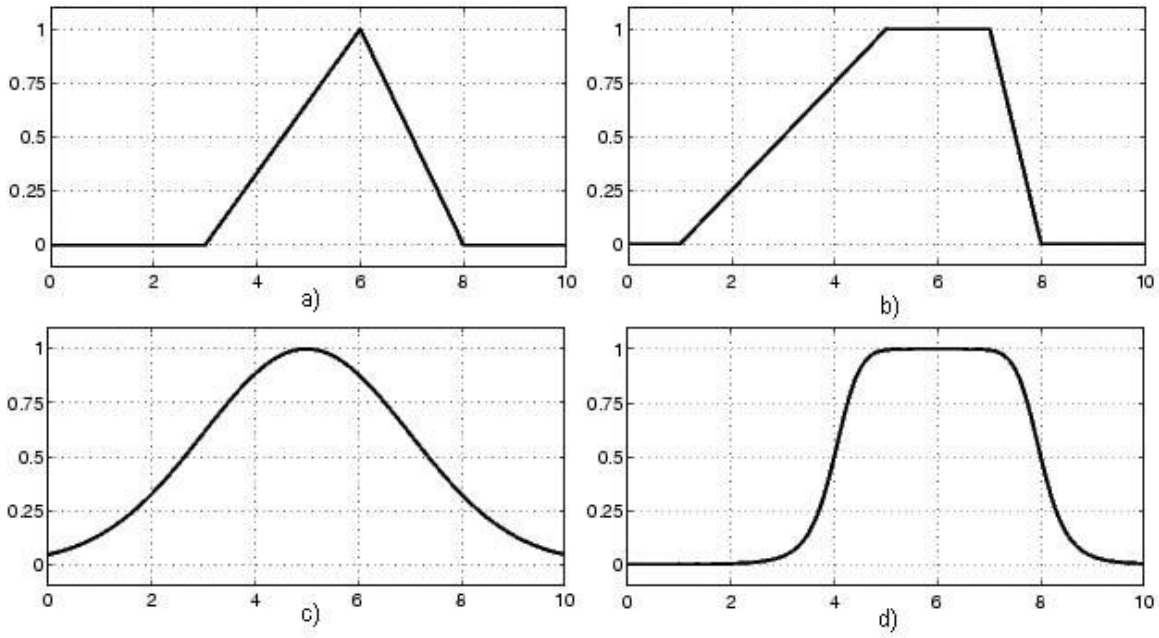
ADAPTİF AĞ TABANLI BULANIK MANTIK ÇIKARIM SİSTEMİ

2.1. Bulanık Mantık ve Bulanık Model

Bulanık küme, mantık ve sistem kavramları 1965 yılında Zadeh tarafından ortaya atılmıştır [27]. Uzun yıllar boyunca kontrol alanında çalışmalar yapan araştırmacı, istediği özelliklerde kontrol sistemi tasarlamaya çalışırken, doğrusal olmayan denklemlerin işin içine girmesi, kullanılan yöntemlerin karmaşıklaşması ve çözümün zorlaşması sonucunda bulanık mantığa yönelmiştir [28].

Geleneksel mantıkta bir kümeyi oluşturan elemanlar hakkında tanımlayıcı yargı, keskin özelliklidir. Yani bir eleman bir kümenin ya elemanıdır ya da değildir. Bu tür kümelere keskin kümeler (crisp sets) denir. Oysaki bulanık mantık yaklaşımında üye olanlar ve olmayanlar şeklinde kesin bir sınıflandırma yoktur. Bir örnek olarak orta yaş grubu insanları ele alalım. Eğer 40 yaşı orta yaş olarak kabul edersek, geleneksel kümelendirmede 30 yaşın altındaki kişiler genç, 30-50 yaş arası orta yaşlı ve 50 yaş üzeri ise yaşlı grubunda tanımlanmış olsun. Keskin kümelendirmeye göre 29,5 yaşındaki birisi genç, 30,5 yaşındaki diğeri ise orta yaşlı olarak sınıflandırılacaktır. Halbuki bu örneklerden ilki orta yaş, ikincisi ise genç yaş kümelerinde de tanımlanabilir. Benzer düşünceden hareketle, bulanık kümelerde kümenin elemanları, bu elemanlara [0,1] aralığında reel değerler atayan üyelik fonksiyonları ile tanımlanmaktadır. Elemanların sahip oldukları değerler üyelik derecesi olarak adlandırılmakta ve bu değerlere göre bir elemanın herhangi bir bulanık kümenin özelliklerini ne derece taşıdıkları belirlenmektedir. Bulanık küme teorisinde üyelik derecesinin 0 ile 1 arasında değerler alması sözel bilgilerin, problemlerin çözümü sırasında sayısal verilerle birlikte kullanılmasını mümkün kılmaktadır. Sözel ifadelerin bulanık modellere katılması bulanık mantığın diğer yöntemlerden ayrıldığı en önemli farklılıklarından birisidir [29].

Bulanık modeller oluştururken değişik formlarda üyelik fonksiyonları seçilebilir. Çok sayıda üyelik fonksiyonu bulunmasına rağmen, literatürde yaygın olarak kullanılan üyelik fonksiyonlarından bazıları, üçgen (triangular), yamuk (trapezoidal), normal dağılım (gaussian) ve çan şekilli (bell-shaped) fonksiyonlarıdır. Şekil 2.1'de bu üyelik fonksiyonları gösterilmiştir. Üçgen ve yamuk eğrileri bir vektörün fonksiyonları olup üçgen 3 ve yamuk 4 sayısal parametreye bağlıdır. Simetrik gaussian fonksiyonu 2 ve çan şekilli üyelik fonksiyonu ise 3 parametreye bağlıdır [30].



Şekil 2.1. Yaygın olarak kullanılan üyelik fonksiyonları a) Üçgen b) Yamuk c) Normal Dağılım d) Çan Şekilli

Bulanık sistemler genel olarak, mevcut verilerden seçilen girdi değişkenlerinden çıktı değişkenlerinin elde edilmesini sağlamak amacıyla bulanık küme ilkelerini kullanan sistemlerdir. Bulanık sistemlerin en büyük avantajı insan deneyimlerinin ve sözel verilerin bulanık modele katılması ile çözüme ulaşılmasıdır. Bu amaçla modelin değişkenleri bulanık alt kümeler ile ifade edilirler ve söz konusu çıkarım için klasik küme işlemlerinin genelleştirilmesiyle elde edilen bulanık küme işlemleri kullanılır [31].

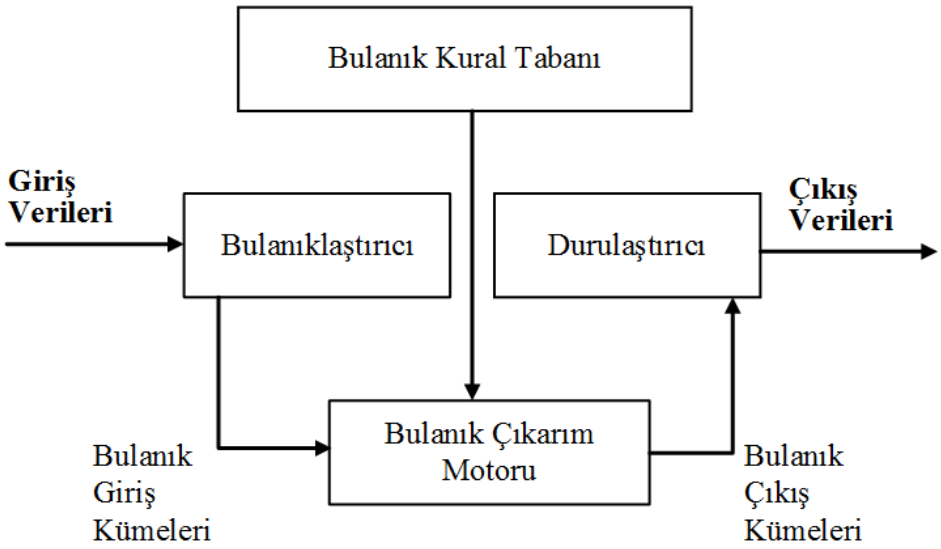
Bulanık model (bulanık çıkarım sistemi), bulanık **Eğer-İse** kuralları adı verilen bulanık kurallara dayanan sistemlerdir. Bu nedenle bulanık çıkarım sistemleri, bulanık kural tabanlı sistemler olarak da adlandırılır. Bazı kaynaklarda bulanık çıkarım sistemleri yerine bulanık model, bulanık çağrışımlı bellek, bulanık mantık kontrolör terimleri de

kullanılmaktadır [32]. Bulanık modelin temeli, bulanık **Eğer-İse** kurallarından anlaşılacağı üzere öncül ve soncul kısımlardan oluşmaktadır. Öncül kısımda sonuca sebep olan giriş değişkenleri ve bunlar arasındaki mantıksal ilişkiler, soncul kısımda ise bu giriş değişkenlerine bağlı olarak ortaya çıkan sonuç değişkenleri yer alır. Genel olarak bulanık kurallar aşağıdaki formdadır:

Kural 1: **Eğer** $x = A_1$ ve $y = B_1$ **İse** $z = N_1$

Kural 2: **Eğer** $x = A_2$ ve $y = B_2$ **İse** $z = N_2$

Burada x ve y öncül kısımdaki girdi değişkenlerince tanımlanan koşulları, z ise soncul kısımdaki çıktı değişkenlerince tanımlanan sonuçları ifade eder. Şekil 2.2'de genel bir bulanık model sisteminin yapısı gösterilmiştir.



Şekil 2.2. Bulanıklaştırma – durulaştırma birimli bulanık model sistemi [27]

Şekilde verilen sistemin bileşenleri aşağıda kısaca tanımlanmıştır:

1) Giriş Verileri: İncelenecek olayın etkilendiği girdi değişkenlerini ve bunlar hakkındaki tüm bilgileri içerir. Buradaki bilgilerin sayısal ve/veya sözel olabilmesinden dolayı genel veri tabanı da denilmektedir.

2) Bulanıklaştırıcı: Sayısal girdi değerlerini sözel olarak nitelendirilmiş bulanık kümelerdeki üyelik derecelerine atayan bir işlemcidir.

3) Bulanık Kural Tabanı Birimi: Veri tabanındaki girişleri çıkış değişkenlerine bağlayan mantıksal **Eğer-İse** türünde yazılabilen kuralların tümünü içerir. Bu kuralların yazılmasında sadece girdi verileri ile çıktılar arasında olabilecek tüm ara (bulanık küme) bağlantılar düşünülür. Böylece, her bir kural girdi uzayının bir parçasını çıktı uzayına mantıksal olarak bağlar ve bu bağlamların tümü kural tabanını oluşturur.

4) Bulanık Çıkarım Motoru Birimi: Bulanık kural tabanında giriş ve çıkış bulanık kümeleri arasında kurulmuş olan parça ilişkilerin hepsini bir arada toplayarak sistemin bir çıkışlı davranmasını temin eden işlemler topluluğunu içeren bir mekanizmadır. Bu motor her bir kuralın çıkarımlarını bir araya toplayarak tüm sistemin girdiler altında nasıl bir çıktı vereceğinin belirlenmesine yarar.

5) Durulaştırıcı: Bulanık işlemler sonucu elde edilen bulanık çıkarım sonuçlarını keskin sayısal çıkış değerlerine dönüştürür.

6) Çıkış Verileri: Bilgi ve bulanık kural tabanlarının bulanık çıkarım motoru vasıtası ile etkileşimi sonunda elde edilen çıktı değerlerinin topluluğunu belirtir.

Buradaki modelde önce, bulanıklaştırıcı katmanındaki mevcut verilerden seçilen giriş değişkenlerinin üyelik dereceleri belirlenir. Bulanık kural tabanında, giriş ve çıkış değişkenlerini mantıksal olarak birbirine bağlayan kurallar oluşturulduktan sonra, bulanık çıkarım motorunda bu kuralların hepsi bir araya toplanarak, girişlerin üyelik derecelerine göre bir çıkışı üretilir. Son olarak durulaştırıcı katmanda, bulanık çıkarım sonuçları sayısal keskin çıkış değerlerine dönüştürülür.

Bulanık mantık ile modellemenin tercih edilmesinin nedenleri aşağıdaki gibi özetlenebilir:

- Anlaşılması kolaydır.
- Dayandığı matematiksel teori basittir.
- Yaklaşım doğaldır ve karmaşıklık değildir.
- Esnektir.
- Eksik ya da yetersiz verilerle işlemler yapılabilmektedir.

- Doğrusal olmayan fonksiyonları modelleyebilir. ANFIS gibi uyarlanabilir teknikler yardımı ile herhangi bir girdi ve çıktı veri kümelerini eşleştirerek bulanık modeller oluşturulabilir.
- Uzman kişilerin görüş ve tecrübelerinden yararlanılabilir.
- İnsanların günlük işlerinde kullandığı dili kullanır ve bu durum bulanık mantığın en büyük avantajıdır.

Muhtelif modeller bulunmasına rağmen, literatürde yaygın olarak kullanılan bulanık mantık çıkarım sistemlerinden ikisi, Mamdani ve Takagi–Sugeno bulanık mantık çıkarım sistemleridir. Aşağıda bu modeller ile ilgili bilgi verilmiştir.

2.1.1. Mamdani Bulanık Modeli

Mamdani tipi bulanık model kolaylıkla oluşturulur ve insan davranışlarına çok uygundur. Bu nedenle oldukça yaygın bir kullanıma sahiptir ve diğer bulanık mantık modellerinin temelini oluşturur. İlk defa bir buhar motorunun kontrolü amacıyla, insan tecrübelerinden elde edilen sözel kontrol kuralları yardımıyla kullanılmıştır [33]. Bu modelde hem girdi değişkenleri hem de çıktı değişkeni üyelik fonksiyonları ile ifade edilir [34]. Mamdani tipi bir bulanık model aşağıda verilen beş adımda oluşturulur;

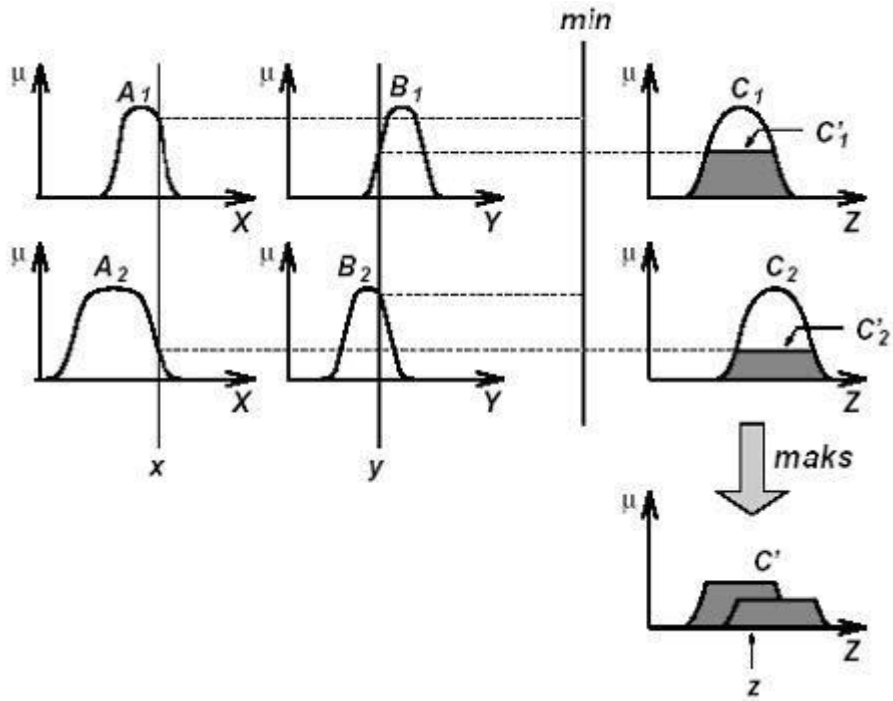
- Bulanıklaştırma: Öncül kısımdaki bütün bulanık ifadeleri kullanarak, girdi değişkenleri için [0 1] aralığında üyelik derecelerinin belirlenmesi
- Bulanık mantık işlemlerini kullanarak kural ağırlıklarının belirlenmesi
- Bulanık küme mantıksal işlemcilerinin (VE, VEYA) uygulanması
- Sonuçların toplanması: Her bir kuralın çıktısını temsil eden bulanık kümelerin birleştirilmesi
- Durulaştırma: Tek bir sayıya dönüştürülmüş toplam bulanık küme sonuçlarının durulaştırılması

Şekil 2.3'de verilen x ve y gibi sayısal iki değişkeni içeren iki kurallı bir Mamdani tipi bulanık modelde, z çıkış değerinin C_i bulanık küme fonksiyonlarından nasıl hesaplandığı gösterilmiştir. Aşağıda bu bulanık modelin hesaplanmasında kullanılan iki kural verilmiştir.

Kural 1: **Eğer** $x = A_1$ ve $y = B_1$ **İse** $z = C_1$

Kural 2: **Eğer** $x = A_2$ ve $y = B_2$ **İse** $z = C_2$

Bulanıklaştırılan giriş bilgileri önceden belirlenmiş kurallara tabi tutulur. Kurallar uygulanırken, şartlar arasındaki "VE" durumunda giriş üyelik değerlerinden en küçük olanı, "VEYA" durumunda ise en büyük olanı alınarak çıkış üyelik fonksiyonunda bu sayının altında kalan alan bulunur. Her bir kurala ait çıkış üyelik fonksiyon bölgeleri bulunduktan sonra bu alanlar çeşitli berraklaştırma yöntemleri yardımıyla kesin çıkış bilgisinin elde edilmesinde kullanılır [35].



Şekil 2.3. Mamdani bulanık mantık modeli [35]

Mamdani tipi bulanık modelin avantajları aşağıdaki gibi özetlenebilir [27]:

- Modelin oluşturulması basittir.
- Diğer bulanık mantık modellemenin temelini oluşturur.
- İnsan davranış ve duyularına uygundur.

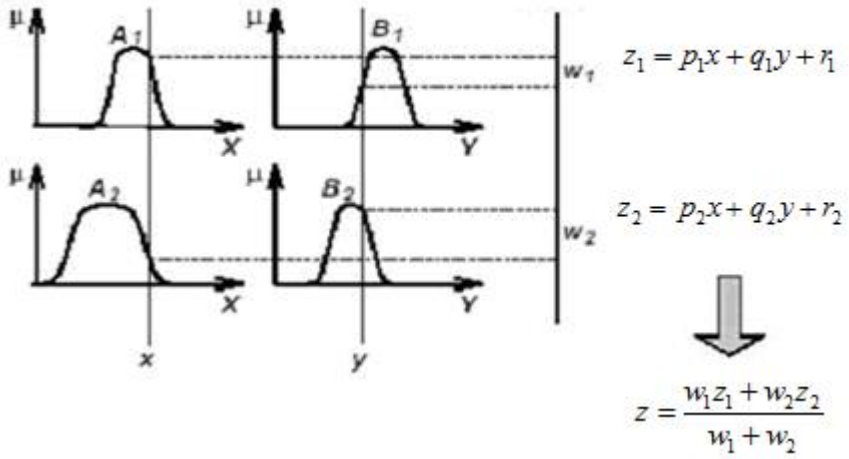
2.1.2. Takagi-Sugeno Bulanık Modeli

Sugeno bulanık modeli (TSK bulanık modeli olarak da bilinir), Takagi, Sugeno ve Kang tarafından önerilip, ilk kez 1985 yılında kullanılmaya başlanmıştır. Mamdani bulanık mantık yönteminin bir uyarlamasıdır. Sugeno tipi bulanık modellemede çıktı üyelik fonksiyonları lineer ya da sabittir. Bu model, çıktı üyelik fonksiyonları sabit olduğu zaman sıfırıncı derece, 1. derece doğru denklemi şeklinde olduğu zaman ise birinci derece Sugeno bulanık model olarak adlandırılır. Bir Sugeno bulanık modelindeki tipik bulanık kural aşağıdaki forma sahiptir,

Eğer $x = A$ ve $y = B$, **İse** $z = f(x, y) = px + qy + r(c)$

Burada A ve B, $z = f(x, y)$' ye bağlı keskin çıkış veren bir fonksiyon olmasına rağmen, giriş değişkenlerinin bulanık kümeleridirler. z çıkış değeri ise genellikle x ve y değişkenlerine bağlı bir polinomdur. Ancak, z çıkış değeri bulanık kuralın girişi tarafından belirtilmiş bölge dahilinde, bir sistemin çıkışını tanımlayabildiği sürece herhangi bir fonksiyon da olabilir.

Şekil 2.4'de Birinci dereceden Sugeno modelinin bulanık çıkarım mekanizması gösterilmektedir. Bu yöntemde hem girişim katsayılarının hem de kural ağırlıklarının uzman kişi tarafından elde edilmesi zor olacağından, bu sayıların bir optimizasyon algoritması yardımıyla bulunması gerekir. Her bir kural keskin çıkışa sahip olduğundan, sonuç keskin çıkış değeri *ağırlıklı ortalama* ile elde edilir. Bu yaklaşım Mamdani modelindeki durulandırma işleminin hesaplama yükünü ve zaman kaybını gideren basit ve fonksiyonel bir yaklaşım olup, sistem modelleme ve kontrolör tasarımına çok uygun bir çıkarım mekanizmasıdır. Pratikte bazen ağırlıklı ortalama operatörü yerine, $z = ((w_1 z_1 + w_2 z_2)/(w_1 + w_2))$ operatörü kullanılır. Bu basitleştirme işlemi, kuralların ateşleme katsayılarının toplamı "1" e yakınsamadıkça $(\sum_i w_i = 1)$, üyelik fonksiyonlarının dilsel anlamının kaybolmasına neden olabilir.



Şekil 2.4. Sugeno bulanık mantık modeli [36]

Sıfırıncı-dereceden Sugeno model komşu üyelik fonksiyonlarının şartları ile yeterince örtüşebildiği sürece kendi giriş değişkenlerine bağlı olarak düzgün bir fonksiyon olur. Mamdani modelde ise, üyelik fonksiyonlarındaki örtüşmeler, ara değerlendirmenin düzgünlüğü üzerinde belirleyici etkiye sahip değildir.

Sugeno tipi bulanık modelin avantajları aşağıda sıralanmıştır [27]:

- Numerik hesaplama için çok uygundur.
- Doğrusal olmayan sistemlerin kontrol edilmesi için doğrusal teknikler kullanılabilir.
- Optimizasyon ve uyarlanabilir tekniklerle birlikte iyi çalışır ve çıktı parametrelerini optimize ederek sonuçları iyileştirir.
- Çıktı uzayında sürekliliği garantiler.
- Matematiksel analiz için uygundur.

Sugeno tipi bulanık modelin dezavantajları ise şunlardır [27]:

- Yüksek derecedeki Sugeno bulanık modelleme kullanıldığında oldukça kompleks bir yapıya sahip olur.
- Girdi ve alt küme sayılarının artması verilerin eğitilmesini zorlaştırır, sonuçların elde edilmesi için belirlenmesi gereken soncul parametrelerin sayısı artar.
- İnsan sezgilerine çok uygun değildir.

Sugeno ve Mamdani bulanık modelleri, giriş değişkenlerinin bulanıklaştırılması ve bulanık mantık işlemleri bakımından tamamen aynıdır. Sugeno bulanık model, parametrelerinin optimize edilebilmesi kolaylığı bakımından Mamdani bulanık modelden daha avantajlıdır. İki sistemi birbirinden ayıran en önemli özellik Mamdani modelde çıkış değişkenlerinin tanımlanmasıdır. Sugeno tipi bulanık modellemede çıkış değerleri lineer ya da sabit fonksiyonken, Mamdani modelde çıkış değerleri üyelik fonksiyonları ile ifade edilir. Bu nedenle, Sugeno model Mamdani modelden daha karmaşık ve gösterim açısından daha elverişli olup, Sugeno tipi bulanık model uyarlanabilir tekniklerle birlikte kullanılabilir.

2.2. Adaptif Ağ Tabanlı Bulanık Mantık Çıkarım Sistemi (Adaptive Neuro-Fuzzy Inference System, ANFIS)

Adaptif ağ tabanlı bulanık mantık çıkarım sistemi (ANFIS), Sugeno tipi bulanık sisteminin, sinirsel öğrenme kabiliyetine sahip bir ağ yapısı olarak temsilinden ibarettir [36]. Yapısında hem yapay sinir ağları hem de bulanık mantık kullanılır [37]. Sinir ağı ile bulanık sistemin kombinasyonu *bulanık sinir* ağı olarak adlandırılır. ANFIS, öğrenme ve adaptasyon işlemlerini kolaylaştırmak için, adaptif sistemlerden yararlanan bulanık Sugeno modelidir. Böyle bir yaklaşım bulanık mantığı daha sistematik ve tecrübeye daha az bağlı hale getirmektedir [38]. ANFIS'in temel amacı, eşdeğer bulanık mantık sisteminin parametrelerini, giriş-çıkış veri kümelerini kullanarak bir öğrenme algoritması vasıtasıyla optimize etmektir. Parametre optimizasyonu, gerçek çıkış ile hedef çıkış arasındaki hata değeri minimum olacak şekilde yapılmaktadır [39].

ANFIS, bulanık çıkarım sistemindeki **Eğer-İse** kuralları ve giriş çıkış bilgi çiftlerinden oluşur. Ancak sistem eğitiminde ve denetiminde YSA öğrenme algoritmaları kullanılır [40,41]. x ve y giriş, z ise çıkış olarak alınırsa temel kural yapısı aşağıdaki şekilde yazılabilir:

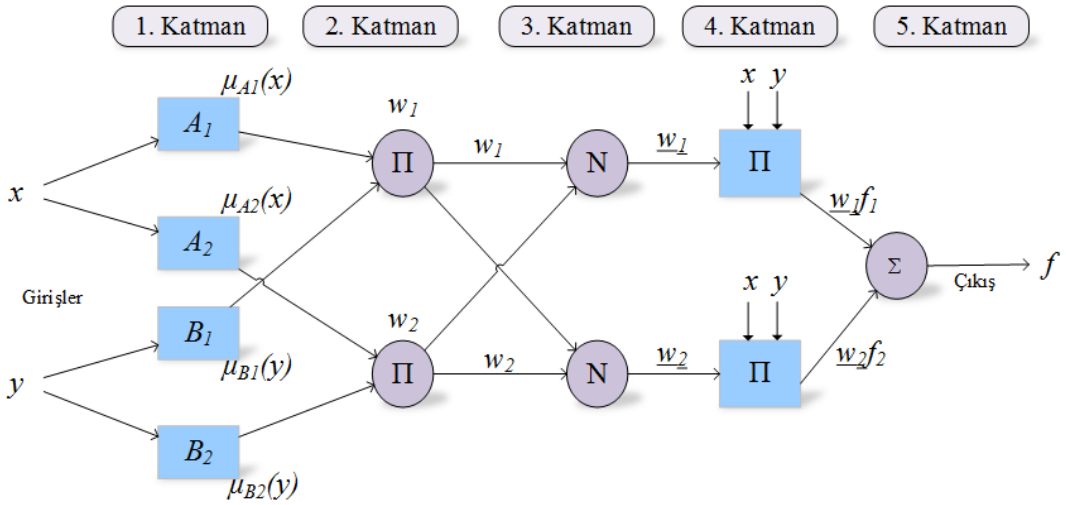
Eğer $x \in A_i$ ve $y \in B_i$ **İse** $z_i = p_i x + q_i y + r_i$

Burada A_i ve B_i sırasıyla öncül kısımdaki x ve y değişken uzayını bulanık alt uzaylara ayıran kümelerin etiketidir. p_i, q_i ve r_i eğitme işlemi boyunca belirlenen tasarım

parametreleridir. z_i ise o kurala ait çıkış değeri olup, giriş değişkenlerinin bir fonksiyonudur. Herhangi bir x, y girdi çifti için sonuç çıkış değeri ise tüm kuralların çıkış değeri olan z_i'lerin ağırlıklı ortalamasıdır [38]. z çıkış değerinin hesaplanması Eşitlik 2.1'de verilmiştir.

$$z = \sum_{m=1}^{k} w_m Z_m \Big/ \sum_{m=1}^{k} w_m \tag{2.1}$$

Aşağıda verilen iki bulanık kurala bağlı olarak, birinci derece bulanık Sugeno modeli için olası ANFIS mimarisi Şekil 2.5'de verilmiştir. ANFIS modelinin katmanlarını genel olarak özetlersek, 1. katmanda giriş verilerine üyelik fonksiyonları uygulanarak bulanıklaştırma işlemi yapılır. 2. katmanda bulanık mantık çıkarım sistemine göre kurallar oluşturulur. 3. katmanda kural katmanından gelen her bir düğüme, ağırlıklı ortalama ile normalizasyon işlemi uygulanır. 4. katmanda ise bulanık sonuçlar sayısal değerlere dönüştürülür ve son olarak 5. katmanda tüm düğümlerin çıkış değerleri toplanarak sistemin tek çıkış değeri üretilir.



Şekil 2.5. Adaptif ağ tabanlı bulanık mantık çıkarım sistemi [42]

Kural 1: **Eğer** $x = A_1$ ve $y = B_1$ **ise** $f_1 = (p_1 x + q_1 y + r_1)$

Kural 2: **Eğer** $x = A_2$ ve $y = B_2$ **ise** $f_2 = (p_2 x + q_2 y + r_2)$

Bu kurallarda p_i, q_i ve r_i her bir kural için denklem sabitleridir [39].

Şekilden görüleceği gibi ANFIS, 5 katmanlı ileri beslemeli bir yapay sinir ağı mimarisine sahiptir. Bu mimari için öğrenme algoritmasının temel görevi, ANFIS çıktısı ile öğrenme verilerini benzetmek için tüm uyarlanabilir parametreleri ayarlamaktır. Eğitim veri seti, sinir ağına tanıtılır ve herhangi bir eğitme algoritması yardımıyla ağ eğitilir. Model çıktısı ile öğrenme verileri arasındaki hata fonksiyonun minimum olduğu şartların belirlenmesi hedeflenir [43].

ANFIS yapısındaki her katmana ait düğüm işlevleri ve katmanların işleyişi sırasıyla aşağıda verilmiştir.

1. Katman

Bulanıklaştırma katmanı olarak adlandırılır. Giriş değerlerini bulanık kümelere ayırmada ANFIS modeli kullanılmıştır. Bu katmandaki hücreler, adaptif (uyarlanabilir) hücrelerdir ve hücre sayısı girdi değişkeni sayısına eşittir. Burada, her bir düğümün çıkışı, giriş değerlerine ve kullanılan üyelik fonksiyonuna bağlı olan üyelik derecelerinden oluşmaktadır. Bu tabakadaki hücrelerin çıktıları O_{1i} Eşitlik 2.2 ve Eşitlik 2.3'de verilmiştir [44].

$$O_{1i} = \mu A_i(x) \qquad i = 1,2 \tag{2.2}$$

$$O_{1i} = \mu B_{i-2}(x) \qquad i = 3,4 \tag{2.3}$$

Burada, A_i ve B_i herhangi bir bulanık küme parametresi, μA_i ve μB_i bu küme parametreleri için üyelik dereceleridir. Çan eğrisi şeklinde üyelik fonksiyonu kullanıldığında, μA_i Eşitlik 2.4 ile belirlenir.

$$\mu A_i = \frac{1}{1 + \left[\left(\frac{x - c_i}{a_i} \right)^2 \right]^{b_i}} \qquad i = 1,2 \tag{2.4}$$

Burada, a_i, b_i ve c_i sırasıyla, çan eğrisi şekilli üyelik fonksiyonun sigması, eğimi ve merkezidir. Gaussian üyelik fonksiyonu kullanıldığında, μA_i Eşitlik 2.5 ile belirlenir.

$$\mu A_i = \exp\left[-\left(\frac{x-c_i}{a_i}\right)^2\right] \qquad i=1,2 \qquad (2.5)$$

a_i ve c_i sırasıyla, gaussian üyelik fonksiyonun sigması ve merkezidir. Bu katmandaki parametreler başlangıç (premise) parametreleri olarak adlandırılır.

2. Katman

Kural katmanıdır. Bu katmandaki her bir düğüm, Sugeno bulanık mantık çıkarım sistemine göre oluşturulan kuralları ve sayısını ifade etmektedir. Bu katmandaki hücreler sabittir ve hücre sayısı kural sayısına eşittir. Hücre girdileri, kuralların öncül kısmındaki değişkenlerinin üyelik fonksiyonu değerleridir ve hücre çıktıları (O_{2i}) kuralların ağırlık derecelerini (w_i) vermektedir. Her bir kural düğümünün çıkısı μ_i Eşitlik 2.6'da verildiği gibi 1. katmandan gelen üyelik derecelerinin çarpımı ile belirlenmektedir.

$$O_{2i} = w_i = \mu A_i(x).\mu B_i(y) \qquad i=1,2 \qquad (2.6)$$

3. Katman

Normalizasyon katmanıdır. Bu katmandaki her bir düğüm, kural katmanından gelen tüm düğümleri giriş değeri olarak kabul etmekte ve her bir kuralın normalleştirilmiş ateşleme katsayısını hesaplamaktadır. Bu katmandaki hücreler de sabit hücrelerdir ve girdileri önceki tabakadan aldıkları ağırlık dereceleridir. Bu tabakada ağırlık dereceleri Eşitlik 2.7'de verilen ifadeye uygun olarak normalize edilir.

$$O_{3i} = \overline{w_i} = \frac{w_i}{w_1 + w_2} \qquad i=1,2 \qquad (2.7)$$

4. Katman

Durulaştırma katmanıdır. Durulaştırma katmanındaki her bir düğümde verilen bir kuralın ağırlıklandırılmış sonuç değerleri hesaplanmaktadır. Bu katmandaki hücreler adaptif hücrelerdir ve bu hücrelerin çıktısı Eşitlik 2.8'deki şekilde hesaplanır.

$$O_{4i} = \overline{w_i}.f_i = \overline{w_i}.(p_i x + q_i y + r_i) \qquad i=1,2 \qquad (2.8)$$

5. Katman

Toplam katmanıdır. Bu katmanda sadece bir düğüm vardır ve $\sum$ ile etiketlenmiştir. Burada, 4. katmandaki her bir düğümün çıkış değeri toplanarak, sonuçta ANFIS sisteminin gerçek değeri elde edilir. Sistemin çıkış değeri ise aşağıda verilen formüle göre hesaplanmaktadır [45-51].

$$O_{5i} = f = \sum \overline{w_i} . f_i = \frac{\sum w_i . f_i}{\sum w_i} \qquad i = 1,2 \tag{2.9}$$

2.2.1. ANFIS Modelini Kimliklendirme Yöntemleri

Problem çözümüne yönelik uygulama aşamasında, ANFIS modelinin oluşturulması ve yapılandırılması süreci çözülmesi gereken temel problemlerdendir. Giriş değişkenlerinin seçilmesi, giriş uzayının bölümlenmesi, girişler için üyelik fonksiyonlarının türünün ve sayısının seçilmesi, kural katmanının oluşturularak kural sayısının belirlenmesi ve üyelik fonksiyonları parametrelerinin başlangıç değerlerinin belirlenmesi gibi tüm bu işlemler ANFIS modelinin oluşturulması aşamasını kapsamaktadır. ANFIS modelinin yapılandırılması için genellikle tecrübeye dayalı bir uzman bilgisi kullanılmaktadır. Ancak, birçok durumda uzmanların bilgisi en doğru tercihleri yaparak ANFIS modeli için en uygun kriterlerin belirlenmesinde yetersiz kalabilmektedir. Farklı denemeler yaparak en uygun özelliklerin seçilmesi süreci ise zaman almakla birlikte uygunluğu da tartışmalıdır. Bu süreci otomatikleştiren ve bulanık modelin performansını ve yorumlanabilirliğini geliştirirken çalışma sürecini önemli ölçüde azaltan çeşitli yöntemler bulunmaktadır. Izgara Bölümleme (Grid Partitioning, GP), Eksiltici Kümeleme (Subtractive Clustering, SC) ve Bulanık C-Ortalamalar Kümeleme (Fuzzy C-Means Clustering, FCM) yöntemleri yaygın olarak kullanılan bu yöntemlerden bazılarıdır. Bu yöntemlerin arkasındaki temel fikir ise, giriş-çıkış örnek verisi arasındaki ilişkiyi öğrenerek bulanık kuralları tahmin etmektir. Giriş uzayını etkili bir şekilde bölümlemek, genel olarak ANFIS'in kural sayısını ve üyelik fonksiyonlarının sayısının azaltarak, modelin hem öğrenme hem de uygulanma aşamasında hızını ve performansını artırmaktadır. İlerleyen kısımlarda bu yaklaşımlar hakkında bilgi verilmiştir.

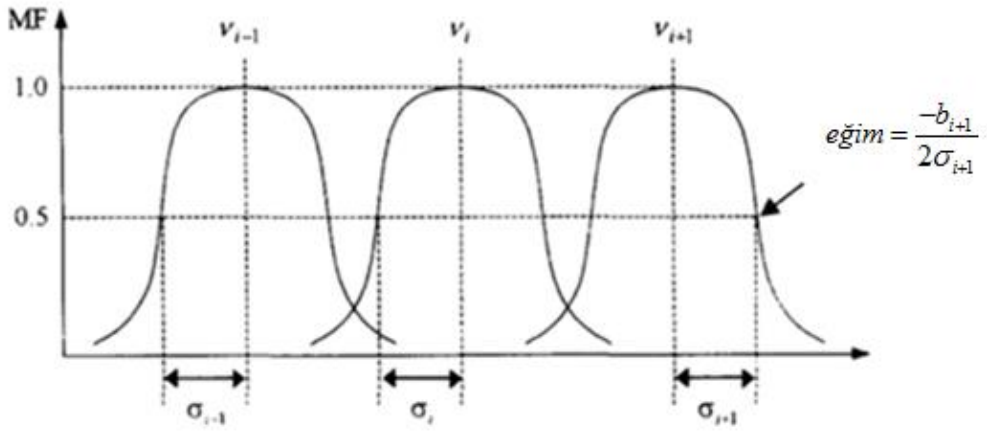
2.2.1.1. Izgara Bölümleme (Grid Partitioning, GP)

Üyelik fonksiyonlarının (MFs) başlangıç parametrelerini oluşturmanın en doğru yolu, seçilen küme sayısı kadar giriş alanını eşit bölümlere ayırmaktır. Her bölüm bulanık bir **Eğer-ise** (if-then) kuralını sunmaktadır. Üyelik fonksiyonlarının merkezleri her veri boyut aralığı boyunca eşit olarak yerleştirilecek şekilde üyelik fonksiyonlarının parametreleri ayarlanır. Her giriş için üyelik fonksiyonları elde edildikten sonra kural kümesini oluşturmak için tüm girişlerin üyelik fonksiyonlarının kombinasyonları numaralandırılır. Verilen c adet üyelik fonksiyonunun (kümenin), küme merkezleri (v_i) ve genişlik değerlerinin (σ_i) hesaplanması için aşağıda kullanılan eşitlikler verilmiştir [52].

$$v_i = U_{\min} + \frac{U_{\max} - U_{\min}}{c-1}(i-1) \qquad 1 \le i \le c \tag{2.10}$$

$$\sigma_i = \frac{U_{\max} - U_{\min}}{2(c-1)} \qquad 1 \le i \le c \tag{2.11}$$

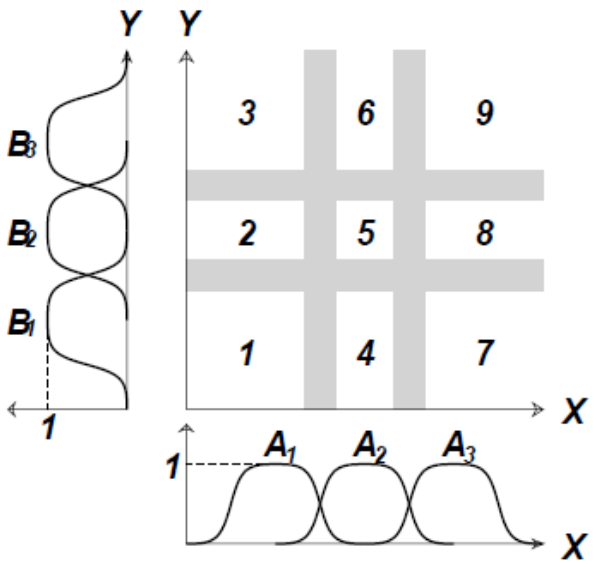
Buradaki $U_{\min}$ ve $U_{\max}$ veri boyutunun maksimum ve minimum değerleridir. Merkez ve genişlik değerleri ayarlanan *gaussian* türünde üyelik fonksiyonları Şekil 2.6'da gösterilmiştir.



Şekil 2.6. Gaussian üyelik fonksiyonlarındaki parametrelerin hesaplanmış hali [52]

Veri kümeleri eşit örneklenmiş ise GP yöntemi basit ve etkilidir. Ancak, çoğu durumda veri dağılımı hakkında ön bilgi bulunmamaktadır. Diğer bir problem ise veri giriş sayısı

arttıkça kural sayısının ve optimize edilmesi gereken parametre sayısının da katlanarak artmasıdır. Örneğin, iki girişli ve her girişine 3 üyelik fonksiyonu uygulanan bir ANFIS modelinde toplam 3^2=9 **Eğer-İse** (If-Then) bulanık kuralı oluşmaktadır. Bu sayı 3 giriş olduğunda 3^3=27 olacaktır. Şekil 2.7'de dokuz kural üreten her girişin üyelik fonksiyonları arasındaki ilişki gösterilmiştir.



Şekil 2.7. GP yöntemi ile oluşturulmuş dokuz kural bölgesi [53]

Bu yaklaşımın dezavantajı, giriş sayısı arttıkça kural sayısının da hızlı bir şekilde artmasıdır [54]. Dolayısıyla, kullanılan bilgisayar işlemcisi ve hafıza gereksinimi de katlanarak artmaktadır. Ayrıca, giriş alanını bölümlemek için kümelerin sayısı hakkında bilgiye de ihtiyaç vardır. Bu nedenlerden dolayı, GP yöntemini en uygun ANFIS modelinin oluşturulmasında kullanmak oldukça zordur. Bu yöntem giriş sayısı az olan ve / veya her bir giriş için tahsis edilen üyelik fonksiyonu sayısı az olan problemler için daha uygundur.

2.2.1.2. Eksiltici Kümeleme (Subtractive Clustering, SC)

SC, üyelik fonksiyonu (küme) merkezlerinin başlangıç konumlarını ve sayısını tahmin ederek veri kümesinden otomatik bulanık kurallar çıkarmak için Chiu [55] tarafından geliştirilmiş bulanık kümeleme yöntemlerinden biridir. SC algoritması, birçok komşu noktaları ile küme merkezi olma potansiyeline sahip bir veri noktası içindeki veri

noktalarının yoğunluk ölçümüne dayanır. Genel olarak algoritma aşağıdaki gibi çalışır [56].

1. M boyutlu bir uzayda n veri noktalı $\{x_1, x_2, ..., x_n\}$ bir diziyi ele aldığımızda, her veri noktası küme merkezleri için bir adaydır. x_i veri noktasındaki yoğunluk ölçümü aşağıdaki gibi tanımlanır.

$$D_i = \sum_{j=1}^{n} \exp\left(-\frac{\left\| x_i - x_j \right\|^2}{\left(\frac{r_\alpha}{2} \right)^2} \right) \tag{2.12}$$

   Burada r_a komşuluk yarıçapını gösteren pozitif bir sabittir. Bir veri noktasına yakın birçok veri noktası varsa yani veri noktasının yoğunluk değeri yüksek ise o veri noktası küme merkezi olma potansiyelindedir.

2. x_{c1} birinci küme merkezi ve D_{c1} ise onun potansiyel değeri (en yüksek yoğunluk değerine sahip olma) olsun. Her bir veri noktası için potansiyel yani yoğunluk değeri aşağıdaki gibi revize edilir.

$$D_i = D_i - D_{c_1} \exp\left(-\frac{\left\| x_i - x_{c_1} \right\|^2}{\left(\frac{r_b}{2} \right)^2} \right) \tag{2.13}$$

   Burada r_b yoğunluk ölçüsünde ölçülebilir azalmaları olan komşulukları tanımlayan pozitif bir sabittir.

Bu kümeleme yönteminde küme sayıları otomatik olarak belirlense de yöntemi kullanan kullanıcı oluşturulan küme sayısını kuvvetle etkileyen r_a parametresini belirlemek zorundadır. Eğer r_a parametresine büyük bir değer atanırsa, genellikle daha az küme ile sonuçlar alınır ve dolayısıyla sıradan bir model elde edilir. Küçük değerli bir r_a parametresi ise aşırı sayıda kümesi olan bir model üretebilir ancak bu modelin genelleme yeteneği iyi olmayacaktır. Buna ek olarak SC yönteminde, verinin kendi sınıflarına göre

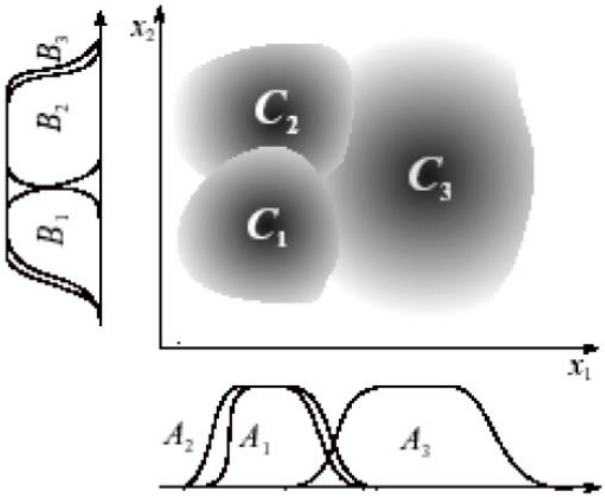
gruplara ayrılması gerekmektedir. Kümeleme yöntemi daha sonrasında kural çıkarmak için tek bir veri grubuna uygulanır [56].

Kümeleme yönteminde, kurallar verilen girişin merkezi ile bire bir eşleştirilerek elde edilen küme merkezinden otomatik olarak oluşturulur. Şekil 2.8'de örnek olarak verilen üç bulanık kural ile oluşturulmuş üç küme merkezi gösterilmektedir.

Kural 1: **Eğer** $a = A_1$ ve $b = B_1$ **İse** $c = C_1$

Kural 2: **Eğer** $a = A_2$ ve $b = B_2$ **İse** $c = C_2$

Kural 3: **Eğer** $a = A_3$ ve $b = B_3$ **İse** $c = C_3$



Şekil 2.8. Kümeleme temelli bölümleme [53]

Bulanık kümeler, kümelerin yansıması ile her kural için ayrı ayrı üretildiğinden kümele yöntemi yorumlanması zor olan bir kural üretme eğilimindedir. Bu nedenle, parametre tanımlama veya verilen kısıtlar altında bulanık kümeleri değiştirme sürecinde daha hassas bulanık modelleme elde etmek için başlangıç sınıflandırıcıyı eğitmede bir öğrenme işlemi uygulanır [56-58].

2.2.1.3. Bulanık C-Ortalamalar Kümeleme (Fuzzy C-Means Clustering -FCM)

Bulanık C-Ortalamaları (Fuzzy C-Means- FCM) yöntemi her bir veri noktasının bir kümeye belirlenen üyelik derecesi kadar ait olmasına sağlayan iteratif tabanlı bir kümeleme algoritmasıdır. Bulanık mantık prensibi gereği her veri, kümelerin her birine

[0,1] arasında değişen birer üyelik değeri ile aittir. Bir verinin tüm sınıflara olan üyelik değerleri toplamı "1" olmalıdır. Nesne hangi küme merkezine yakın ise o kümeye ait olma üyeliği diğer kümelere ait olma üyeliğinden daha büyük olacaktır. Bu yöntem ilk olarak 1973 yılında Dunn tarafından tanıtılmış ve önceki kümeleme yöntemlerine bir geliştirme olarak 1981'de Bezdek tarafından geliştirilmiştir [59]. Bu yöntem genellikle başlangıç üyelik fonksiyonlarını belirlemek için kullanılır [52, 57]. FCM aşağıdaki amaç fonksiyonunu minimize etmeye dayalı danışmanlı bir kümeleme yöntemidir.

$$J(U,V)=\sum_{i=1}^{n}\sum_{j=1}^{c}\left(\mu_{ij}\right)^{m}\left(d_{ij}\right)^{2} \tag{2.14}$$

Burada:

$X = x_1, x_2, \dots, x_n$: n adet veri noktasının dizisi,

$V = v_1, v_2, \dots, v_c$: küme merkezlerinin elemanları,

$\mu_{ij}\varepsilon[0,1]$: küme merkezi v_j ile x_i veri noktasının bir üyelik derecesi,

d_{ij}: x_i ile v_j arasındaki öklid uzaklığı, $d_{ij} = \|x_i - v_j\|$ dir.

m: bulanıklık indeksi, her bir veri noktasının üyelik derecelerinin bulanıklıklarını kontrol etmek için kullanılır.

$U = (m_{ij})_n xc$: bulanık bölüm matrisi, her bir veriden tüm küme merkezlerine ait üyelik derecelerinin tümünü kapsar.

Bulanık bölümleme, üyelik derecesi μ_{ij} ve küme merkezleri v_j güncellenerek yukarıda verilen J fonksiyonunu optimize etmek için yapılır.

$$v_j=\frac{\sum_{i=1}^{n}\left(\mu_{ij}\right)^{m}x_i}{\sum_{i=1}^{n}\left(\mu_{ij}\right)^{m}} \tag{2.15}$$

$$\mu_{ij}=\frac{1}{\sum_{k=1}^{c}\left(\frac{d_{ij}}{d_{jk}}\right)^{2/(m-1)}} \tag{2.16}$$

Algoritma yukarıda verilen iki koşulu kullanarak durdurma kriterleri sağlanana kadar iteratif olarak çalışır. Güncellenmiş ve önceki amaç fonksiyonu J arasındaki fark önceden tanımlanmış minimum eşik değerinden az ise veya maksimum iterasyon sayısına ulaşıldığı anda durdurma kriterleri sağlanmış olur. FCM yöntemini kullanırken kullanıcı oluşturulacak kümelerin sayısının da belirlemesi gerekmektedir. Bununla beraber, küme geçerlilik kriteri, veri seti için en uygun küme sayısını bulmak için kullanılır.

2.2.2. ANFIS Modelinin Eğitilmesi

ANFIS'in başlangıç ve sonuç parametreleri olmak üzere eğitilmesi gereken iki parametre türü vardır. Başlangıç parametreleri, Eşitlik 2.4'de $\{a_i, b_i, c_i\}$ olarak verilen üyelik fonksiyonlarına aittir. Buradaki a_i üyelik fonksiyonlarının varyansı, c_i üyelik fonksiyonlarının merkezi ve b_i ise diğer bir eğitim parametresidir. Başlangıç parametrelerinin toplam sayısı tüm üyelik fonksiyonlarındaki parametrelerinin toplam sayısına eşittir. Sonuç parametreleri ise Eşitlik 2.8'de $\{p_i, q_i, r_i\}$ olarak gösterilen ve Şekil 2.5'deki ANFIS ağının durulaştırma katmanında kullanılan çıkış polinomuna ait parametrelerdir. ANFIS'in eğitim sürecinde başlangıç ve sonuç parametrelerinin güncellenmesi için birçok yöntem önerilmiştir. ANFIS yapısının parametrelerini eğitmek için Jang, kendi hesaplama karmaşıklığına göre aşağıda verilen dört yöntemi önermiştir [40]:

1. Sadece Geriye Yayılım (Backpropagation, BP) algoritmasının kullanılması: Tüm parametreleri BP ile güncellenmiştir.
2. BP'nin ve sadece tek geçiş için En Küçük Kareler Tahmini (Least Squares Estimation, LSE) yönteminin kullanılması: Sonuç parametrelerinin başlangıç değerlerini almak için sadece en başta bir defalığına LSE uygulanır ve sonrasında tüm parametreleri güncellemek için BP eğitim sürecinde tek algoritma olarak kullanılır.
3. BP ve LSE yöntemlerinin birleşiminden oluşan hibrid öğrenme algoritmasıdır.
4. Sıralı LSE: tüm parametreleri güncellemek için genişletilmiş Kalman filtresi (EKF) algoritması kullanılmıştır.

Bu yöntemler ile ANFIS parametrelerinin güncellenmesinde BP, LSE, En Dik İniş (Gradient Descent, GD), Levenberg Marquardt (LM) ve Kalman Filtresi (Kalman Filter,

KF) gibi türeve dayalı algoritmalar yaygın olarak kullanılmaktadır. Ancak, türeve dayalı algoritmalarda her adımda eğim hesaplanması nispeten zordur ve kullanılması gereken zincir kuralı çoğunlukla yerel minimuma takılma problemine yol açmaktadır. Ayrıca türeve dayalı algoritmalarda parametrelerin yakınsaması oldukça yavaş ve çözümün başarımı büyük oranda başlangıç değerlerine bağlı olmaktadır. Bu nedenle türeve dayalı geleneksel algoritmalar ile ANFIS'in parametrelerinin eğitilmesi ve güncellenmesi temel sorunlardan biridir.

ANFIS'in başarılı bir şekilde eğitilmesi ve parametrelerin güncellenmesi konusunda farklı yöntemlerin uygulanmasına ihtiyaç olup, bu konuda yapılmış muhtelif çalışmalar ve önerilmiş yöntemler bulunmaktadır. Bu yöntemlerden bazıları yapay zekâ optimizasyon algoritmalarından olan Genetik Algoritma (GA), Parçacık Sürüsü Optimizasyon Algoritması (PSO), Karınca Koloni Optimizasyon Algoritması (ACO) ve Diferansiyel Gelişim Algoritması (DGA) gibi sezgisel algoritmalara dayalıdır [2]. ANFIS parametrelerinin optimizasyonunda, BP, LSE, KF yada birden fazla matematiksel optimizasyon yönteminin birleşmesinden oluşan hibrid öğrenme algoritmaları gibi değişik yöntemler de kullanılmaktadır [51]. Ayrıca, birçok çalışmada evrimsel algoritmalar ve GD, LSE ve KF gibi türev tabanlı algoritmaların birlikte kullanıldığı hibrid yöntemler de önerilmiştir. Örneğin, Shoorehdeli ve arkadaşları ANFIS'in başlangıç parametrelerini PSO ile sonuç parametreleri ise GD ile eğiterek hibrid bir optimizasyon algoritması önermişlerdir [60,61]. Araştırmacılar, diğer bir çalışmalarında ise PSO ve LSE yöntemlerini birlikte kullanarak ANFIS parametrelerini güncellemişlerdir [62].

2.2.2.1. Klasikleşmiş Algoritmalar

Yaygın olarak kullanılan gelenekselleşmiş algoritmalardan ikisi BP ve Hibrid öğrenme (HB) algoritmalarıdır. Aşağıda bu algoritmalar hakkında bilgi verilmiştir.

2.2.2.1.1. Geriye Yayılım Algoritması

İlk olarak Rumelhart ve arkadaşları (1986) tarafından önerilen BP öğrenme modeli, yapay sinir ağı öğrenme modelleri arasında en çok kullanılanlardandır. BP modelinin keşfi yapay sinir ağları için çok önemli tarihsel gelişmelerden biri olmuştur [63]. Hatayı geriye

yayma ağları, temel olarak girdi örneği ile birlikte ağa sunulmak üzere sistemin istenen çıktısını gerektiren denetimli öğrenme stratejisini kullanan, birçok katmandan oluşan tamamıyla bağlantılı ileri beslemeli ağlardır. Başka bir deyişle, bir katmandaki her nöron alt katmandaki bütün nöronlara bağlanır. Hiçbir geri besleme bağlantısı ve aynı katmandaki nöronlar arasında da hiçbir bağlantı yoktur.

BP öğrenme modeli, çıkıştan alınan değerler ile beklenen değerlerin durumlarına göre hatayı çıkıştan geriye doğru azaltmaya çalıştığından dolayı bu ismi almıştır. Örnek veri setinden elde edilen hata değeri ile eğim düşüm bilgisini hesaplayarak parametre güncelleme işlemini yapmaktadır. Bir BP algoritmasının temel işlem adımları Şekil 2.9'da verilmiştir.

Adım 1. REPEAT
Adım 2. Eğitim kümesinden rasgele bir sonraki örneği seç ve ağ girişine giriş vektörü uygula
Adım 3. Ağın çıkışını hesapla
Adım 4. Ağın çıkışı ile hedef çıkış arasındaki hatayı hesapla
Adım 5. Ağın ağırlıklarını geriye gönderilen hata yardımıyla güncelle
Adım 6. UNTIL Durdurma kriteri sağlanıyor **ise** dur,
Aksi takdirde Adım 1'e git.

Şekil 2.9. Geriye yayılım algoritmasının temel adımları

BP algoritmasında öğrenme mekanizması, ağın gerçek ve istenen çıktıları arasındaki hatayı en aza indirgeyen iteratif eğim düşme yöntemine dayanır. Bu algoritmada hatalar çıkıştan geriye doğru gönderilerek ağırlıkların ayarlanmasıyla hata azaltılmaya çalışılmaktadır. Öğrenme kuralında ağ çıkışında hesaplanan hata, ağırlıkların yeni değerlerinin hesaplanmasında kullanılmaktadır. Geri yön geçişinde giriş parametreleri için güncelleme formülü aşağıdaki gibidir;

$$\Delta\alpha = -\eta \frac{\partial E}{\partial \alpha} \tag{2.17}$$

Burada α herhangi bir giriş parametresini, η öğrenme oranını, E ise ağın çıkışındaki hata değerini göstermektedir [64].

2.2.2.1.2. Hibrid Öğrenme Algoritması

ANFIS ağına ait parametrelerin güncellenmesinde, iki aşamalı bir öğrenme algoritması olan HB algoritması kullanılmıştır. HB algoritmasında, ANFIS ağ yapısına ait parametreler, başlangıç ve sonuç parametreleri olarak iki kısımda ele alınır. Toplam parametre kümesi; $S = S_1 + S_2$ şeklinde ifade edilecek olursa, S_1 başlangıç parametrelerine, S_2 de sonuç parametrelerine karşılık gelir. GD ve LSE yöntemlerinin birlikte kullanımı ile ANFIS'in tüm parametrelerini güncellenerek algoritmanın öğrenme süreci tamamlanır. GD lineer olmayan başlangıç parametrelerinin düzenlenmesinde, LSE ise lineer sonuç parametrelerinin düzenlenmesinde kullanılmaktadır [65,66]. HB algoritmasının öğrenme aşamaları ileri yön geçişi ve geri yön geçişi olmak üzere ikiye ayrılır. Tablo 2.1.'de bu öğrenme süreci özetlenmiştir.

Tablo 2.1. Hibrid öğrenme algoritmanın iki geçişli öğrenme süreci

	İleri Yön Geçişi	**Geri Yön Geçişi**
Başlangıç Parametreleri	Sabit	GD
Sonuç Parametreleri	LSE	Sabit
Sinyaller	Düğüm Çıkışları	Hata Sinyalleri

HB algoritmasının ilk aşamasında (ileri yön geçişinde) ANFIS ağına ait başlangıç parametreleri, ikinci aşamada (geri yön geçişinde) ağın sonuç parametreleri güncellenir. HB algoritmasının ileri yön geçişini oluşturan ve LSE yöntemi kullanılmak suretiyle gerçekleştirilen kısmında, ağın girişindeki üyelik fonksiyonlarına ait parametreler ya da kısaca S_1 ile belirtilen giriş parametreleri sabit tutulur. Böylece ağın çıkışı, S_2 parametre kümesi içinde bulunan çıkış parametrelerinin, lineer bir kombinasyonu haline gelir. Modellenecek sisteme ait P tane giriş-çıkış verisi ya da eğitim verisi ile S_1 giriş parametrelerinin verilen değerleri kullanılarak ağın çıkışı, matrisel formda Eşitlik 2.18'deki gibi ifade edilebilir.

$$A\theta = B \tag{2.18}$$

Matrisel eşitlikte yer alan θ vektörü, S_2 çıkış parametrelerinin elemanlarından oluşan bilinmeyen vektördür. Bu denklem, standart lineer en küçük-kareler (least-squares)

problemini gösterir ve θ için en iyi çözüm, $\| A\theta - B \|^2$ denkleminin minimum değeri olan, en küçük kareler tahmini (LSE), θ^* dır.

$$\theta^* = (A^T A)^{-1} A^T B \tag{2.19}$$

Burada A^T, A matrisinin transpozu ise ve eğer $A^T A$ singuler değilse, $(A^T A)^{-1} A^T$, A matrisinin sahte inversidir. Bunun yerine, yinelemeli olarak LSE formülü de kullanılabilir. Özel olarak Eşitlik 2.18'deki A matrisinin *i.* satırı a_i^T ile *B* matrisinin *i.* elemanı b_i^T ile gösterilir ise, bu taktirde θ vektörü iteratif olarak aşağıdaki gibi hesaplanabilir.

$$\theta_{i+1} = \theta_i + S_{i+1} a_{i+1} (b_{i+1}^T - a_{i+1}^T \theta_i) \tag{2.20}$$

$$S_{i+1} = S_i - \frac{S_i a_{i+1} a_{i+1}^T S_i}{1 + a_{i+1}^T S_i a_{i+1}} \qquad i = 0,1,\ldots,P-1 \tag{2.21}$$

Burada en küçük kareler tahmini θ^*, θ_P vektörüne eşittir. Eşitlik 2.20'de ihtiyaç duyulan başlangıç şartları şunlardır; $\theta_0 = 0$ ve $S_0 = \gamma I$. Burada γ, pozitif büyük bir sayı, I ise MxM boyutunda birim matristir.

HB algoritmasının geri yön geçişinde ise S_2 çıkış parametreleri sabit tutulur ve ağın çıkışındaki hata sinyali geriye doğru yayılarak, eğim düşümü (gradient descent) yöntemiyle giriş parametreleri güncellenir. Geri yön geçişinde ise giriş parametreleri için güncelleme formülü Eşitlik 2.17'de verilmiştir.

2.2.2.2. Sezgisel Algoritmalar

Büyük boyutlu gerçek dünya problemleri genellikle zor problemler sınıfındadır. Bu problemlerin parametreleri arasında doğrusal olmayan ilişkiler söz konusu olmaktadır ve yine bu türdeki problemlerin bazılarının matematiksel olarak modellenmesi dahi mümkün değildir. Genel olarak, kabul edilebilir bir zaman dilimi içerisinde en iyi çözümün bulunmasının garanti edilemediği problemler zor problemler olarak tanımlanır. Türeve dayalı geleneksel optimizasyon algoritmaları ile bu tür problemlerin çözümü ya çok uzun

süreler gerektirebilmekte ya da mümkün olamamaktadır. Bu nedenle gerçek dünya problemlerinin büyük zorluğu, araştırmacıların ilgisini doğada ve insanda var olan zeki sistematikleri ve davranışları taklit eden yapay zekâ optimizasyon algoritmalarına kaydırmıştır.

Yapay zekâ optimizasyon algoritmaları sezgisel algoritmalardır. Sezgisel kelimesinin İngilizcesi olan "*heuristic*", Yunanca kökenlidir ve bulmak, keşfetmek anlamına gelmektedir. Ancak günümüzde sezgisel, herhangi bir şeyin bulunmasını garanti etmeyen bir "arama" (seeking) metodu olarak tanımlanmaktadır. Sezgisel, en iyi çözüme yakın çözümleri araştıran bir tekniktir ve bu işlemi en iyi sonucu bulmayı garanti etmeksizin makul bir hesaplama maliyeti ile yapar. Günümüze kadar, ayrık problemlerin çözülmesinde sezgisel metotları kullanan çok fazla sayıda çalışma yapılmıştır. Bu problemlerin çözümünde sezgisel teknikler kullanılarak en iyi çözüme yakın çözümler bulunabilmesi, modern sezgisel yaklaşımların artarak gelişmesini sağlamış ve bu metotların gücünde ve etkinliğinde önemli bir yükseliş ortaya çıkmıştır. Genetik algoritma, benzetilmiş tavlama algoritması, karınca koloni algoritması ve parçacık sürüsü optimizasyon algoritması araştırmacılar tarafından yaygın olarak kullanılan sezgisel algoritmalardan bazılarıdır.

2.2.2.2.1. Genetik Algoritma

1970'li yıllarda temel ilkeleri John Holland tarafından ortaya konulan GA pek çok problem türüne başarı ile uygulanmaktadır [67]. GA, optimizasyon veya arama probleminde tam yada yaklaşık sonuçlar bulabilmek için kullanılan sezgisel bir algoritmadır. Bu algoritma kalıtım, mutasyon, seçim ve çaprazlama gibi evrimsel biyolojideki tekniklerden esinlenerek geliştirilmiştir. GA arama uzayı büyük olan ve ayrıca değişken sayısı çok fazla olan, çok boyutlu problemlere bile oldukça rahat uygulanabilmektedir. Problemlerin arama uzaylarında, uzayın tamamını aramak yerine daha iyi olabilecek değerleri deneme eğilimi sonucu makul sayılabilecek süreler içerisinde optimal sonuçlar üretebilme kabiliyetine sahiptir.

GA, popülasyon temelli bir optimizasyon algoritma olup doğal evrime benzetilir. Bu nedenle Üreme (Reproduction), Çaprazlama (Crossover), Mutasyon (Mutation) gibi doğal evrimde kullanılan operatörleri içerir. Popülasyonu oluşturan aday çözümlerin

algoritmadaki karşılığı kromozomlardır. Bu kromozomlar çeşitli evrim işlemleri sayesinde daha iyi sonuçları temsil eden çözüm adaylarına dönüşürler. Bu işlem kabul edilebilir bir uygunluk değerine ulaşıncaya kadar veya önceden belirlenen bir işlem süresi veya nesil sayısı gibi kriterler karşılanıncaya kadar sürdürülür.

GA'da kromozomlar (aday çözümler) temsil ettiği çözüme ait ayrık veya sürekli değerlere sahip değişkenleri tutan katarlardır. Uygunluk fonksiyonu ise kromozomların kalitesini ölçen amaç fonksiyonudur. Genetik Algoritmanın temel adımları Şekil 2.10'da verilmiştir.

Adım 1. Rastgele değerlerden başlangıç popülasyonunu üret

Adım 2. Kromozomların uygunluk değerlerini hesapla

Adım 3. Şartlar sağlanıncaya kadar devam et (Uygunluk değeri, işlem süresi vb)

- **(i)** Popülasyonda kötü uygunluk değerine sahip kromozomları belirle
- **(ii)** Çocuk kromozomlar için ebeveyn kromozomları tayin et
- **(iii)** Çaprazlama işlemi ile ebeveyn kromozomlardan çocuk kromozomlar üret
- **(iv)** Çocuk kromozomları mutasyona uğrat.
- **(v)** Popülasyona yeni dahil olan kromozomların uygunluk değerlerini hesapla

Şekil 2.10. Genetik algoritmanın temel adımları

Başlangıç popülasyonu rastgele aday çözümlerden oluşur. Popülasyonun büyüklüğü için belirli bir kriter yoktur. Ancak popülasyondaki birey sayısının çok fazla olmasının problemin çözüm zamanına veya elde edilecek çözümün iyiliğine etkisi olmadığı değerlendirilmektedir. Bu yüzden genellikle popülasyon 20-50 arasında alınır [68].

Üreme, uygunluk (fitness) değerlerine bakılarak rastlantısal yöntemlerle seçilen bireylerden yeni bir popülasyon oluşturma işlemidir. Bu işlem, ilerleyen jenerasyonlarda daha yüksek uygunluk değerlerine sahip bireylerin oluşmasına neden olur. Bu işlem en uygun olanın hayatta kalması olarak tanımlanır.

Çaprazlama, çoğunlukla rastgele olarak seçilen iki bireyin kromozomları çaprazlanarak gerçeklenir. Bu işlemde, bireylerin kromozomunu oluşturan dizilerin değişik kısımlar yer değiştirerek yeni döl üretimi sağlanır. Bu döl popülasyonunda daha az uygunluk değerine

sahip “zayıf” bireylerin yerine konabilir. Çaprazlama, genetik algoritmada en önemli operatördür ve jenerasyonda yeni çözümlerinin üretiminden sorumludur.

Mutasyon, bireyin kromozomunu oluşturan dizideki tek bir elemanın değerinin rastgele olarak değişmesidir. Mutasyon, çözümün alt optimal noktalara takılmasını önleyen ve çok düşük olasılık değeri ile uygulanan operatördür.

Genel olarak bir GA çözüm bilgisinin hiç olmadığı veya çok az olduğu bir durumla aramaya başlar. Çözüm çevreden gelen etkileşime ve genetik operatörlere bağlıdır. GA, aramaya paralel bir şekilde, birbirinden bağımsız noktalardan başlar, bu nedenle alt optimal çözümlere takılma olasılığı azdır. Bu nedenle GA, karmaşık arama problemleri (birden çok alt çözüm kümesi olan) için en iyi optimizasyon tekniklerinden biri olarak bilinir. GA’yı diğer Evrim Algoritma (EA) türlerinden farklı kılan özellikleri;

- Eşeyli üreme yöntemini,
- Mutasyon ve çaprazlama operatörlerini,
- Stokastik veya determistik seçim yöntemlerini,
- Problemin çözümü için problemin kendisi yerine kodlanmış bir dizisini kullanması gösterilir.

Bu şekilde GA diğer EA türlerinden daha esnek bir yapı sağlar. Bir evrim işlemi, potansiyel çözüm uzayında, popülasyonu oluşturan kromozomlar içinde en uygun kromozomu arama işlemidir. Böyle bir arama iki zıt amacı dengelemeyi gerektirir. Bu amaçlar; en iyi çözümlerin aranması ve arama uzayının genişletilmesidir. Genetik algoritma ile geleneksel optimizasyon teknikleri (özellikle nümerik metotlar) arasında da çok önemli farklılık vardır. Bu farklılıklar aşağıdaki gibi özetlenebilir:

- GA, optimize edilecek olan parametrelerin kendileri ile değil kodlamış dizileri üzerinde çalışır. Pek çok durumda ikili (binary) kodlama kullanılır. Fakat genetik algoritmalar için bu bir gereklilik değildir. Gerçel sayı kodlama, ağaç yapılı kodlama (tree coding) gibi farklı kodlama sistemleri de kullanılabilir.
- GA, bir popülasyon içinde arama yapar. Bu popülasyon, problemin bütün olası çözümlerini temsil eden uzayı oluşturur. Başlangıç popülasyonu genellikle rastgele üretilen bireyleri içerir.

- GA, problemin çözümünü belirlenen çözüm uzayında aramak için bir uygunluk fonksiyonu (fitness function) kullanır. Bu uygunluk fonksiyonu klasik optimizasyon tekniklerinde kullanılan amaç fonksiyonuna benzetilebilir.
- GA, sonuca ulaşmak için rastlantısal yöntemler kullanır.

Genetik algoritmaları cazip kılan avantajlarından bazıları şunlardır:

Öğrenme: Genetik algoritma, global arama tekniklerinde yaygınca kullanılır ve en iyisi olarak bilinir. Mevcut performans ölçütlerini kullanarak verilen bir arama uzayında arama uzayını genişletme ve en iyiyi arama özelliklerini kullanma yeteneğine sahiptir. Bu özelliklerini çaprazlama, mutasyon ve üretim gibi genetik operatörlerle kullanarak öğrenme yeteneğine sahiptir.

Genetik kod yapısı: Genetik algoritma doğrudan parametrelerle değil, kodlanmış parametre dizisiyle çalışır. Bu, kullanıcıya problemleri bir değişken optimizasyon problemi gibi çözmesine imkan verir.

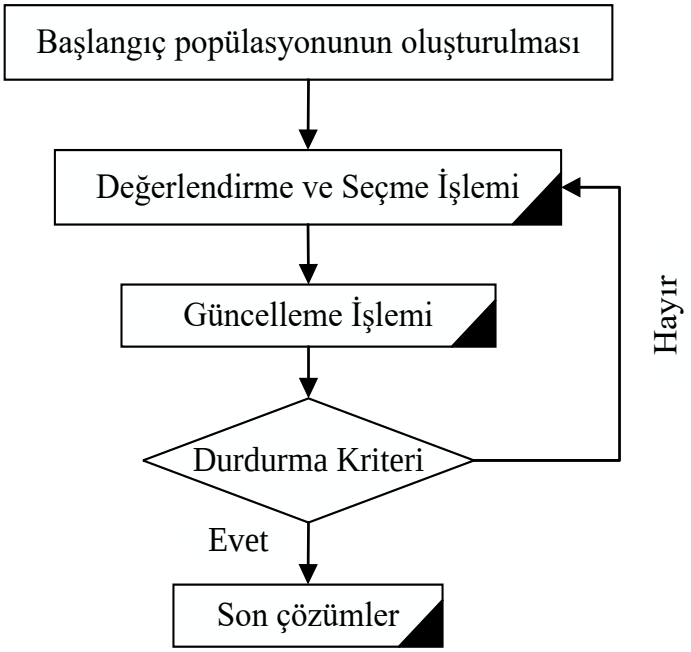
Çözümlerin optimalliği: Pek çok gerçek hayat problemlerinin Çok Modellilik (Multimodal) ve doğrusal olmama gibi özellikleri vardır. Geleneksel arama teknikleri böyle arama uzaylarında yetersiz kalır. Genetik algoritma ise böyle karmaşık arama uzaylarında optimale yakın çözümler bulma yeteneğine sahiptir. Genetik algoritma, mühendislik, bilim, ekonomi çok değişik alanlardaki problemler için gürbüz (Robust) bir optimizasyon aracı olarak son yıllarda büyük bir önem kazanmıştır.

Genetik algoritmanın uygulama alanlarından bazıları; haberleşme şebekeleri tasarımı, elektronik devre dizaynı, gaz boruları şebekeleri optmizasyonu, görüntü ve ses tanıma, veri tabanı sorgulama optimizasyonu, uçak tasarımı, fiziksel sistemlerin kontrolü, gezgin satıcı problemlerinin çözümü, ulaşım problemleri, optimal kontrol problemleridir [69].

2.2.2.2.2. Parçacık Sürü Optimizasyon Algoritması

Parçacık sürü optimizasyon algoritması (PSO) balık ve kuş sürülerinin yiyecek kaynaklarına doğru yaptıkları davranışlardan esinlenilerek 1995'de Kennedy ve Eberhart tarafından geliştirilmiştir [70]. Temel olarak sürü zekasına dayanan bir algoritmadır. Sürü halinde hareket eden hayvanların yiyecek ve güvenlik gibi durumlarda, çoğu zaman

rasgele sergiledikleri hareketlerin, amaçlarına daha kolay ulaşmalarını sağladığı görülmüştür. PSO algoritması özellikle sürekli parametre optimizasyonunda ve çok boyutlu arama alanlarında yaygın olarak kullanılmaktadır. PSO bireyler arasındaki sosyal bilgi paylaşımını esas alır. Arama işlemi genetik algoritmalarda olduğu jenerasyon sayısınca yapılır. Her bireye parçacık denir ve parçacıklardan oluşan popülasyona da sürü (swarm) denir. Parçacık sürüsünü oluşturan parçacıklar evrimsel algoritmadaki bireylere karşılık gelir. PSO algoritmasının temel blok şeması Şekil 2.11'de verilmiştir.



Şekil 2.11. Parçacık sürü optimizasyon algoritmasının temel akış diyagramı

İlk olarak, parçacık popülasyonu rasgele olarak üretilir. Her parçacık potansiyel çözümü temsil eder. Parçacıkların pozisyonu ise pozisyon vektörü ile gösterilir. Bir parçacık sürüsü rasgele belirlenen başlangıç pozisyonlarından problem uzayına doğru hareket eder ve hareketleri her bir parçacığın vektörel hareket hızı ile gösterilir. Bu hareketleri sırasında her parçacık kendi en iyi pozisyon bilgisini tutar, bunun yanında sürünün en iyi pozisyon bilgisi en iyi uygunluk değerine karşılık gelir ve vektör değeriyle gösterilir ve saklanır. PSO modelinde parçacıklar arasındaki sosyal bilgi paylaşımı geliştirme amaç edinilmiştir. Parçacıklar aday çözümün uygunluğunu iteratif olarak hesaplar ve en iyi kaynağın yerini hafızasına alır. x ve v parçacığın çok boyutlu uzayda konum ve hız

vektörleridir. Her bir aday çözüm (parçacık) bir sonraki pozisyonunu Eşitlik 2.22'de verilen hız vektörünün değerine göre ayarlar.

$$\vec{x}(t+1)=\vec{x}(t)+\vec{v}(t+1) \tag{2.22}$$

Hız değerinin güncellenmesinin hesaplanmasında, kendi tecrübesini (yerel en iyi değeri) ve sürü tecrübesini (küresel en iyi değeri) kullanır. Bu şekilde hız değişimi parçacığın, hem kendi en iyi pozisyonunun hem de küresel en iyi pozisyonun çevresinde arama yapmasını sağlar. Güncellenmiş bu hız vektörüne göre her bir parçacık kendi pozisyonunu Eşitlik 2.23'e göre değiştirir.

$$\vec{v}(t+1)=w\vec{v}(t)+\phi_1 rand(0,1)\left(\vec{p}(t)-\vec{x}(t)\right)+\phi_2 rand(0,1)\left(\vec{g}(t)-\vec{x}(t)\right) \tag{2.23}$$

Burada w yeni hız değerinin hesaplanmasında bir önceki hız değerinin $\vec{v}(t)$ etkisini belirleyen ağırlıktır. $\emptyset_1$ ve $\emptyset_1$ sırasıyla $\vec{p}(t)$ ve $\vec{g}(t)$'nin önem derecelerini belirlerler. $(\vec{p}(t))$ yerel en iyi ve $(\vec{g}(t))$ küresel en iyi çözümlerdir [64].

Başlangıçta parametrelerin, uygunluk (amaç) fonksiyonunun ve durdurma kriterinin tanımlanması gerekir. Daha sonra, parametrelerin temsilinin, başlangıç popülasyonunun oluşturulmasının, uygunluk değerlendirilmesinin, seçme ve güncelleme işlemlerinin yapılması gerekir. Algoritmanın temel adımları Şekil 2.12'de verilmiştir.

Adım 1. Popülasyonun başlatılması
Adım 2. REPEAT
Adım 3. Parçacıkların uygunluk değerlerinin hesaplanması
Adım 4. Sürüdeki en iyi parçacıkların güncellenmesi
Adım 5. En iyi parçacıkların belirlenmesi
Adım 6. Parçacıkların hızlarının hesaplanması
Adım 7. Parçacıkların pozisyonlarının güncellenmesi
Adım 8. UNTIL Durdurma kriteri sağlanıyor **ise** dur,
Aksi takdirde Adım 2'ye git.

Şekil 2.12. Parçacık sürü optimizasyon algoritmasının temel adımları

3. BÖLÜM

BENZETİLMİŞ TAVLAMA ALGORİTMASI İLE ANFIS AĞININ EĞİTİLMESİ

3.1. Benzetilmiş Tavlama Algoritması

Benzetilmiş Tavlama (Simulated Annealing, SA) ilk olarak Kirkpatrick, Gelatt ve Vecchi tarafından önerilmiş, olasılık tabanlı sezgisel bir algoritmadır [71]. SA algoritması katıların fiziksel tavlama işlemi ile kombinatoryal optimizasyon problemlerinin çözümü arasındaki benzerlik üzerine dayalıdır. SA algoritması, pek çok değişkene sahip fonksiyonların en büyük veya en küçük değerlerinin bulunması ve özellikle pek çok yerel minimumlara sahip doğrusal olmayan fonksiyonların en küçük değerlerinin bulunması için tasarlanmıştır. Bu algoritma, katı cisimlerin soğurken atomlarının mükemmel şekilde dizilip potansiyel enerjiyi minimize etmesini örnek aldığından ve özellikle metallerin tavlama işlemini andırdığından bu ismi almıştır. Diğer olasılıksal yaklaşımlar (genetik algoritmalar, tabu arama vb.) gibi en iyi çözümün en kısa zamanda üretimini sağlar. Bu sebeple, özellikle matematiksel modellerle gösterilemeyen kombinasyonel problemlerin eniyileme uygulamalarında tercih edilir. SA algoritması; elektronik devre tasarımı, görüntü işleme, yol bulma problemleri, seyahat problemleri, malzeme fiziği benzetimi, kesme ve paketleme problemleri, akış çizelgeleme ve iş çizelgeleme problemlerinin çözümlerinde başarılı sonuçlar vermiştir.

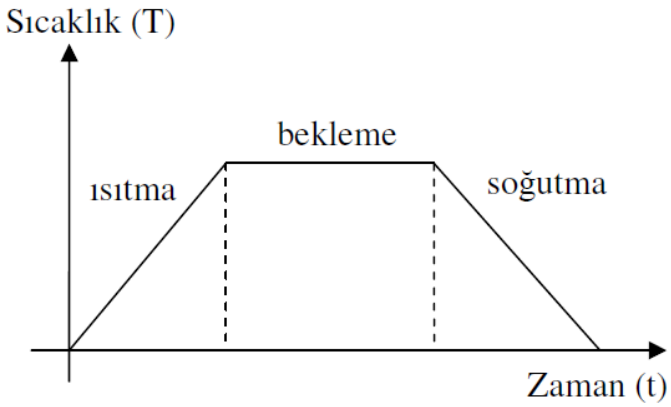
Algoritmanın çalışma mantığını anlamak için, en alçak noktası aranılan bol delikli bir golf sahasını örnek olarak göstermek mümkündür. Golf sahasının en alçak noktasını bulmak için sahaya konulan bir top ile sahanın eğimi yönünde ilerlemek gibi basit bir yöntem kullanılırsa, topun deliklerden birine takılma olasılığı yüksektir. Bu sorunu çözmek için, arazi olduğu gibi sallanırsa, top arada bir deliklere girse de arazi sürekli sallandığı için top delikten tekrar çıkmaktadır. Zamanla sallama hızı ağır ağır azaltılarak,

sallama tamamen durduğunda topun sahanın en alçak noktasında (global minimum) yada yakın bir yerlerde olduğu farz edilir. Gerçek dünyadaki katı cisimlerde de durum bu örnektekine benzerdir. Golf sahası örneğindeki sallama hareketi cisimlerin sıcaklığına karşılık gelir. Bir gazı soğuturken atomlar bir süre sonra nasıl ki periyodik aralıklarla dizilip potansiyel enerjiyi minimize ediyorlar ise (kristalleşme) SA algoritması da aynı yöntemi kullanarak enerjiyi değil tanımlanan bir amaç fonksiyonunu minimize etmeye çalışmaktadır [72].

SA, yerel arama yöntemlerinin yerel bir minimuma ulaştıktan sonra global minimum için daha fazla arama yapmamasından kaynaklanan eksikliğini gidermeye çalışan bir yöntemdir. Ayrıca, bir istisnası dışında yerel arama yöntemindeki aynı temel adımları kullanır. Soğutma işlemi bu algoritmada daha iyi sonuçların bulunmasını sağlayacak yeni komşu çözümlerin üretilmesini sağlayan üstel (exponential) bir ifadedir.

3.1.1. Fiziksel Tavlama Süreci

SA algoritması katıların fiziksel tavlama işlemi ve kombinatoryal optimizasyon problemlerinin çözümü arasındaki benzerlik üzerine dayalıdır. Isıl işlem sayesinde malzemenin belirli bir sıcaklığa kadar ısıtılması (ısıtma) bu sıcaklığın uygun bir süre tutulması (bekleme) ve belirli bir programa uygun olarak sıcaklığın düşürülmesi (soğutma) ile üç kademede özellik değişimleri sağlanır. Şekil 3.1'de ısıl işlem kademeleri gösterilmiştir.



Şekil 3.1. Isıl işlem kademeleri

Tavlama, katının ergime noktasına kadar ısıtılması ve devamında mükemmel kafes yapılı durumda kristalize olana kadar yavaşça soğutulması işlemidir. Malzemelerin atomları yüksek sıcaklıklarda yüksek enerji seviyelerindedir ve düzgün yerleşimler için daha fazla hareket serbestliğine sahiptirler. Düzgün yapılı bir kristal sağlandığında sistem minimum enerjiye sahiptir. Sıcaklık azaltıldıkça atomik enerji düşer. Eğer soğutma işlemi çok hızlı gerçekleşirse kristal yapıda bozukluklar ve düzensizlikler ortaya çıkacaktır. Bu nedenle soğutma işleminin dikkatle yapılması gerekmektedir [73]. Fiziksel tavlama işlemi Monte Carlo tekniği üzerine dayalı olarak Metropolis ve arkadaşları tarafından modellenmiştir [74]. Verilen bir T sıcaklığında sistem enerjilerinin olasılık dağılımı aşağıda verilen termodinamik kanunu ile belirlenir:

$$P(E) = e^{-E/(kT)} \tag{3.1}$$

Burada, E sistem enerjisi, k ise Boltzmann sabitidir.

Küçük bir karışıklıkla sistem durumunda değişiklik yaratılması halinde Metropolis algoritmasına göre sistemin yeni enerjisi hesaplanır. Eğer enerji azalmış ise sistem bu yeni duruma geçer. Eğer enerji artmış ise yeni durumun kabul edilip edilmemesine, Eşitlik 3.1'de verilen olasılık formülü kullanılarak şu şekilde karar verilir: Düzgün (uniform) dağılımdan (0,1) aralığında rasgele bir δ sayısı üretilir ve Eşitlik 3.2'de verilen şart sağlanırsa durum, yeni durum olarak kabul edilir. Aksi takdirde, mevcut durum değiştirilmez.

$$\delta \leq e^{\Delta E/T} \tag{3.2}$$

Burada, ΔE, iki durumun enerji seviyeleri arasındaki farktır ve bu kabul kriteri *Metropolis kriteri* olarak bilinir. Eşitlik 3.1'e göre, yüksek sıcaklıklarda tüm enerji durumları için $P(E)$, 1'e yakınsar. Düşük sıcaklıklarda bile sistemin yüksek enerji seviyesine sahip olması küçük bir olasılıkla görülebilir. Bu nedenle enerjilerin istatistiksel dağılımı sistemin bir bölgesel enerji minimumundan çıkmasına izin verir.

3.1.2. SA Algoritmasının Çalışması

Kombinatoryal optimizasyon problemi ile tavlama işlemi arasındaki benzerlik de, katının durumları optimizasyon probleminin muhtemel çözümlerini temsil eder ve bu durumların enerjileri, çözümler için hesaplanan amaç fonksiyon değerlerine karşılık gelir. Minimum enerji durumu problem için optimal çözümü ifade eder, hızlı soğutma işlemi ise yerel optimum olarak görülebilir [75]. Benzetilmiş tavlama ve optimizasyon problemleri arasındaki ilişki Tablo 3.1'de gösterilmektedir.

Tablo 3.1. Tavlama süreci ve optimizasyon problemi arasındaki ilişki

Tavlama Süreci	**Optimizasyon Problemi**
Sıcaklık	İterasyon Sayısını Kontrol Eden Kontrol Parametresi
Fiziksel Sistemin Durumu	Problemin Çözümü
İçinde Bulunan Durumun Enerjisi	Amaç Fonksiyonunun Değeri
Minimum Enerji	Optimal Çözüm

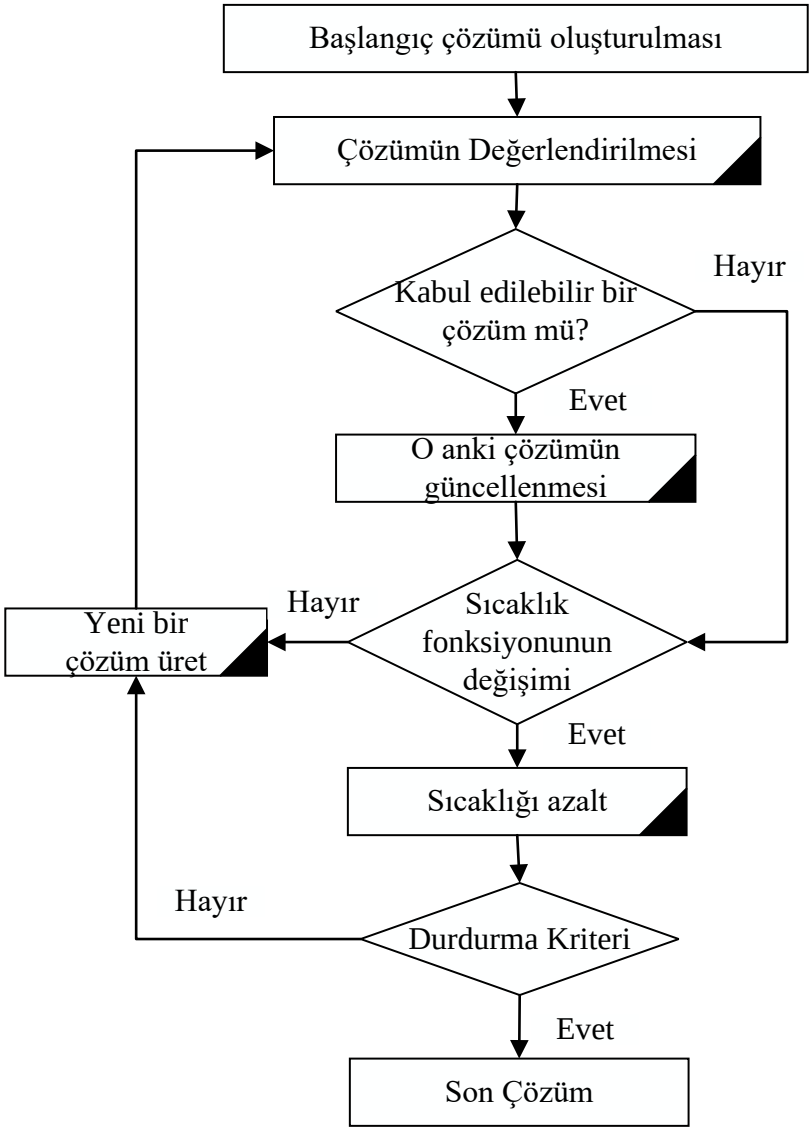
SA iteratif tabanlı bir algoritmadır, yani algoritma çözüm uzayında sayıların vektörü formunda ifade edilen bir çözümü sürekli olarak geliştirmeye çalışır. Standart bir SA algoritmasının temel adımları ise Şekil 3.2'de verilmiştir. SA algoritmasının bir probleme uygulanması aşamasında karar verilmesi gereken önemli kavramları iki gurupta toplamak mümkündür [76]:

a) Probleme özgü seçimler: Problem, muhtemel tüm çözümleri içerecek şekilde formüle edilmelidir. Çözümlerin temsili yapılmalı, minimize edilecek bir amaç fonksiyonu ile komşuluk üretim mekanizması tanımlanmalı ve bir başlangıç çözümü üretilmelidir.
b) Soğutma planına ait seçimler: Sıcaklık parametresi T' nin başlangıç değerine, soğutma oranı ve güncelleme kuralına, her bir sıcaklıkta icra edilecek iterasyon sayısına karar verilmelidir.

Adım 1. Bir başlangıç çözümü üret, *S*.
Adım 2. Bir *S'∈N(S)* çözümü seç ve amaç fonksiyon değerlerindeki farkı hesapla.
$\Delta = C(S) - C(S')$
Adım 3. Eğer,
(i) *S'*, *S* den daha iyi ($\Delta > 0$), **yada**
(ii) $\delta \leq e^{\Delta E/T}$ **ise**,
S' nü yeni çözüm olarak tayin et (*S*←*S*')
Aksi takdirde, mevcut çözümü tut.
Adım 4. Sıcaklığı güncelle.
Adım 5. Eğer Durdurma kriteri sağlanıyor **ise** dur,
Aksi takdirde Adım 2' ye git.

Şekil 3.2. SA algoritmasının temel adımları [77]

SA algoritmasının performansı büyük oranda seçilen soğutma planına bağlıdır. Literatürde çeşitli soğutma planları önerilmiştir. Önerilen en eski plan Kirckpatrick ve arkadaşlarının fiziksel tavlama ile olan benzerliğe dayanarak ileri sürdükleri plandır. Bu tavlama planına göre, maddenin sıvı safhaya ulaştığında tüm parçacıkların rasgele düzenlenmesini taklit etmek için, *T* sıcaklık parametresinin başlangıç değeri denenen tüm hareketler kabul edilecek kadar yüksek seçilmiştir. Sıcaklık parametresinin değerini azaltmak için ise, $T(t+1) = r.T(t)$ bağıntısı kullanılmıştır. Burada r değeri 1'den küçük fakat 1'e yakın bir sabittir ve genellikle değeri 0,8 ile 0,99 arasında seçilir. Bu sıcaklık fonksiyonu ile sıcaklık parametresinin değeri sıfıra yaklaştıkça daha da yavaş azalmaktadır. Sıcaklık parametresinin her değerinde gerçekleştirilecek iterasyon sayısı sabit bir üst sınıra göre belirlenerek problemin fiziksel tavlamadaki ısıl dengeye karşılık gelen bir denge durumuna ulaşması amaçlanmaktadır. Bu tavlama planında, sıcaklık parametresinin her değerinde elde edilen çözüm belirli sayıda ardışık sıcaklık değişimleri boyunca aynı kalırsa SA algoritması durdurulmaktadır. Buna göre elde edilen son durum fiziksel tavlamadaki *donma durumuna (frozen state)* karşılık gelmektedir. Kombinasyonel eniyileme problemlerinin çözümündeki bu tavlama sürecini gösteren benzetilmiş tavlama algoritmasının akış diyagramı Şekil 3.3'de gösterilmiştir.



Şekil 3.3. SA algoritmasının akış diyagramı

3.1.3. SA Algoritmasına ait Parametreler

Sıcaklık parametresinin azalan değerlerinin sonlu bir sırası için, sonlu uzunluklu homojen Markov zincirleri üretilerek oluşturulan bir SA algoritmasının uygulaması, sonlu zaman alacaktır. Yakınsamanın sağlanabilmesi için bu algoritmada kullanılan parametreler kümesi uygun bir şekilde belirlenmek zorundadır. SA algoritmasında kullanılan parametrelerin ve değerlerinin belirlenmesi tavlama veya soğutma planı olarak tanımlanmaktadır. Bu parametrelerin her biri tavlama benzetiminin çözmeye çalıştığı problemin yapısına göre değişiklik göstermekte hatta bu parametrelere yenileri eklenebilmektedir. Tavlama planı ile aşağıdaki parametreler belirlenmektedir [78].

1. T sıcaklık parametresinin başlangıç değeri,
2. Sıcaklığın hangi yöntemle azaltılacağını belirlemek için kullanılan $T(t)$ sıcaklık fonksiyonu,
3. Soğutma oranı ve güncelleme kuralının tanımlanması,
4. Her sıcaklıkta gerçekleştirilmesi gereken M iterasyon sayısı,
5. Algoritmayı sonlandırmak için gerekli *durdurma kriteri*.

Başlangıç sıcaklığı: Başlangıç sıcaklığı tavlama benzetimiyle ulaşılmaya çalışılan sabitlerin aranacağı genişliği ifade etmektedir. Bu ısı ne kadar yüksek olursa taranacak genişlik o kadar yüksek olacak, dolayısıyla arama yoğunluğu düşecektir. Tavlama benzetiminde de başlangıç sıcaklığının belirlenmesi oldukça önemlidir. Başlangıç sıcaklığının gereğinden fazla bir değere çıkartılması gereksiz işlem sürelerine neden olurken; düşük sıcaklıklarda başlanması, algoritmanın yerel minimumlarda takılma riskini arttırmaktadır. Bu açıdan eğer başlangıç sıcaklığının değeri hakkında kesin bir sınır belirlenemiyorsa, mümkün olduğunca yüksek değerlerin atanması daha uygun olacaktır.

Soğutma oranı: Soğutma oranı α ile temsil edilmektedir. 0 ile 1 arasında herhangi bir değer verilebilen α, mevcut sıcaklıkla çarpılarak azalmasını sağlar. Soğutma oranı 0'a yaklaştıkça sistem daha hızlı soğumakta, 1'e yaklaştıkça sistemin soğuması yavaşlamaktadır. Fiziksel tavlama da sıcaklığın hızlı veya yavaş soğutulmasının kendileri göre avantajları ve dezavantajları bulunmaktadır. Örneğin; yavaş soğutulan metal daha esnek olurken, soğutma hızı arttıkça bu esneklik kendisini sertliğe ve dayanıklılığa bırakmaktadır. Metalin fonksiyonuna göre en doğru soğutma oranları seçilmektedir. Benzetilmiş tavlamada ise soğutma oranı ile hız ve kesinlik arasında denge kurulmaya çalışılır. Bu oranın yüksek olması durumunda çözüme ulaşma hızı azalacak fakat kesin çözüme yaklaşacak veya kesin çözümü bulacaktır. Oranın düşük olması durumunda ise çok kısa zamanda sonuca ulaşmakla beraber çözümün doğruluk payında düşme görülecektir.

Algoritmanın kullanım alanına göre hız ve kesinlik arasında dengenin sağlanması gerekmektedir. Sonucun doğruluğunun sonuca ulaşma zamanına göre çok daha önemli olduğu durumlarda soğutma oranı 1'e yakın seçilerek sistemin yavaş soğuması

sağlanmalıdır. Genellikle dönemlik tahminlerde kullanılabilen böyle bir yaklaşım ile kesin (veya kesine yakın) sonuçlar elde edilmektedir. Hem doğruluğun hem de hızın çok önemli olduğu askeri araçlar, gemiler vs. karmaşık sistemlerde ise işlem gücü arttırılarak ve işlemler paralel çalıştırılarak kabul edilebilir zaman aralığında kabul edilebilir hata oranlarına ulaşılabilmektedir.

İterasyon sayısı: İterasyon sayısı, sıcaklığın düşürülmeden önce, her bir sıcaklık seviyesinde gerçekleştirilecek döngü miktarıdır. Döngü miktarı belirlenen limite ulaştığında sıcaklık α değeriyle çarpılarak düşürülecektir.

İterasyon sayısı ile soğutma oranı ters etkiye sahiptir. Döngü sayısının artması ve soğutma hızının yükselmesi ile soğutma çizgisi merdiven şeklini alacaktır. Merdiven basamaklarının derinliği soğutma oranının düşük olmasıyla orantılı olarak artacaktır. Bu durum ise arada kalan bazı noktaların aranma olasılıklarını sıfıra indirmektedir. Bu "imkânsızlığın" giderilmesi için soğutma oranını yükseltilerek soğuma hızı yavaşlatılmalı ve döngü sayısı azaltılmalıdır. Böylece her bir sıcaklık katmanı arasındaki mesafeler olabildiğince yaklaşacak ve tüm uzayın aranma olasılığı artacaktır.

Durdurma Kriteri: Durdurma kriteri, döngünün belli bir koşulu sağladığında durdurulmasını ifade etmektedir. Fiziksel tavlamada işlem belli bir sıcaklığa ulaştığında otomatik olarak durdurulmaktadır.

Belli bir sıcaklığa ulaşma durumunda döngünün durdurulması benzetilmiş tavlamada da en sık kullanılan yöntemdir. Genellikle sıcaklık 0 ^{0}C'ye indiğinde, yani parçaların hareketi durduğunda, döngü durdurulmaktadır. Bu kuralla beraber belli bir döngü sonucunda herhangi bir iyileştirme olmuyorsa da döngü durdurulabilir. Durdurma kriteri için kullanılabilecek bir diğer yöntem ise zaman kısıtlı döngülerdir. Döngüler zaman yönünden kısıtlanarak belli bir zaman aralığı içinde buldukları en iyi sonuca odaklanabilir.

3.1.3.1. Teorik Soğutma Stratejileri

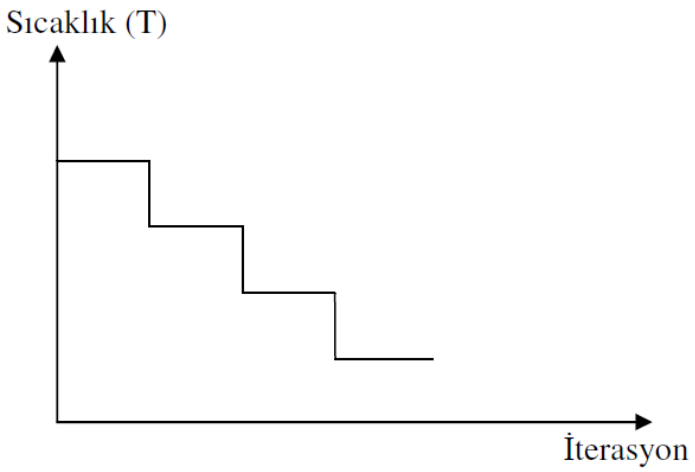
Teorik soğutma stratejileri benzetilmiş tavlama algoritmasının 1.0'a eşit olasılıkla global optimal çözümlere yakınsamasını sağlar. Teorik soğutma stratejilerinde iki tip formülasyon ortaya çıkmıştır.

Homojen algoritma: Bu algoritmada sıcaklık, sıfıra erişinceye kadar sonsuz kademeler dizisi seklinde azaltılır. Verilen bir sıcaklık değerinde gerçekleştirilen iterasyon sayısı belirlenir.

Homojen olmayan algoritma: Bu algoritmada da yine sıcaklık sıfıra erişinceye kadar sonsuz kademeler dizisi seklinde azaltılır. Bununla birlikte her bir kademe sadece tek bir iterasyon hareketini temsil etmektedir.

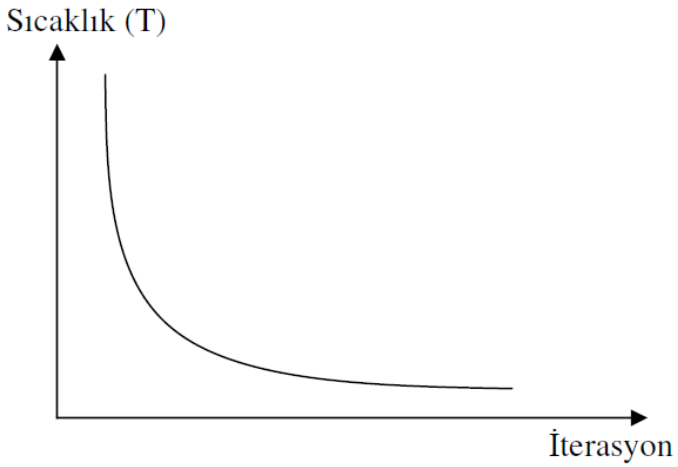
3.1.3.2. Basit Soğutma Stratejileri

Pratikte basit soğutma stratejileri teorik soğutma stratejilerine göre daha başarılı neticeler üretmektedir. Bunun sebebi teorik olanların yakınsamayı ispat etmeye çalışması fakat bununda yakınsamayı yavaşlatması olabilir. Basamak tip sıcaklık düşürme stratejisi: Bu kategori, belirli bazı kurallara göre sıcaklığı bir kademe düşürerek, mevcut sıcaklık değerinde sabit bir Markov zincir uzunluğundaki (L) uygulamaları içerir. Şekil 3.4'de basamak tipte sıcaklık düşürme stratejisi gösterilmiştir.



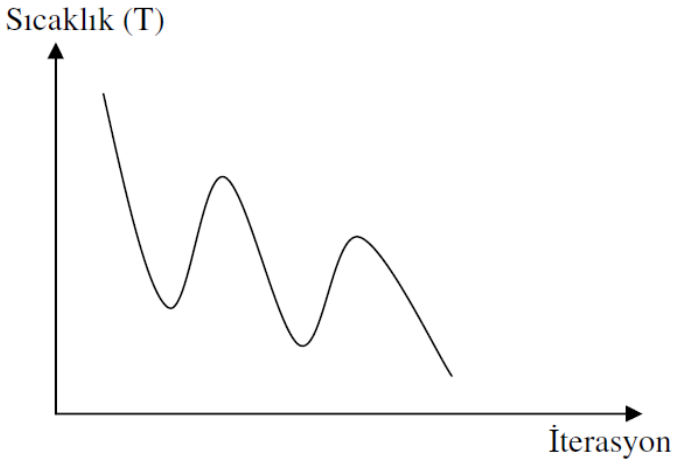
Şekil 3.4. Basamak tipte sıcaklık düşürülmesi

Sürekli sıcaklık düşürme stratejileri: Bu kategori her sıcaklık değerinde sadece bir hareketin yapıldığı soğutma tarifelerini içerir. Sıcaklık T, her iterasyondan sonra düşürülerek sürekli azaltılır. Şekil 3.5'de sürekli sıcaklık düşürme stratejisi gösterilmiştir.



Şekil 3.5. Sürekli olarak sıcaklığın düşürülmesi

Monotonik olmayan sıcaklık düşürme stratejileri: Monoton olmayan sıcaklık düşürme stratejisinde, her iç değişimden sonra sıcaklık yavaş yavaş azaltılmaktadır. Komsu çözümlerde herhangi bir değişim gözlenmez ise sıcaklık belirli bir değere kadar arttırılır. Şekil 3.6'da monoton olmayan sıcaklık düşürme stratejisi gösterilmiştir.



Şekil 3.6. Monoton olmayan şekilde sıcaklığın düşürülmesi

Monoton olmayan sıcaklık düşürme stratejisi yukarıda bahsettiğimiz diğer soğutma stratejilerinin dezavantajlarını yok etmek içim geliştirilmiştir. Önceki soğutma tarifelerinde, komşuluk genellikle rastgele alternatif çözümler üretmek için araştırılmakta ve sıcaklığın düşük değerlerinde, mevcut çözümden daha kötü çözümlerin kabul edilebilirlik olasılığı çok azalmaktadır.

Son yıllara kadar benzetilmiş tavlama algoritması, çoğunlukla ayrık formdaki optimizasyon problemlerine uygulanmakta ve başarılı sonuçlar alınmakta iken, son zamanlarda yapı dinamiği, kontrol sistemleri, elman ağı gibi faklı mühendislik sahalarından değişik nümerik optimizasyon problemlerine de uygulanmıştır [77,79,80].

3.2. Trombofili Hastalığı ile Genetik Bozukluklar Arasındaki İlişkinin Tespit Edilmesi

3.2.1. Tıp Bilişimi

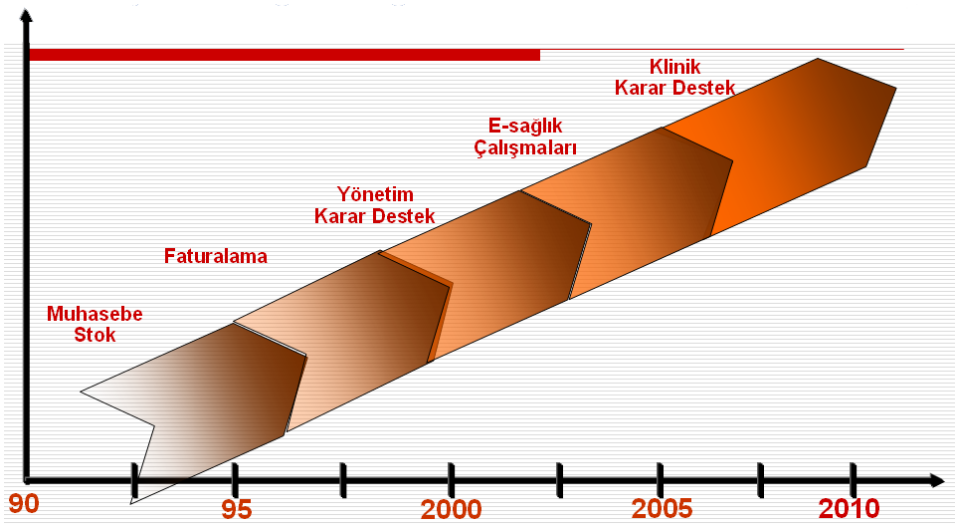
Günümüzde, sosyal, ekonomik, bilimsel ve teknolojik gelişmeler, tüm insanlığı, giderek artan bir hızla, küresel bir bilgi toplumu olmaya doğru götürmektedir. Büyük bir hızla artan tıp bilgisi ve buna paralel olarak çoğalan ve gelişen ölçü ve görüntüleme yöntemleri, giderek otomatikleşen tıbbi test, analiz ve monitorizasyon cihazları, bireyler ve hastalar için toplanılan tıbbi veri ve bilgileri de büyük bir hızla arttırmaktadır.

Tıp Bilişimi (TB, Medical Informatics), Bilgi teknolojilerini kullanarak yapılan tıbbi çalışma, eğitim, iletişim, veri ve bilgi toplama, veri ve bilgi işleme, bilgi yönetme, tıbbi karar verme ve bilimsel çözümleme yöntemlerini içeren bir bilim dalıdır. TB, tıp, bilişim teknolojileri ve farklı disiplinlerinin kesiştiği bir yerde bulunmaktadır [81]. TB, tıp alanındaki bilgilerin (data, information, knowledge) etkili ve etkin kullanımı, bilgilerin yaygınlaştırılması, analizi, yeni yapılanmalara imkan sağlayacak şekilde yönetilmesi için değişik bilim dalları ile etkileşimli bir şekilde günümüz bilgisayar ve iletişim teknolojisinin en üst düzeyde kullanılmasını amaçlamaktadır. Bilgi teknolojileri ve sağlık bilimlerinin kesiştiği noktada bulunan tıbbi bilişim, şimdiye kadar çeşitli tıbbi araştırma ve geliştirme konularında olduğu gibi, kuramsal ve uygulamalı eğitimde de rol almıştır.

TB teknikleri, sağlık merkezlerinin veri toplama, işleme ve değerlendirme yetilerini üstel bir şekilde artırarak küresel ölçekte doğru, ayrıntılı ve güvenilir sonuçlara ulaşılmasını sağlamaktadır. Hemen tamamı gözleme, elde edilen bulguların sınıflandırılıp değerlendirilmesine ve istatistiki çalışmalara dayalı olan tıp biliminde, verilerin toplanabilmesi, süratle işlenip tasnif edilebilmesi, standardizasyonu ve daha sonra kolayca ulaşılabilmesi önem taşımaktadır. TB uygulamaları bu açıdan global veri

tabanları ve bilgilerin tasnif edilip süratle erişilebilmesinde sağladığı kolaylıklarla modern tıp biliminin vazgeçilemez bir alt bilim dalı olmuştur.

Küresel ölçekte bakıldığında, tıp bilişiminin genel olarak idari süreçlere yönelik bilgi kayıtları ve hastane yönetim sistemleri gibi uygulamalarının yanı sıra, daha çok bilimsel ve akademik faaliyetlere yönelik görüntü işleme ve analizi, sinyal analizi, karar destek sistemleri, tıpta yapay zekâ ve tıpta veri madenciliği gibi birçok uygulamaları vardır. Şekil 3.7' de gösterildiği gibi Türkiye' de tıp bilişimi araçlarının kullanımı 90' lı yıllarda muhasebe/stok uygulamaları ile başlamış, 2000'li yıllarda e-sağlık uygulamaları ile yeni bir boyut kazanmıştır. Dünyadaki gelişmelere bakıldığında ise bundan sonraki süreçte eğilimin klinik karar destek sistemlerine yönlenmesi gerektiği öngörülmektedir. Dolayısıyla ülkemizde de bu alanda yapılacak çalışmalara ihtiyaç duyulmaktadır. Klinik karar destek sistemleri, önceki süreçlerdeki uygulamalar ile kıyaslandığında, daha zor, akademik ve bilimsel düzeyde uzmanlık gerektiren uygulamalardır. Bu tür uygulamalar, tıp fakültesi ve bilişim mühendisliği alanındaki akademisyenlerin ortaklaşa gerçekleştireceği bilimsel Ar-Ge çalışmalarını zorunlu kılmaktadır. Dünyada klinik karar destek sistemlerine yönelik çok sayıda çalışma bulunmasına rağmen, ülkemizde bu tür akademik çalışmaların sayısı oldukça sınırlıdır.



Şekil 3.7. Türkiye' de tıp bilişimi uygulamalarının gelişimi

3.2.1.1. Tıp Bilişimi Uygulama Alanları

Tıp bilişimi, klinik karar destek sistemlerinin tasarımından, gerekli yazılım ve donanım araçlarının geliştirilmesine kadar geniş bir yelpazede çalışma alanı içerir. Tamamı gözleme, elde edilen bulguların sınıflandırılıp değerlendirilmesine ve istatistikti çalışmalara dayalı olan tıp biliminde, verilerin global bazda homojen olarak toplanabilmesi, süratle işlenip tasnif edilebilmesi, standardizasyonu ve daha sonra kolayca ulaşılabilmesi belirgin önem taşımaktadır. Tıbbı bilişim uygulamaları bu açıdan global veri tabanları ve bilgilerin tasnif edilip süratle erişilebilmesinde sağladığı kolaylıklarla modern tıp biliminin vazgeçilmez bir alt bilim dalı olmuştur. Tıp bilişimi içerisinde yer alan uygulama alanlarından bazıları şunlardır:

- Klinik Bilgi Sistemleri
- Sağlık Bilgi Standartları
- Tıbbi Kayıt Sistemleri
- Hastane Yönetim Sistemleri
- Hemşirelik Bilgi ve Karar Destek Sistemleri
- Tıbbi Bilişim Eğitimi
- Bilgisayar Destekli Tıp Eğitimi
- Görüntü İşleme ve Analizi
- Sinyal Analizi
- Tıbbi Bilgi Ağı (Teletıp, Internet)
- Karar Destek Sistemleri
- Tıbbi Uzman Sistemler
- Tıpta Yapay Zekâ
- Tıpta Veri Madenciliği

Gelişen bilişim teknolojileri ışığında sağlıkta bilişimin konuları artmakta, bu da tıp bilişimini sürekli olarak uygulama alanları genişleyen bir konuma sokmaktadır [82].

3.2.1.2. Tıbbi Karar Destek Sistemleri (Medical Decision Support Systems)

Karar verme süreci incelendiğinde, karar vericinin, karar vermekle yükümlü olduğu probleme ilişkin geçmiş deneyimleri ve mevcut durumu değerlendirerek seçim yapması

gerektiği görülür. Bu süreç sonucu probleme uygulanacak çözüm yolu ile elde edilenler kimi zaman istenmeyen sonuçlar olabilir. Bunun nedeni probleme ilişkin yeterli bilginin olmaması veya karar verme sürecinde yetersizlik olabilir. Karar verme süreci sonunda doğru sonuçlara ulaşabilmek için, probleme ilişkin detaylı bilgiye sahip olunması gerektiği gibi problem üzerinde de deneyime sahip olunması gerekmektedir.

Hekimler tıbbi problemler ile karşılaşan kişilere teşhis koymak ve ilgili tedaviyi uygulamak ile görevli kişilerdir. Hekimlerin karar verme süreci incelendiğinde geçmiş bilgilerin ve deneyimlerin etkili olduğu görülecektir. Dolayısı ile deneyimsizlik, insani durumlar ve benzeri anlık veya kalıcı problemler nedeniyle kararlar gerektiği yönde verilemeyebilir. Bunun sonucunda da hata toleransı çok düşük olan tıp alanında ölüme kadar varabilen istenmeyen sonuçlar ile karşılaşılabilir.

Tıbbi Karar Destek Sistemi, sağlık çalışanlarına teşhis ve tedavide yardımcı olmak için geliştirilmiş bilgisayar temelli yazılım sistemleridir. Günümüzde gittikçe yaygınlaşarak kullanılan Karar Destek Sistemleri aşağıdaki alanlarda hekimlere yardımcı olmaktadır:

- Tanı – Tedavi
- Bilginin saklanması – çağırılması
- Uyarı ve anımsatma
- Tedavi planlama, değerlendirme, reçete sunma
- Görüntü tanıma ve yorumlama

Günümüz karar destek sistemleri Şekil 3.8’ de gösterildiği gibi yapay zekâ (YZ), yazılım mühendisliği ve simülasyon gibi farklı disiplinler kullanmaktadır. Tıbbi veri miktarının çok büyük boyutlara ulaşması, uzun değerlendirme sürelerinin gerekmesi, algıda monotonluk, hastalık çeşitlerinin ve tedavi yöntemlerinin çokluğu gibi nedenlerden dolayı insan hataları da sıklıkla ortaya çıkabilmektedir. Çok miktardaki tıbbi verilerin doğru bir şekilde değerlendirilmesi ise çok sayıda nitelikli uzman hekime ihtiyaç duyulmasına neden olmaktadır. Oysa tüm dünyada uzman hekimlerin sayısı sınırlıdır. Bu nedenle, tıbbi verileri zeki yaklaşımlarda içerecek biçimde hızla ve doğru bir şekilde değerlendirebilecek sınıflandırma ve değerlendirme yeteneğine sahip karar destek sistemlerinin geliştirilmesi son derece önemlidir. Uzman hekim değerlendirmesinin yanı

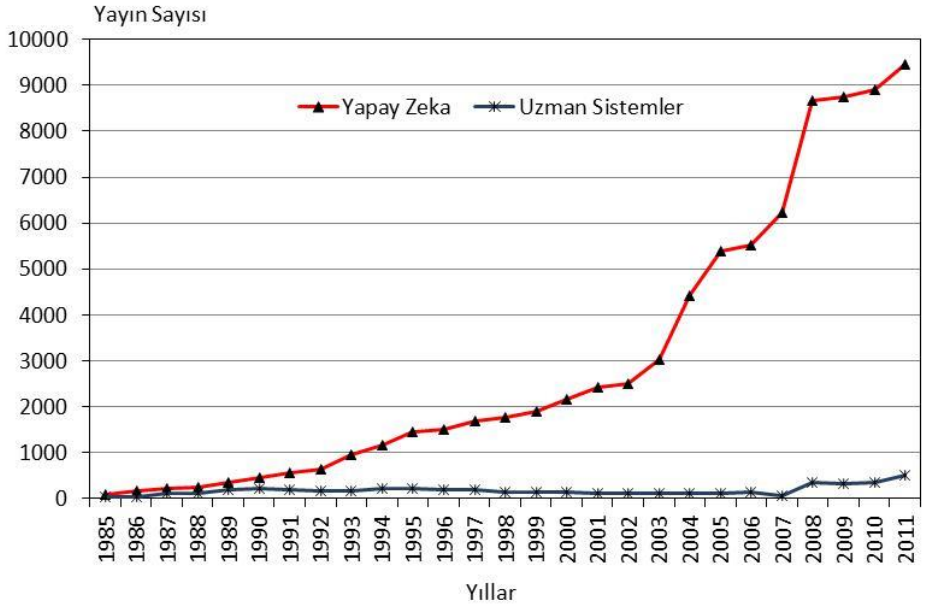
sıra zeki bir sınıflandırma ve değerlendirme sisteminin kullanılması gerek teşhis gerekse tedavideki başarıma önemli katkılar sağlayacaktır [83].
Genel olarak, karar destek sistemleri aşağıdaki avantajlara sahiptirler:

- Birbirini etkileyen hastalıkların tespitinde başarılı sonuçlar sağlarlar,
- Kompleks çoklu hastalık durumundaki yanılgıları minimize edebilirler,
- Eğitim ve araştırmada yeni araçlar sağlar,
- Hastalara da yardımcı olabilir,
- Maliyet, zaman ve tıbbi hataların azaltılmasını sağlar.

Geleneksel istatistiksel yöntemler genellikle lineer ilişkiye sahip değişkenler arasındaki ilişkilerin tespitinde başarılı olabilmekte, büyük boyutlu veriler ve doğrusal olmayan ilişkilerin tespit edilmesinde başarılı olamamaktadır. Ayrıca, veriler üzerinde sınırlı sonuçlara ve yargılara erişilebilmesine imkan verebilmektedir. Hastalıkların sınıflandırılması, görüntülerden bilgi çıkarımı, hastalık seyrinin kestirimi, tedavi planlama ve hasta verileri arasındaki bilinmeyen ilişkilerin tespit edilmesi gibi tıbbi uygulamaların zorluğu araştırmacıları daha modern ve güçlü yöntemler kullanmaya sevk etmiştir. Bu amaçla, araştırmacıların ilgisi doğada ve insanda var olan mekanizmaları taklit eden YZ yöntemleri üzerine yoğunlaşmıştır.

Şekil 3.8. Günümüz karar destek sistemlerinin ilişkili olduğu bilim dalları

Pubmed' de yıllara göre gerçekleştirilen yayın sayılarına bakıldığında, YZ yöntemlerinin tıp uygulamalarına yönelik yayın sayısının hızla artığı görülmektedir.



Şekil 3.9. Pubmed' de yıllara göre gerçekleştirilen yapay zekâ ile ilgili yayın sayısı

YZ yöntemleri, öğrenme kabiliyeti, genelleme yapabilmesi, adaptasyon kabiliyeti, sonuç çıkarımı, doğrusal olmayan kompleks ilişkileri kurabilmesi, gürültüye karşı tolerans ve farklı tekniklerle entegrasyon kabiliyetleri bulunması gibi sahip oldukları pek çok üstünlükler nedeniyle modern TB uygulamalarının temel bileşenlerinden birisi haline gelmiştir. Literatürde YZ tekniklerinin hastalık teşhisi, karar verme, örüntü tanıma ve sınıflandırma gibi tıbbi problemlere uygulanmasına yönelik çok sayıda çalışmaya rastlanmaktadır [84-112].

3.2.2. Trombofili Hastalığının Genel Yapısı

İnsanda damar hasarından sonra kanamanın durabilmesi için gerçekleşen normal pıhtılaşma sürecine *Hemostaz* denir. Kanın bir atardamar veya toplardamar içinde ya da bir organda pıhtılaşmasına ise *Tromboz*, pıhtıya ise *trombus* denir. *Trombofili* (Thrombophilia: trombozu sevme) ise tromboza eğilim yaratan tabloları tanımlamakta kullanılan bir terimdir. *Tromboz* gelişimi multifaktoriyel olup, çok sayıda edinsel ve kalıtsal faktörün değişik mekanizmalarla tromboz oluşumuna neden olduğu bilinmektedir. Tablo 3.2'de edinsel risk faktörleri, Tablo 3.3'de ise kalıtsal risk faktörleri verilmiştir [113, 114-125]. Tromboz genellikle bir damarın iç yüzündeki bir bozukluk sonucu meydana gelir. Bu bozukluk o tabakada oluşan bir yozlaşma (aterom, travma) nedeniyle olabileceği gibi,

kaza sonucu ya da ameliyatlar sırasında da oluşabilir. Bu duruma bazı zehirler, ilaçlar, mikrobik hastalıklar, bazı kan bozuklukları ve uzun süre yatakta kalmayı gerektiren ya da hastayı zayıf düşüren hastalıklar da yol açabilir.

Tablo 3.2. Edinsel trombofili nedenleri

Arteriyel tromboz nedenleri	**Venöz tromboz nedenleri**
İleri yaş	İleri yaş
Ateroskleroz	Genel cerrahi girişim
Sigara içme	Ortopedik cerrahi girişim
Hipertansiyon	Travma
Diabetes mellitus	İmmobilizasyon
Antifosfolipid sendromu	Antifosfolipid sendromu
LDL kolesterol yüksekliği	Konjestif kalp yetersizliği
Hipertrigliseridemi	Nefrotik sendrom
Sol kalp yetersizliği	Obezite
Atrial fibrilasyon	Malignite
Oral kontraseptif kullanımı	Varisler
Östrojen kullanımı	Gebelik
Lipoprotein(a) yüksekliği	Postpartum dönem
Polistemi	Oral kontraseptif kullanımı
Hipervizkozite sendromları	Östrojen kullanımı
Lökostazis sendromları	Behçet Hastalığı

Tromboz sözcüğü yalnız olarak kullanılırsa, çevresel (kol, bacak gibi) organlardaki atar ya da toplardamarlarda veya bu damarları besleyen daha büyük damar gruplarında olan tıkanmalar anlaşılır. Koroner tromboz ya da serebral tromboz terimleri ise kalbi ve beyni besleyen damarlardaki trombozları ifade etmekte kullanılır.

Tromboz her yönü ile önemli bir hastalık olup, tıbbın birçok dalının ilgili olduğu bir konudur. *Tromboz* şayet yüzeydeki toplardamarlarda ise bu varisdir ve nispeten az tehlikelidir. Ancak, *Tromboz* bacak ve kalçanın iç toplardamarlarında oluşursa daha ciddi problemler doğurur. Derindeki damarlardan kopan kan pıhtısı kalbe doğru akar, kalpten geçen pıhtı akciğerde damar tıkanıklığına (embolizm) sebep olur. Buda kan akışının yavaşlaması neticesinde kalbin zorlanması ve hatta ölüm demektir. Nadiren bu pıhtı beyin damarlarını tıkayarak beyin kanaması veya kalp damarlarının tıkanması sonucu kronar tıkanmasına neden olur. Takriben insanların % 0,2'sinde tromboz görülmekte olup, kadınlarda erkeklere göre daha yaygındır.

Tablo 3.3. Kalıtsal trombofili nedenleri

Bozukluk	**Toplumdaki Sıklığı (%)**	**Trombozlu Hastalarda Sıklığı (%)**
Antitrombin eksikliği	0,02	1
Protein C eksikliği	0,2	3
Protein S eksikliği	0,1	1-2
APC direnci/FV Leiden mutasyonu	3-6	20
Hiperhomosisteinemi	5-10	10-25
Protrombin 20210 alleli	1-2	6
FVIII yüksekliği	1	25

Venöz trombozlar (toplardamar trombozları) arteriyel trombozlara göre daha sık görülmekte ve daha tehlikeli olmaktadır. Arteriyel tromboz tehlikesi daha çok ileri yaşlarda ortaya çıkarken, venöz trombozlar on yaşla doksan yaş arasında görülebilmektedir.

Arteriyel sistemde endotel hasarı ve trombositlerin fonksiyonel bozukluklarının önemli rol oynadığı, venoz sistemde ise daha çok staz ve pıhtılaşma sistemine ait bozuklukların tromboz gelişimine neden olduğu bilinmektedir [113]. Toplardamarda oluşan pıhtılaşmaların nedeni patolojik nedenlerdir. Bu nedenler bölgesel ya da genel olabilir. Etkisi atardamarlara, toplardamarlara ya da her ikisine birden olabilir.

Doğumdan sonra görülen toplardamar trombozları oldukça sıktır. Özellikle doğum güç olmuşsa, ya da herhangi bir terslik olmuşsa tromboz olasılığı artar. Bulaşıcı hastalıklardan sonra yan etki olarak da toplardamar trombozu görülebilir. Tifo, verem, çocuk felci, beyin yangısı gibi virüs ve bakteri hastalıkları bu duruma neden olabilir. Uzun süren kalp hastalıkları, özellikle, toplardamar trombozuna, tromboz da ölüme neden olabilecek akciğer amboluslarına yol açabilir.

Kanserle toplardamar trombozları arasında ilginç bir ilişki vardır. İlerlemiş kanser hastalarının son zamanlarında toplardamar trombozları görülebilir. Ayrıca kanser ameliyatı geçirmiş hastalarda da ameliyat sonrasında bu tür trombozlar olağandır. Toplardamar trombozları, bazen pankreas ve akciğer kanserlerinin etken belirtileridir. Kanser ve belirtileri bu durumdan çok sonra ortaya çıkar. Hatta metastaz denilen ikincil kanser odakları oluştuktan sonra bu belirtiler görülebilir.

Toplardamar trombozlarında en çok korkulan sonuç akciğer ambolusudur. Fakat akciğer ambolusunun her zaman ölümle sonuçlanmayacağı, üstelik böyle bir durumla karşılaşma tehlikesinde olan hastalara bu açıdan özen gösterilmesiyle ambolus tehdidinin azaldığını belirtmek yerinde olur. Akciğerde oluşan ambolus akciğerin bir kısmının ya da tamamının kanla beslenmesini engeller. Akciğer amboluslarının bir kısmı vücudun savunma sisteminin etkisiyle çabucak ve kendi kendine erir.

Toplardamar trombozlarının tedavisi daha çok önleme çabalarına dayanır. Bunlar ameliyat sonrasında hastayı çabuk ayağa kaldırıp yürütmek, kan, pıhtılaşmasını önleyen ilaçların kullanılması ve oluşmuş bir trombusun ameliyat yöntemleriyle alınmasıdır. Uzun süre yatakta kalması gerekli hastalar, özellikle bir toplardamar dolaşım bozukluğu ya da varis hastalığı varsa, yatakta esnek çorap (varis çorabı gibi) giyerek yatmalıdırlar.

Son yıllarda ağızdan alınan doğum kontrol hapları ile tromboz arasında bir ilişki olduğu kanısı doğmuştur. 1967'de yapılan bir incelemede böyle bir ilişkinin var olabileceği açıklanmıştır. Hapların yapısında bulunan östrojenin bu duruma yol açtığı anlaşılmıştır. 1961'de kalp ve beyin damarları hastalıklarında bu madde bol olarak kullanılmış ve tromboambolus olaylarının artmasına yol açmıştır.

3.2.2.1 Kalıtsal Trombofili

Kalıtsal trombofililer yani pıhtılaşma bozuklukları genel olarak basit tanımlamayla kanın pıhtılaşmaya eğilim gösterdiği bazı hastalıklardır. Bu hastalıklarda kanın pıhtılaşmasına bağlı damar tıkanıkları kalp, akciğer ve beyin gibi organlarda pıhtı oluşması, gebeliklerin düşük veya ölü doğum veya rahim içi gelişme geriliği ile sonuçlanması gibi problemler yaşanmaktadır. Kalıtsal trombofililer genetik bozukluklara bağlı olarak ırsi yani kuşaktan kuşağa geçebilen ve doğuştan edinilmiş hastalıklardır.

Trombotik olayların ailesel olabileceği yüzyılın başında dikkati çektiği halde, pıhtılaşma sistemi konusundaki bilgilerin netleşmesi ile ancak 1960’ların sonunda kalıtsal trombofililer aydınlatılmaya başlanmıştır. 1965’de antitrombin eksikliği, 1981’de protein C eksikliği ve 1984’de protein S eksiklikleri tanımlanmıştır. Bu üç eksiklik, kalıtsal trombofililerin sadece %15’ini oluşturmaktadır [126,127]. 1993’de APC direnci [128] ve 1994’de Faktor V Leiden mutasyonunu tanımlamaları trombofilili ailelerin %50’sinde,

trombozlu hastaların ise %20'sinde etyolojinin aydınlatılmasını sağlamıştır [129]. Yine 1994'de hiperhomosisteineminin [130], 1996'da protrombin geninde bir mutasyonun (protrombin 20210 alleli) kalıtsal trombofiliye yol açtığı gösterilmiştir [131].

Bunlar dışında her geçen gün kalıtsal trombofiliye neden olduğu iddia edilen bozukluklar tanımlanmaktadır. Bununla beraber günümüzde hala kalıtsal trombofili düşünülen vakaların %40-60'ında tüm incelemelere rağmen nedeni ortaya koymak mümkün olamamaktadır. Kalıtsal trombofili nedenleri Tablo 3.3'de özetlenmiştir.

Kalıtsal trombofili nedenlerini genetik olarak taşıyan bireylerde tromboz riski artmakla birlikte, yaşamları boyunca hiç bir trombotik atak geçirmemeleri de mümkündür. Bu durum, tek başına kalıtsal nedenlerin yeterli olmadığını, tromboz gelişiminde bazı edinsel faktörlerin de katkısı olduğunu göstermektedir [114, 116-118, 132].

Trombofilinin multifaktöriyel nedenlere bağlı olması nedeniyle tanı konmasında zorluklar yaşanmaktadır. Örneğin klinik olarak derin ven trombozu (DVT) tanısı konulan hastaların çok az kısmında gerçek hastalık bulunduğu ve bazı yüksek riskli olgularda ise lokal bulgu ve semptomlar olmadan DVT geliştiği görülmektedir. Yine bazı hastalarda ani gelişen ve ölümcül seyirli Pulmoner Emboli (Akciğer embolisi) görülebilmektedir.

Klinik tanısal yaklaşımda hastaların öyküsüne ilave olarak fiziksel muayene ve rutin laboratuvar testleri yapılmaktadır. Seçilen hasta alt guruplarına ise kalıtımsal trombofili sebeplerine yönelik genetik testler uygulanmaktadır. Ancak kalıtsal trombofili tanısı için yapılacak testler oldukça zahmetlidir ve pahalıdır [133]. Ayrıca, uygun testler kullanılmazsa yanıltıcı sonuçlar elde edilebileceğinden bu hastalara uygulanacak testlerinde doğru seçilmesi de önemlidir. Genetik test yapılacak hastaların seçiminde göz önünde bulundurulan unsurlar aşağıdaki gibi özetlenebilir:

- 0-45 yaşından önce oluşan ve nedeni açıklanamayan tromboembolizm atakları olanlarda
- Alışılmadık bölgelerde (serebral, üst ekstremite, batın içi damarlarında) tromboz gelişenlerde
- Tekrarlayıcı, gezici veya masif tromboz öyküsü bulunanlarda
- Ailesinde tromboembolizm öyküsü saptananlarda

- Warfarine bağlı deri nekrozu öyküsü olanlarda
- Neonatal tromboz öyküsü olan kişilerde

Genetik tetkik aşamasında genellikle 4 farklı bölgeye bakılmaktadır.

Faktör V: Faktör V kan koagülasyon sistemindeki en önemli proteinlerden biridir. Faktör V aktive edilmiş protein C (APC)' nin kofaktörü olarak fonksiyon yapar. APC ile birlikte Faktör VIIIa'yı inaktive eder. Faktör V Leiden mutasyonu otozomal dominant olarak kuşaktan kuşağa iletilir. Bu mutasyon varlığında faktör V'in APC tarafından inaktive edilmesi engellenir ve kanda yoğun pıhtılaşması meydana gelir. Tromboz tehlikesi doğum kontrol hapı kullanımı, sigara, ileri yaş ya da pıhtılaşma sistemini etkileyen diğer tromboz etkenleri (Prothrombin 20210) gibi risk faktörlerinin varlığında çok daha fazla artar.

Faktör II (Prothrombin 20210A): Prothrombin 20210 mutasyonu Faktör V polimorfizminden sonra tromboz oluşumuna en sık neden olan polimorfizimdir. Tromboz, kandaki protrombin miktarının artmasıyla ortaya çıkar.

MTHFR enzimi protein ve polyamine sentezi ve metilasyon reaksiyonları için gerekli olan metil alış-verişinde rol oynar. MTHFR varyantları nöral tüp defekti, homosisteinemi, erken yaşta kalp damar hastalığı, erken yaşta venöz tromboz ortaya çıkma riskini arttırır.

MTHFR C677T: MTHFR C677Tnin homozigot olduğu durumlarda yüksek plazma homosistein ve düşük folat seviyesi gözlenmektedir. Enzim aktivitesi % 30 civarlarındadır.

MTHFR A1298C: MTHFR A1298C, homocysteine ve plazma folat konsantrasyonunu etkiler fakat etkisi C677T formundan daha düşüktür. Enzim aktivitesi %65 civarlarındadır.

Seçilen hastalara uygulanacak testlerin seçimi olguların olasılık olarak "Kuvvetle" ve "Zayıf" trombofilik olmasına göre yapılır. Aşağıda bu sınıflandırmada genel olarak dikkate alınan hususlar verilmiştir:

Kuvvetle trombofilik:

- İlk venöz tromboz atağının 50 yaşından önce görülmesi,
- Tekrarlayan trombotik atak öyküsü,
- 50 yaşından önce tromboz saptanmış birinci derece akraba mevcudiyeti.

Zayıf trombofilik:

- İlk idiopatik venöz tromboz atağının 50 yaş ve üzerinde görülmesi,
- Tromboz yönünden negatif aile öyküsü.

Buna göre uygulanması önerilen testler Tablo 3.4'de verilmiştir.

Tablo 3.4. Uygulanması önerilen testler

Test Önerilen Durum	**Kuvvetle Trombofilik**	**Zayıf Trombofilik**
Aktive protein C direnci	Evet	Evet
Protrombin mutasyonu	Evet	Evet
Antifosfolipit antikor	Evet	Evet
Plazma homosisteini	Evet	Evet
Antitrombin eksikliği	Evet	Hayır
Protein C eksikliği	Evet	Hayır
Protein S eksikliği	Evet	Hayır

Günümüzde geniş ölçekli DNA analizlerinin yapılabiliyor olması tek bir nükleotid düzeyinde venöz trombozun genetiğinin çalışılmasına imkan sağlamaktadır. Bu durumun tromboz riskinin genetik yapısında, kişisel tromboz kestiriminde benzersiz imkanlar sağlaması beklenmektedir [134].

Çalışmalar, çoklu parametrelere bağlı kompleks hastalıkların tanısının tahmin edilmesinde, genel popülasyonda ilgili risklerin genliğinin düşük olması ve genotiplerin tamamlanmamış özellikleri (imcomplete penetrance of genotypes) nedeniyle genetik tetkikler sınırlı bilgi sağlandığını göstermektedir. Ancak bu tür çalışmalar genellikle aynı anda bir genin etkisini incelemektedir. Ancak bu tür kompleks hastalıkların tanısında genetik ve çevresel faktörlerin birlikte değerlendirilmesi hastalığın etiyolojisi için daha anlamlı ve önemli sonuçlar sağlayacağına inanılmaktadır [135]. Ancak çoklu

parametrelerin trombofili hastalığının tanısındaki etkilerinin araştırılmasına yönelik çalışmalar henüz sınırlıdır [136,137]. Ayrıca henüz kişinin venöz tromboz riskinin doğru tahmin edilmesine yönelik risk modelleri de mevcut değildir ve bu konu oldukça zordur [138]. Venöz trobbozun matematiksel olarak tahmin edilmesine yönelik yalnızca birkaç çalışma bulunmaktadır. Çalışmalar göstermektedir ki, kişisel emboli riskinin etkili bir şekilde tahmin edilebilmesi ancak klinik, çevresel ve genetik değişkenleri de dikkate alan sofistike modeller ile mümkün olabilecektir [139]. Souto ve Sorio Aralık 2012'de blood dergisindeki Haan ve arkadaşlarının çalışmasına yönelik eleştirel makalelerinde aşağıdaki hususlar üzerinde çalışmanın kişisel risk faktörlerinin belirlenmesinde daha etkili olabileceğini önermişlerdir:

- Henüz bilinmeyen yeni genetik risk faktörlerinin araştırılması,
- Risk analizinde taşıyıcılarda güçlük etkileri olan ender görülen genetik varyantların birleştirilmesi,
- Koruyucu tüm faktörlerin hastalıktan korunmada kişisel risk faktörleri ile daha doğru bir tahmin için birleştirilmesi,
- Klinik uygulama açısından DNA genotipleme için yeni laboratuvar metotlarının geliştirilmesi,
- DNA verisinin yanı sıra, trombin generasyonu, tromboelastimetri, platelet hacmi ve Faktör III plazma hacmi gibi yeni markerlerdan elde edilen biyolojik bilgilerin de kullanılması,
- Klinik, genetik, genomik ve biyolojik parametreleri içeren çok boyutlu yapıların entegre edildiği daha özel matematiksel modellerin geliştirilmesi.

Bu hastalığın ürik asit, LDH, Fibrinojen, Ddimer gibi serum biyokimyasal parametreler ve tam kan sayımı, MCV ve MPV gibi hemogram parametreleri ile ilişkisinin ortaya konması hususunda yapılan çalışmalar ise oldukça sınırlıdır.

Biyokimyasal ve Hemogram tetkikleri ile belirlenen parametreler ile hastalığın varsa ilişkisinin ortaya konması gerek klinik değerlendirmeler ve gerek ise genetik tetkiklerden hareketle hastalığın tanısında önemli kolaylıklar sağlayabileceği değerlendirilmektedir.

Penco ve arkadaşları, YSA kullanarak venöz torombozun önemli risk faktörlerinin belirlenmesine yönelik bir çalışma yapmışlardır. Yaptıkları çalışmada YSA'nın diğer

analitik metotlara göre daha başarılı sonuçlar verdiğini göstermişlerdir [140]. Shanthi ve ark. beyinde troembolik tıkanma olasılığını tahmin etmek için çok katmanlı ileri beslemeli YSA modeli kullanmışlardır. Hastanın yaş, cinsiyet, hipertansiyon, diabet, sigara ve alkol kullanımı, kalp damar tıkanıklığı, görüntü kaybı gibi 25 parametrenin dikkate alındığı bu çalışmada YSA'nın beyinde troembolik tıkanıklık ihtimalinin belirlenmesinde başarılı bir şekilde kullanılabileceği gösterilmiştir [141]. Dahabiah ve ark. Klinisyenlerin daha önce değerlendirdiği ultrason görüntülerinden venöz trombolinin karakterize edilmesi üzerine bir çalışma yapmışlardır [142]. Markus ve ark. Doppler ultrason sinyallerinin sinyal ilmeme teknikleri ile değerlendirilerek embolik sinyallarin tespi üzerine bir çalışma yapmışlardır. Dikkate alınan veriler üzerinde başarılı sonuçlar elde edilmesine rağmen, dikkate alınan yöntem çevrim için değerlendirme yapabilecek bir otomasyon sistemi olmaması nedeniyle geliştirilmeye muhtaçtır [143]. Kemeny ve ark. Yaptıkları çalışmada doppler ultrason sinyallerinin kullanarak emboli tespiti için YSA kullanmışlardır [144]. Yaptıkları çalışmada sonuçların çok başarılı olmadığını ancak bu başarısızlıkta uygun sinyal-gürültü oranının uygun seçilmemesinin etkili olduğunu ifade etmişlerdir. Bununda bazı eşikleme algoritmaları ile aşılabileceğini ve emboli tespitinde otomasyon sistemlerinin geliştirilmesine büyük ihtiyaç duyulduğunu ifade etmişlerdir.

3.2.3. Deneysel Çalışmalar ve Analiz

Bu tez çalışmasında, Trombofili hastalığı tanısı konmuş bu hastaların klinik verileri ile genetik bozukluk olup olmama durumu arasındaki ilişkinin belirlenmesi amacıyla Adaptif Ağ Tabanlı Bulanık Mantık Çıkarım Sistemi (ANFIS)'nin kullanılması araştırılmıştır. Elde edilen sonuçlar sınıflandırma algoritmalarına ait sonuçlarla karşılaştırılarak yorumlanmıştır.

3.2.3.1. Makine Öğrenmesi Algoritmaları ile Başarım Analizi

Tez çalışması kapsamında Erciyes Üniversitesi Tıp Fakültesi Genetik Anabilim dalından temin edilen 180 hastaya ait veriler dikkate alınmıştır. Veri setinde uzman hekimlerin öneri ve yönlendirmeleri ile genetik bozukluk olup olmaması hususunda etkili olabileceği öngörülen 47 farklı parametreye ait bilgilere yer verilmiştir. Ancak, veriler üzerine yapılan incelemelerde hastaların tamamının dikkate alınan 47 parametreye ait verilerinin

bulunmadığı görülmüştür. Analizlerin genellenebilir sonuçlar üretebilmesi amacıyla mümkün olduğunca fazla sayıda hasta için mevcut olan ve istatistiksel olarak anlamlılık düzeyi yüksek olan veriler üzerine gerekli çalışmaların yapılması kararına varılmıştır.

Veriler üzerinde yapılan incelemede hangi parametre verilerinin kaç hastada mevcut olduğu tespit edilmiştir. Çalışmamızda uygulayacağımız mühendislik modeline uygun sayıda eğitim ve test veri seti belirleyebilmemiz için yeterli miktarda hasta sayısı olması gerekmektedir. Bu kapsamda, elimizde bulunan verileri dikkate aldığımızda 70 adet hastanın çalışmamız için yeterli olduğu düşünülerek 72 veya daha az sayıda hastada bulunan parametreler veri setinden çıkartılmıştır. Sonuç olarak, uzman hekimlerin görüşü de alınarak 180 hastaya ait 39 adet parametrenin dikkate alınmasına karar verilmiştir.

ERÜ Genetik Anabilim dalı öğretim üyeleri ile yapılan çalışmalarda, *Demir* ve *Demir Bağlama* parametrelerinin birbirini temsil edebilir parametreler oldukları öngörülmüştür [145,146]. İstatistiksel yöntemlerden olan *Roc* analizi verilerimize uygulanmış, elde edilen sonuçlar doğrultusunda *Demir Bağlamanın* genetik bozukluk varlığının tespitinde *Demir* parametresinden daha etkili olduğu görülerek, *Demir* parametresi veri setinden kaldırılmıştır. *Roc* analizinde her değişken için ayrı ayrı *Roc* eğrileri çizilmiş ve eğri altında kalan alanları hesaplanmıştır. Eğri altında kalan alanlar *Demir* ve *Demir Bağlama* parametreleri için sırasıyla 0,569 ve 0,585 olarak elde edilmiştir. Dolayısıyla genetik bozukluğun ayırt ediciliğinin belirlenmesinde *Demir Bağlamanın* daha yüksek bir kestirim gücü olduğu belirlenmiştir. Aynı şekilde ERÜ Genetik Anabilim dalı öğretim üyelerinden alınan bilgiler doğrultusunda birbirini temsil etme durumunun *Glukoz-Alkalen Fosfataz, LDL-APC Resistansı* ve *Kolestrol-Magnezyum* parametreleri arasında da olduğu anlaşılmıştır [147,148]. Yine bu parametreler için de yapılan *Roc* analizi neticesinde *Glukoz, APC Resistansı* ve *Magnezyum* parametrelerinin genetik bozukluk varlığının tespitinde daha etkili oldukları anlaşılmıştır. Böylelikle *Alkalen Fosfataz, LDL* ve *Kolestrol* parametreleri de veri setinden çıkarılmıştır. Sonuç olarak veri setinde, parametre sayısı 35 olan toplam 180 hastaya ait veriler bırakılmıştır. Tablo 3.5'de dikkate alınan 35 adet parametre ve bu parametre değerine sahip hasta sayıları verilmiştir.

Yapılan bu çalışmalardan sonra kalan hasta verileri içerisinde hala 35 parametrenin tüm değerleri dolu olmayan hastaların bulunduğu görülmüştür. Oluşturulacak model için kullanılacak örnek veri seti boyutunu daha yüksek tutabilmek amacıyla, hastaların boş

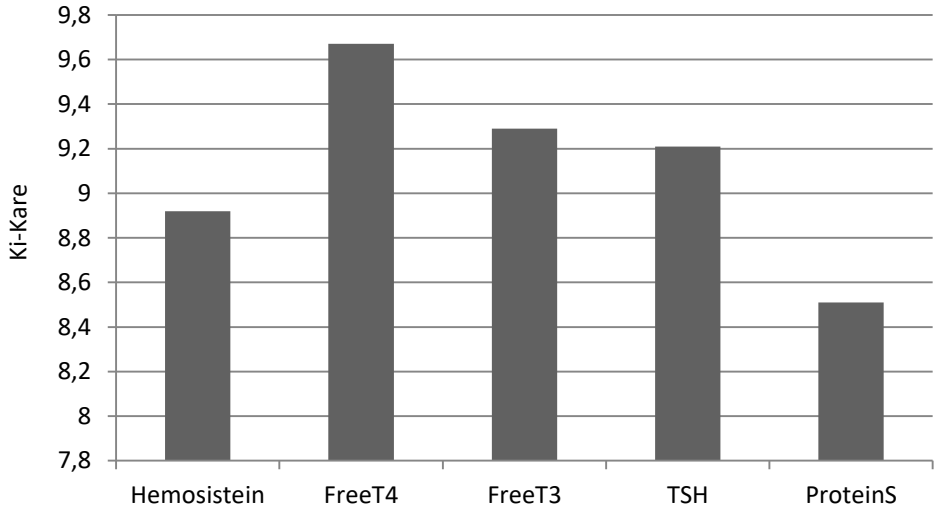
olan parametre değerlerinin *Eksik Veri Atama* (*Missing Value Imputation)* yönteminin *Beklenti-Maksimizasyonu* (*Expectation-Maximization)* algoritması ile rastgele doldurulması yoluna gidilmiştir [149]. Ancak, bu yöntemle üretilen parametre değerlerinin büyük oranda klinik olarak anlamlı olmadığı görülmüş ve üretilen bu değerler çalışmada kullanılmamıştır.

Tablo 3.5**.** Parametreler ve bu parametre değerine sahip hasta sayıları

Parametre	**Hasta Sayısı**
Kreatinin	182
PROTEIN S	182
ALT	179
AST	177
BUN	173
Glukoz	171
APC Resistansı	167
TSH	164
FREET4	157
Hemosistein(HPLC İle)	154
Sodyum	153
Total_Bilirubin	148
Direk_Bilirubin	146
LDH	144
FREET3	143
ACA IGG	139
ACA IGM	138
Kalsiyum	134
Albümin	133
Total Protein	130
Potasyum C	129
Ürik_asit	114
Potasyum	112
GGT	107
Klor	103
Fosfor	99
HBS AG (Doğrulama Dahil)	89
Magnezyum	86
Vitamin B-12	85
Trigliserit	82
HDL Kolesterol	81
Sedimantasyon	80
CPK	78
Demir Bağlama	78
ANTİ-HCV1	72

Genellenebilir sonuçlar elde edebilmek için eğitme ve test gurubunda yer alacak hasta sayılarının belirli bir miktarda olması gerektiği bilinmektedir. Bu nedenle toplamda en az 100 hastaya ait bir veri seti ile çalışmanın uygun olacağı kanaatine varılmıştır. Bu kapsamda yapılan incelemeler neticesinde yine ERÜ öğretim üyeleri ile yapılan çalışmalar neticesinde, klinik olarak anlamlı olabilecek 8 adet parametre değeri bulunan (*ACA IGG,ACA IGM, APC Resistansı, FreeT3, FreeT4, Hemosistein, Protein S* ve *TSH*) 103 adet hasta olduğu görülmüştür. Bu hastaların 60'ında genetik bozukluk mevcut olup, 43 tanesinde ise genetik bozukluk bulunmamaktadır.

103 hastaya ait veriler üzerinde istatistiksel çalışmaların *Öznitelik Seçimi* (*Feature Selection)* yöntemi uygulanarak 8 adet parametrenin anlamlılık düzeyleri belirlenmiştir. Bu yöntemde Ki-kare *(X^2)* dağılımı testi kullanılmıştır. Bu test ile sınıf değişkeni ve her öznitelik arasındaki bağımsızlık araştırılmıştır. Elde edilen test istatistikleri sıralanarak öznitelik seçimi yapılmıştır [150]. Öznitelik seçimi kriteri olarak $p<0,10$ düzeyi kullanılmıştır. Yapılan bu çalışma ile sadece *FreeT3, FreeT4, Hemosistein, Protein S* ve *TSH* parametrelerinin model için anlamlı oldukları görülmüştür. Anlamlı olarak bulunan bu 5 parametrenin *Öznitelik Seçimi* yöntemiyle bulunan Ki-kare değerleri Şekil 3.10'da verilmiştir. Yapılan istatistiksel ön çalışmalar sonucunda, başarım analizleri için dikkate alınan veri setinde 103 adet örnek ve 5 adet öznitelik yer almıştır.



Şekil 3.10. Öznitelik Seçimi yöntemiyle bulunan ki-kare değerleri

Lojistik Regresyon, Sınıflandırma Ağaçları, Rastgele Orman ve *Destek Vektör Makinaları* yöntemleri kullanılarak, Trombofili tanısı konmuş 103 hastaya ait verilerden

hareketle, belirlenen 5 parametre ile hastalarda genetik bozukluk bulunup bulunmadığının tespit edilmesine yönelik analizler yapılmıştır. Hasta sayısının sınırlı olması nedeniyle, modelin genelleştirebilme yeteneğini doğru bir şekilde değerlendirebilmek için *çapraz doğrulama (cross validation)* yöntemlerinden *K parçalı* yöntemi kullanılmıştır. *K parçalı çapraz geçerlilik* yönteminde amaç bir deneyi bağımsız koşullarda yineleyerek sonuçlarının geçerliliğini sınamadır. Örneğin, istatistiksel sınıflandırma problemlerinde, bir veri kümesini aşağı yukarı eşit *k* tane kümeye bölüp, her bir *k* için, sınıflandırıcıyı oluşturmak için, *k*-1 kümeyi kullanıp ve arta kalan *k*'ıncı küme üzerinde test işlemini yapmaktır. Bu çalışmada, 5-kat çapraz doğrulama yöntemi kullanılmış ve elde edilen sonuçlar Tablo 3.6'da verilmiştir. Sonuçlardan ilgili yöntemlerin bu problem üzerindeki performanslarının genel olarak düşük olduğu görülmüştür. Özellikle literatürde en etkili sınıflandırma algoritması olarak bilinen *Destek Vektör Makinaları* yönteminin de anlamlı sonuçlar vermediği görülmüştür. Çalışmada kullanılan sınıflandırma algoritmaları için *Statistica 7 (Stat Soft Inc., Tulsa, USA)* ve *R 3.0.1* (www.r-project.org) yazılımının *caret* kütüphanesi kullanılmıştır [151]. Destek Vektör Makinaları yönteminde, C-SVM (C-Style Soft Margin Support Vektor Machine) algoritması ve radyal tabanlı çekirdek (kernel) fonksiyonu kullanılmıştır. Cost parametresi 0,5 olarak optimize edilmiştir [152].

Duyarlılık kriteri genetik bozukluk var olarak bulunan hasta sayısının gerçekte genetik bozukluk bulunan hasta sayısına oranını, Seçicilik kriteri ise genetik bozukluk yok olarak bulunan hasta sayısının gerçekte genetik bozukluk bulunmayan hasta sayısına oranını vermektedir. Bu iki kriter sınıflandırma modelinde iki önemli başarım ölçütüdür. Tablo 3.6'da verilen sonuçlar incelendiğinde *Destek Vektör Makinaları* için başarım oranı yüzdesi diğerlerine göre yüksek olsa da, seçicilik değerinin 0 olması bu yöntem için elde edilen sonuçların da uygulanabilir olmadığını ortaya koymaktadır.

Tablo 3.6. Farklı yöntemlere ait sınıflandırma sonuçları

Yöntem	**Başarım (%)**	**Duyarlılık**	**Seçicilik**
Lojistik Regresyon	56,3	0,6500	0,4419
Sınıflandırma Ağaçları	56,3	0,6500	0,4419
Rastgele Orman	56,3	0,7833	0,2558
Destek Vektör Makinaları	57,2	0,9833	0,0000

3.2.3.2. ANFIS ile Başarım Analizi

Sınıflandırma algoritmaları ile elde edilen sonuçları karşılaştırmak üzere Yapay zekâ tekniklerinden ANFIS kullanılarak ilave çalışmalar yapılmıştır. Bu çalışmada ANFIS için nonlinear üyelik fonksiyonları ve verilerin hazırlanmasında çapraz geçerlilik (cross validation) yöntemi kullanılmıştır. Toplam 103 hastaya ait veri aşağı yukarı olmak üzere 5 parçaya bölünmüştür. Her bir parça kadar tekrar edilerek 4 parça veri eğitim için ve kalan 1 parçada test için kullanılarak elde edilen sonuçlar kaydedilmiştir. Sınıflandırma modelini oluşturmak ve oluşturulan modelin test işlemin yapmak için Matlab ortamında gerekli kodlama çalışmaları yapılmıştır. Kodlama içerisinde ANFIS modelinin tüm giriş değerlerine 0 ve 1 aralığında veri normalizasyonu uygulanmıştır. ANFIS modelinin en uygun üyelik fonksiyonu, üyelik fonksiyonu sayısı, iterasyon sayısı ve modelin çıkış fonksiyonunu belirleme çalışmaları yapılmıştır. ANFIS parametrelerini eğitmek için BP ve SA algoritmaları kullanılmıştır. Modelin birçok kez farklı parametreler ile koşturulmasından sonra ANFIS modelinin üyelik fonksiyonu *trimf*, üyelik fonksiyonu sayısı sırasıyla 5 giriş için *[2 3 3 3 3]*, iterasyon sayısı *100* ve çıkış fonksiyonu *constant* olarak seçilmiştir. Oluşturulan ANFIS modelinde 103 hastadaki 5 giriş parametresi ve bir çıkış parametresi için başarım sonuçları Tablo 3.7'de verilmiştir.

Tablo 3.7. ANFIS modeli ile başarım sonuçları

Yöntem	**Başarım (%)**	**Duyarlılık**	**Seçicilik**
ANFIS-BP	63,1	0,7333	0,4883
ANFIS-SA	64,6	0,7676	0,5348

3.2.4. Genel Değerlendirme ve Sonuçlar

Bu tez çalışmasında, Trombofili hastalığı tanısı konmuş hastaların klinik verileri ile genetik bozukluk olup olmama durumu arasındaki ilişkinin belirlenmesi amacıyla Adaptif Ağ Tabanlı Bulanık Mantık Çıkarım Sistemi (ANFIS)'nin kullanılması araştırılmıştır. Bu amaçla, Erciyes Üniversitesi Tıp Fakültesi Hastanesi Genetik Anabilimdalından temin edilen 180 hastaya ait veriler kullanılmıştır. Hastalığın tespiti için önem arz eden parametrelere ait verilerin tüm hastalar için mevcut olmaması nedeniyle, bu hastalardan 103'ünün verileri bu çalışma için kullanılabilmiştir.

Hasta verileri üzerinde, *Lojistik Regresyon, Sınıflandırma Ağaçları, Rastgele Orman* ve *Destek Vektör Makinaları* yöntemleri kullanılarak yapılan analizlerde, başarımın en fazla %56,3 olduğu ve ayrıca seçicilik değerlerinin düşük olduğu görülmüştür.

ANFIS kullanılarak elde edilen sonuçlar incelendiğinde, modelinin başarım oranının, sınıflandırma algoritmalarına göre daha yüksek olduğu görülmektedir. ANFIS modelinin SA algoritması kullanılarak eğitilmesi ile başarım oranının % 64,6 olarak gerçekleştiği, duyarlılık ve seçicilik değerlerinin de diğer yöntemlere göre daha başarılı olduğu görülmüştür. Elde edilen sonuçlardan, yapay zekâ yöntemlerinin istatistiksel yöntemlere göre daha başarılı sonuçlar ürettiği görülmüştür.

Sınıflandırma algoritmalarının doğrusal olmayan karmaşık problemler üzerindeki başarımları genel olarak sınırlı olabilmektedir. ANFIS, bulanık sistemlerin ve sinirsel ağların bilinen en iyi özelliklerini kendi içyapısında birleştirmektedir. Bulanık sistemlerin, giriş bilgisini ağ üzerine dağıtarak araştırma uzayını azaltması, Sinirsel ağların ise geri yayılım davranışlarını kullanarak kontrol parametrelerinin doğrusal olmayan problemler için en uygun değerlerinin bulunması ANFIS modeline önemli üstünlükler kazandırmaktadır.

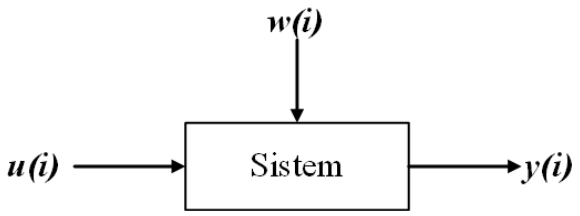
ANFIS'in eğitilmesi, sınıflandırma algoritmalarına göre daha fazla işlem sayısı ve etkin hafıza kullanımı gerektirmektedir. Ancak, ANFIS bir kez eğitildiğinde, girişlerine uygulanan veriler için çok hızlı bir şekilde sonuç verebilmektedir. Dolayısıyla, yüksek başarım düzeyi ve uygulamada girişlere karşılık çok kısa sürelerde cevap verebilmesi ANFIS yönteminin önemli bir üstünlüğü olarak ortaya çıkmaktadır.

Genel olarak, belirlenmesi gereken parametre sayısının fazla olması ve ağın eğitilmesinde kullanılan geri yayılım algoritmasının türeve dayalı olması nedeniyle bölgesel optimuma takılma riski içermesi gibi hususlar ANFIS’in başarımını olumsuz etkileyebilen temel unsurlar arasındadır. Bu kapsamda, ANFIS ağının geri yayılım algoritması yerine yapay zekâ optimizasyon algoritmaları kullanılarak eğitilmesi ve ilgili problem üzerindeki başarımının incelenmesi ise tez kapsamında sunulmuştur.

3.3. Dinamik Sistemlerin Kimliklendirilmesi

3.3.1. Sistem Kimliklendirmenin Tanımı ve Amacı

Dış dünya ve kendi iç elemanları arasındaki ilişkilerin belirlenebildiği yapılara sistem adı verilir. Bir sistemin herhangi bir andaki çıkışı, o sistemin daha önceki giriş değerlerine göre şekilleniyorsa bu sisteme dinamik sistem adı verilir [153]. Tipik bir dinamik sistem, *u(i)* giriş işareti ile *w(i)* bozucu işaretinden etkilenerek *y(i)* gibi bir çıkış işareti üretir. Şekil 3.11'de dinamik bir sistemin temel şeması gösterilmiştir.



Şekil 3.11. Bir dinamik sistem şeması

Bu dinamik sistemin kontrol edilebilmesi için, *u(i)* ve *y(i)* işaretlerinden yararlanılarak istenilen işaretin elde edilmesi gerekir. Bu yüzden sistemin matematiksel modelinin bilinmesi gerekir. İşte bu aşamada sistem kimliklendirme yönteminden yararlanılır.

Günümüzde birçok endüstriyel alanda, örneğin; kağıt, demir veya cam üretiminde, sistemin güvenli, verimli ve istenilen kalitede üretim yapılabilmesi için kontrol düzeneğinin kurulması gerekir. Kontrol edilecek dinamik sistemin matematiksel modelinin bilinmesi zorunludur. Bu dinamik sistemin önceki bilgilerden ve gözlemlerden faydalanarak bir çeşit matematiksel modelinin oluşturulması işlemine sistem kimliklendirme adı verilir [154,155]. Sistem kimliklendirme, dinamik bir sistem modelinin gerçek sistemden alınan giriş-çıkış ölçümleriyle belirlenmesi olarak tanımlanabilir [156]. Sistem kimliklendirmenin amacı, belirli bir sisteme veri giriş-çıkışı ilişkisinden yararlanarak, daha sonra sistem üzerinde tekrar yapılacak çalışmalarda kullanılabilecek güvenilir bir matematiksel model oluşturmaktır. Sistem kimliklendirme sayesinde elde edilen model ile;

1. Sistemin ileri aşamasındaki çıkış işareti tahmin edilebilir.
2. Sistem davranışı kontrol edilebilir, örneğin, hava durumu tahmini yapılabilir.

3. Sistem davranışı istenilen yönde kontrol edilebilir ve yönlendirilebilir.
4. Sistem modeline bağlı bir denetleyici tasarlanabilir, örneğin, uçak ve gemi denetleyici tasarımı yapılabilir.

Böylece sistem kimliklendirme ile birçok alanlarda karşılaşılan problemleri çözüme kavuşturmak mümkün olur.

Sistem kimliklendirme ile dinamik bir modelin matematiksel modelinde yer alan ancak bilinmeyen parametreler deneysel verilerden faydalanılarak tahmin edilir. Sistemin etkin bir biçimde kontrol edilebilmesi, çıkış değeri hakkında her zaman bilgi sahibi olunabilmesi açısından sistem kimliklendirmenin ve dolayısı ile tahmin edilen parametre değerlerinin mümkün olan en az hata ile elde edilmesi gerekir. Sistemin giriş ve çıkış değerleri arasında doğrusal olmayan bir ilişki mevcutsa ve istenmeyen bozucu girişler sistemi etkiliyorsa, parametre değerlerinin minimum hata ile belirlenmesi güçleşir. Dolayısı ile sistem kimliklendirme işlemi hem zaman hem de elde edilen performans açısından araştırmayı verimsizleştirir. Sistem kimliklendirme yönteminin araştırmayı verimsizleştirdiği noktada, yapay zekâ yöntemlerinden faydalanılarak, araştırma verimli hale getirilebilir [157].

3.3.2. Dinamik Sistemlerin Kimliklendirilmesi için SA Algoritması ile ANFIS Ağının Eğitilmesi

Sistem kimliklendirme bir sistemin giriş ve çıkış sinyalleri arasındaki davranışını tahmin etmek için geliştirilmiş bir model oluşturma sürecidir. Dinamik sistemler simülasyon ve tahmin işlemlerinin gerçekleştirilmesini kapsayan, mühendislik dalları, biyoloji, fizyoloji, meteoroloji, ekonomi, model tabanlı kontrol tasarımı ve örüntü tanıma gibi geniş bir uygulama alanına sahiptir. Dinamik sistemlerin çıkışı sistemin daha önceki giriş ve çıkış değerlerine bağlı olduğundan matematiksel olarak ancak karmaşık denklemler ile modellenebilmekte olup bu süreç genel olarak zordur. Bu sebeple dinamik sistem kimliklendirme uygulamalarında başarıyla kullanılabilecek yöntemlerin geliştirilmesi önemlidir. Bulanık sinir ağları ve yapay sinir ağlarına dayalı yaklaşımlar dinamik sistemlerin kimliklendirilmesinde yaygın olarak kullanılan yöntemlerdendir. Ancak, sahip olduğu bazı üstün özellikler nedeniyle dinamik sistemlerin kimliklendirilmesi

amacıyla ANFIS kullanılmasına yönelik çalışmalar son yıllarda giderek artmaya başlamıştır [158,159].

Yapay zekânın bir dalı olan ANFIS, örneklerden öğrenebilme ve öğrenmiş olduğu bilgiyi belirli kurallara bağlı olarak ağ üzerinde dağıtabilme yeteneği gibi esnek ve güçlü yapısı sayesinde makinalı öğrenme, sınıflandırma ve optimizasyon gibi alanlarda başarı ile kullanılmaktadır. ANFIS, bulanık sistemlerin ve sinirsel ağların bilinen en iyi özelliklerini birleştiren ağ tabanlı bir yapıdır. Bulanık sistemlerin, giriş bilgisini ağ üzerine dağıtarak araştırma uzayını azaltması, sinirsel ağların ise geri yayılım davranışlarını kullanarak kontrol parametrelerinin doğrusal olmayan problemler için en uygun değerlerinin bulunması ANFIS modeline önemli üstünlükler kazandırmaktadır.

ANFIS modelinin eğitilmesi, temelde modelin başlangıç ve sonuç parametrelerine ait en uygun değerlerin belirlenmesi sürecidir. Bu kapsamda, ANFIS'in eğitimi için BP, LSE, GD, LM ve KF gibi türeve dayalı algoritmalar yaygın olarak kullanılmaktadır. Ancak, türeve dayalı algoritmalarda her adımda eğim hesaplanması nispeten zordur ve kullanılması gereken zincir kuralı çoğunlukla yerel minimuma takılma problemine yol açmaktadır. Ayrıca türeve dayalı algoritmalarda parametrelerin yakınsaması oldukça yavaş ve çözümün başarımı büyük oranda başlangıç değerlerine bağlı olmaktadır. Bu nedenle türeve dayalı geleneksel algoritmalar ile ANFIS'in parametrelerinin eğitilmesi ve güncellenmesi temel sorunlardan biridir. Bu nedenle, ANFIS'in başarılı bir şekilde eğitilmesi ve parametrelerin güncellenmesi konusunda farklı yöntemlerin uygulanmasına ihtiyaç olup, bu konu araştırmacıların ilgisini çekmiştir. Bu kapsamda, son yıllarda çeşitli yöntemler önerilmiştir. Bu yöntemlerden bazıları yapay zekâ optimizasyon algoritmalarından olan GA, PSO, ABC ve ACO gibi sezgisel algoritmalardır [2].

GA, PSO, ACO ve ABC gibi algoritmalar popülasyon tabanlıdır. Popülasyon temelli algoritmaların çözümlerin bir seti ile çalışması genellikle onların küresel optimumun bulunduğu bölgeye hızla erişmesine olanak sağlar. Ancak, bu algoritmaların yüksek oranda olasılık tabanlı arama stratejilerine sahip olması nedeniyle bulundukları bölgedeki optimum çözümü bulabilmeleri genellikle uzun süreler gerektirebilmektedir. SA gibi algoritmalar ise iteratif tabanlıdır ve genellikle rastgele seçilen bir çözüm adayını iyileştirmek suretiyle küresel optimumu bulmaya çalışırlar. Bu tür algoritmalar genel

olarak bulundukları bölgenin optimumunu popülasyon tabanlı algoritmalara göre daha kısa sürelerde bulabilmektedirler.

Tez çalışmasının bu bölümünde, ANFIS ağının başarılı bir şekilde eğitilmesi için iteratif tabanlı bir algoritma olan SA algoritmasının kullanılmasına yönelik yeni bir yöntem önerilmiştir. Önerilen yöntem doğrusal ve doğrusal olmayan dinamik sistemlerin kimliklendirilmesi problemine uygulanmıştır. Ayrıca, önerilen yöntemin performansı popülasyon tabanlı GA, türev tabanlı BP ve literatürden alınan farklı yöntemlerin performansları ile karşılaştırılmıştır. Elde edilen sonuçlardan önerilen yöntemin diğer yöntemlerden daha başarılı olduğu görülmüştür.

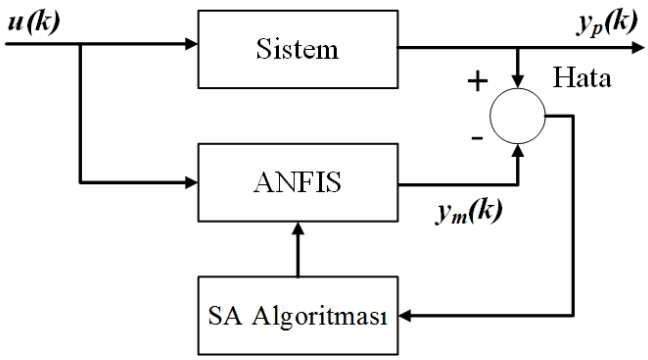
3.3.3. Önerilen Yöntemin Genel Yapısı

Bu bölümde ANFIS parametrelerinin eğitilmesi için iteratif tabanlı SA algoritmasının kullanılması açıklanmıştır. ANFIS'in başlangıç ve sonuç parametreleri olmak üzere eğitilmesi gereken iki parametre türü vardır. Başlangıç parametreleri Eşitlik 2.4'de $\{a_i, b_i, c_i\}$ olarak verilen üyelik fonksiyonlarına aittir. Buradaki a_i üyelik fonksiyonlarının varyansı, c_i üyelik fonksiyonlarının merkezi ve b_i ise diğer bir eğitim parametresidir. Başlangıç parametrelerinin toplam sayısı tüm üyelik fonksiyonlarındaki parametrelerinin toplam sayısına eşittir. Sonuç parametreleri ise Eşitlik 2.8'de $\{p_i, q_i, r_i\}$ olarak gösterilen parametrelerdir.

Şekil 2.5'de verilen temsili ANFIS mimarisinde 2 giriş, toplam 4 üyelik fonksiyonu ve iki kural vardır. Bu yapı için üyelik fonksiyonu olarak *gaussmf* seçtiğimizi varsayarsak, her bir üyelik fonksiyonunda 3 parametre bulunduğundan, bu mimariden toplamda 12 başlangıç parametresi vardır. Her bir kural için 3 sonuç parametresi bulunduğundan toplamda 6 sonuç parametresi vardır. Sonuç olarak Şekil 2.5'de verilen ANFIS mimarisinde optimize edilmesi gereken toplam parametre sayısı 18'dir ve bu parametreleri eğitmek için bu parametreler SA algoritması ile optimize edilmiştir.

Bu bölümde önerilen yöntem dinamik sistemlerin kimliklendirme problemlerine uygulanmıştır. ANFIS'den elde edilen çıkış ile doğrusal veya doğrusal olmayan dinamik sistemden alınan çıkış arasındaki hata değerinin en az olması için SA algoritması ile

ANFIS parametreleri optimize edilerek en iyi modelin oluşturulması sağlanmıştır. Bu kimliklendirme mimarisini gösteren blok diyagram Şekil.3.12‘de verilmiştir.



Şekil 3.12. Sistem kimliklendirme için ANFIS modelinin SA algoritması ile eğitilmesi planı

Eşitlik 3.3 kullanılarak elde edilen Ortalama Karekök Hata (Root Mean Square Error, *RMSE)* hata fonksiyonu çözümün hata değerinin belirlenmesinde kullanılmıştır. Burada kullanılan $y_m(k)$, k zamanında ANFIS tarafından elde edilen çıkışı ve $y_p(k)$ ise sistemin gerçek çıkışını göstermektedir. N ise uygulamada kullanılan örnek sayısını göstermektedir. Amaç, ANFIS ağının giriş ve sonuç parametrelerini ayarlayarak etkin hatayı (*RMSE*) minimize etmektir.

$$RMSE = \sqrt{\frac{\sum_{k=1}^{N}\left(y_p(k) - y_m(k)\right)^2}{N}} \tag{3.3}$$

Başlangıçta ANFIS’in tüm parametreleri için rastgele çözüm uzayı oluşturulur. Sonrasında her bir iterasyon için SA algoritması kullanılarak parametreler güncellenir. Eğitim aşamasında, önce *u(k) (k=0,1,...)* giriş sinyali dizisinin bir elemanı, mevcut çözümden sağlanan ağırlıklarla tasarlanmış ANFIS ve sistemin her ikisinin girişlerine uygulanır. Sonra ANFIS ve sistem çıkışları arasındaki hata değeri hesaplanır. Giriş sinyali dizisinin tüm elemanları için hesaplanan hata değerleri yardımıyla Eşitlik 3.3’de verilen etkin hata değeri, modelin mevcut durumunun iyiliğinin ölçüsü olarak kullanılmıştır. Mevcut durumdan sağlanan etkin hata değeri ile yeni üretilen komşu çözümle sağlanan etkin hata değeri karşılaştırılır. Eğer yeni çözüm ile sağlanan etkin hata değeri daha düşük

ise yeni çözüm mevcut çözüm olarak atanıp, araştırmaya devam edilir. Aksi takdirde Metropolis kriterine göre seçim yapılır.

Eğitmede kullanılacak giriş dizisinin seçimi eğitme başarısı üzerinde önemlidir. Giriş dizisinin elemanları genellikle sinüzoidal ve rastgele giriş değerlerinden seçilmektedir [160]. Bu çalışmada gerçekleştirilen simülasyon çalışmalarında düzgün dağılıma sahip rastgele *u(k)* giriş sinyal dizisi kullanılmıştır.

3.3.4. Simülasyon Sonuçları

Bu bölümde ANFIS modelinin SA algoritması ile eğitilmesine yönelik üç farklı dinamik sistem kullanılarak uygulamalar yapılmıştır. Birinci örnek 3.dereceden doğrusal bir sistemdir. İkinci ve üçüncü örnekler ise doğrusal olmayan sistemlerdir.

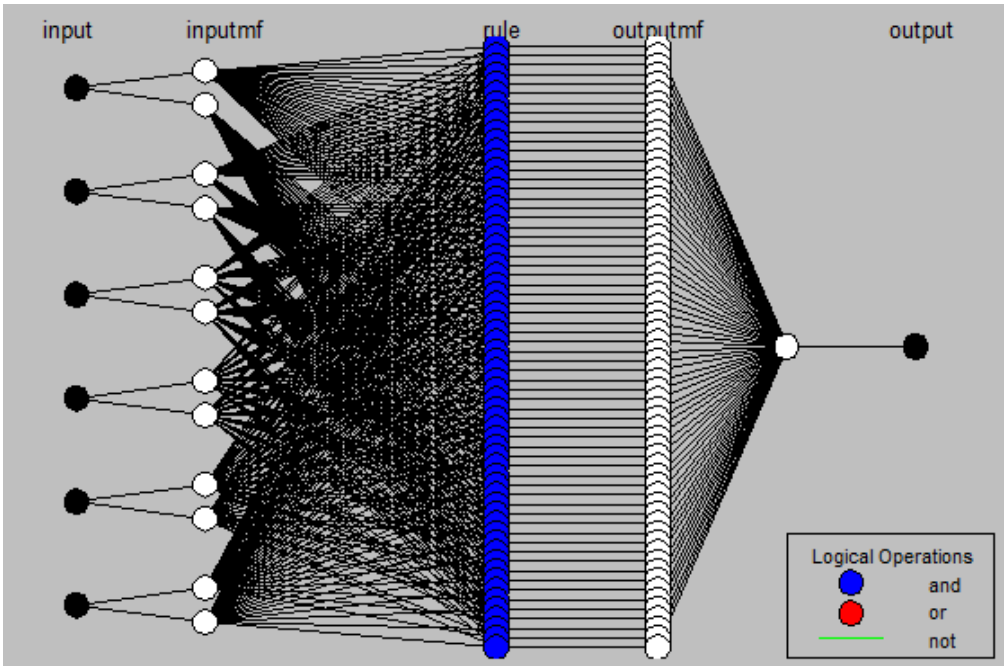
Kullanılan her bir dinamik sistem için 250 adet giriş sinyali *u(k) (k=0,1, ,.., 250)* rastgele olarak (-2, 2) aralığında üretilmiştir. Üretilen 250 adet giriş sinyali kullanılarak her bir dinamik sistemin gerçek çıkışı $y_p(k)$ hesaplanmıştır. Dolayısıyla 250 adet giriş ve çıkış sinyalinden oluşan bir veri seti hazırlanmıştır. Bu veri setlerinden rastgele seçilen 200 tanesi eğitim için, kalan 50 tanesi ise test için kullanılmıştır. 200 adet giriş sinyali $u(k)$ ANFIS ağına uygulanarak, sistemin gerçek çıkışı $y_p(k)$ ve ANFIS ağının çıkışı $y_m(k)$ arasındaki hata değeri Eşitlik 3.3'e göre hesaplanmıştır. Hesaplanan hatayı minimize etmek için ANFIS ağının parametreleri SA algoritması ile eğitilmiştir. Eğitilen ANFIS modelinin performansı ise kalan 50 adet veri seti ile test edilmiştir.

Dinamik sistemlerin kimliklendirilmesi için yapılan bu çalışmada, Izgara Bölümleme (GP) yöntemi ile sırasıyla (6,3,4) giriş ve 1 çıkıştan oluşan ANFIS yapıları modellenerek kural ve parametre sayıları belirlenmiştir. Modellenen bu yapılar sırasıyla Şekil 3.13, Şekil 3.14 ve Şekil 3.15'de verilmiştir. Oluşturulan ANFIS yapılarının her girişi için 2 parametreli *gauss* türünde üyelik fonksiyonu kullanılmıştır. Dolayısıyla kullanılan ANFIS modelinde 2^n adet kural bulunmaktadır [161]. Bu şekilde birinci sistem için 24 başlangıç parametresi ve 448 sonuç parametresi toplamda 472 parametre ile 64 kuraldan oluşan ANFIS yapısı kullanılarak sistem modellenmiştir. İkinci sistem için 12 başlangıç parametresi ve 32 sonuç parametresi toplamda 44 parametre ile 8 kuraldan oluşan ANFIS yapısı kullanılmıştır. Son olarak üçüncü sistem için 16 başlangıç parametresi ve 80 sonuç

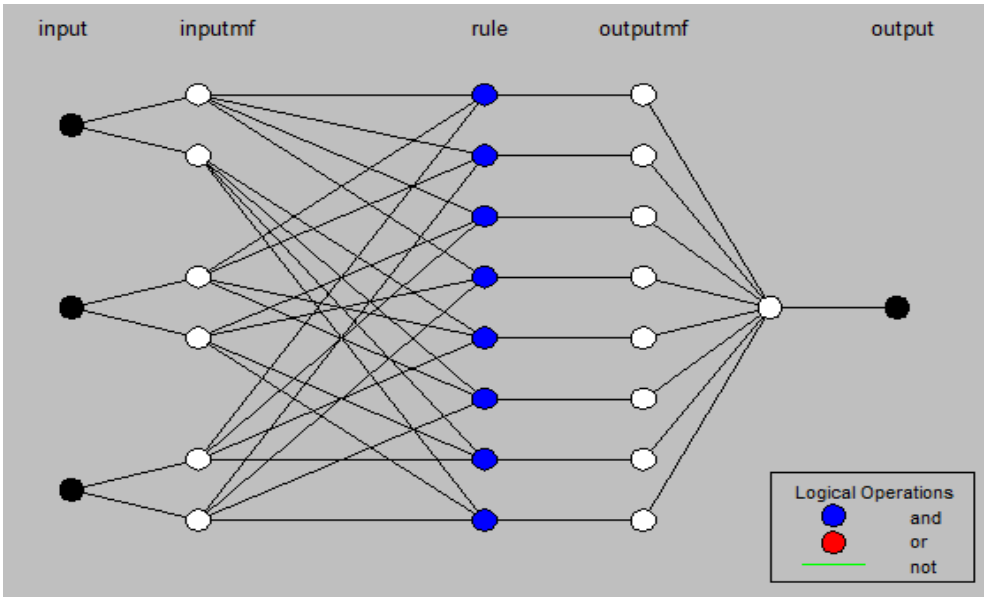
parametresi toplamda 96 parametre ile 16 kuraldan oluşan ANFIS yapısı kullanılarak sistem modellenmiştir.

ANFIS modelinin girişleri, dinamik sistemlerin *i.* dereceden önceki tüm giriş ve çıkış değerleridir. Her örnek için ANFIS 15 kez eğitilmiş ve her seferinde rastgele seçilen çözüm uzayı ile eğitime başlanmıştır. Ayrıca her örnek için elde edilen *RMSE* hata değerlerinin ortalaması alınarak, ortalama *RMSE* ($RMSE_{AVG}$) değeri bulunmuştur.

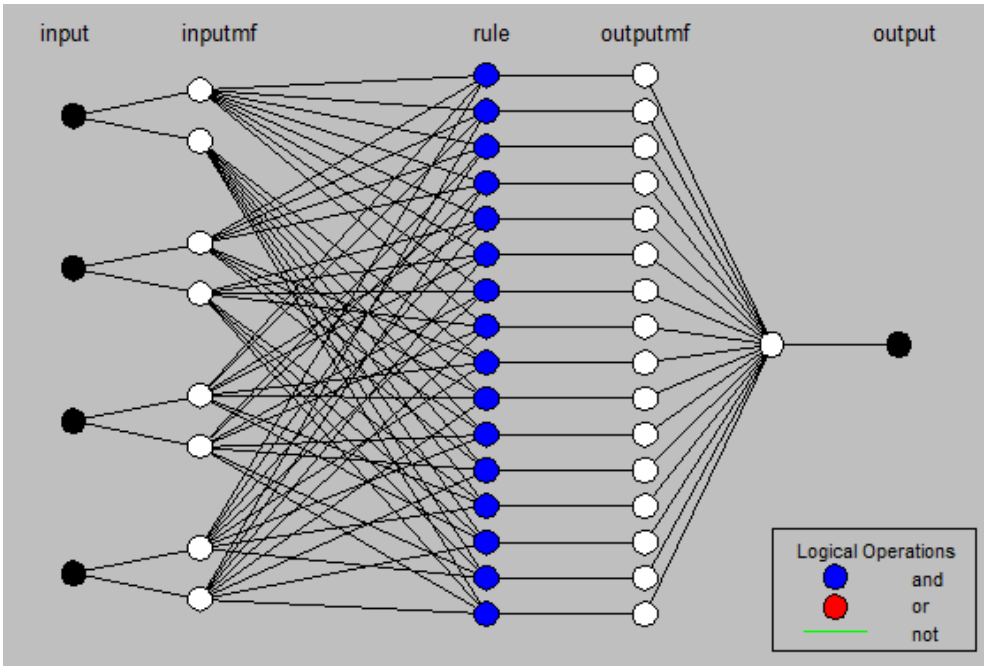
Çalışmada, SA algoritması için gerekli olan çeşitli parametre değerlerine karar vermek için birçok deneme gerçekleştirilmiştir. Bu denemeler sonucunda SA algoritması için sıcaklık noktalarının sayısı 10, her bir sıcaklık noktasında gerçekleştirilen iterasyon sayısı 8 ve sıcaklık azalma parametresi 0,2 olarak alınmıştır.



Şekil 3.13. Sistem 1 için 6 girişli ANFIS modeli



Şekil 3.14. Sistem 2 için 3 girişli ANFIS modeli



Şekil 3.15. Sistem 3 için 4 girişli ANFIS modeli

Önerilen yaklaşımın performansının farklı yöntemlerle karşılaştırılması amacıyla, ANFIS ağı dikkate alınan sistemlerin kimliklendirilmesi için popülasyon tabanlı GA ile de eğitilmiştir. Bu amaçla, kullanılan GA modelinde popülasyon büyüklüğü 200 çaprazlama oranı 0,8 ve mutasyon oranı 0,01 olarak seçilmiştir.

Ayrıca, önerilen yaklaşımın performansı dinamik sistemlerin kimliklendirilmesinde yaygın olarak kullanılan Elman geribeslemeli yapay sinir ağına ait literatürden alınan sonuçlarla da karşılaştırılmıştır [162,163]. Dikkate alınan çalışmalarda Kalınlı ve arkadaşları, Elman ağının dinamik sistemlerin kimliklendirilmesi amacıyla Tabu Arama Algoritması (TS), SA algoritması ve Paralel Tabu Arama Algoritması (PTS) kullanılarak eğitilmesine dayalı bir yaklaşım önermiş ve geleneksel BP algoritmasına göre daha başarılı sonuçlar elde etmiştir [162,163].

3.3.4.1. Dinamik Sistem 1

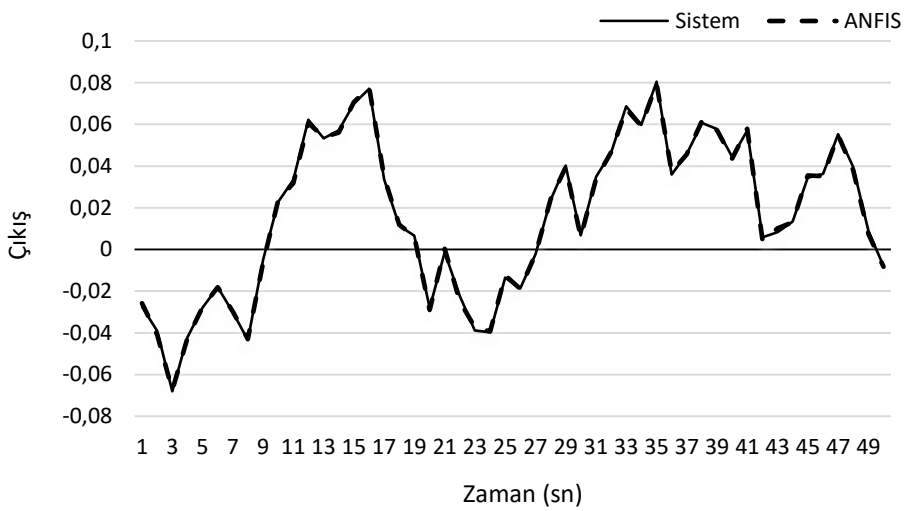
Simülasyon için kullanılan birinci örnek, Eşitlik 3.4 ile tanımlanan 3. dereceden doğrusal bir sistemdir [162].

$$y(k) = A_1 y(k-1) + A_2 y(k-2) + A_3 y(k-3) + B_1 u(k-1) + B_2 u(k-2) + B_3 u(k-3) \quad (3.4)$$

Bu eşitlikteki katsayılar $A_1 = 2.627771$, $A_2 = -2.333261$, $A_3 = 0.697676$, $B_1 = 0.017203$, $B_2 = -0.030862$ ve $B_3 = 0.014086$ olarak belirlenmiştir. Sistem 1 için simülasyon çalışmaları sonucunda elde edilen *RMSE* hata değerleri ve literatürden alınan farklı yöntemlere ait hata değerleri Tablo 3.8'de verilmiştir. Tablo 3.8'de verilen sonuçlardan bu çalışmada önerilen yaklaşımın performansının diğer yöntemlere göre oldukça yüksek başarıma sahip olduğu açıkça görülmektedir. ANFIS ağının GA ile eğitilmesi durumunda geleneksel BP algoritmasına göre hatanın % 29,50 iyileştiği, SA ile eğitilmesi durumunda ise hata değerinin % 61,22 daha düşük olduğu görülmektedir. ANFIS ağının SA algoritması kullanılarak eğitilmesi durumunda elde edilen hata değerinin ise Elman Ağı için literatürdeki en başarılı sonucu veren uygulamaya göre % 47,62 oranında daha başarılı olduğu da görülmektedir. ANFIS ağının SA algoritması ile eğitilmesine yönelik olarak gerçekleştirilen 15 farklı denemede elde edilen hata değerlerinin eğitim ve test işlemleri için standart sapması sırasıyla 5,4870E-04 ve 0,0012 olarak bulunmuştur. ANFIS ağının SA algoritması ile eğitilmesine yönelik gerçekleştirilen denemelerden birisine ait sistem ve ANFIS ağının cevapları Şekil 3.16'da verilmiştir.

Tablo 3.8. Doğrusal sistem için sonuçların karşılaştırılması

Model	Eğitim Hatası	Test Hatası
Elman - BP [162]	-	7,6754E-02
Elman - TS [162]	-	8,1664E-03
Elman - PTS [163]	-	6,4693E-03
Elman - SA [162]	-	4,9653E-03
ANFIS - BP	5,8892E-03	6,7061E-03
ANFIS - GA	3,9306E-03	4,7279E-03
ANFIS - SA	**2,1126E-03**	**2,6009E-03**



Şekil 3.16. Sistem 1 için, sistemin ve SA algoritması ile eğitilen ANFIS modelinin çıkışları (*RMSE* hata değeri = 1,0378E-09)

3.3.4.2. Dinamik Sistem 2

Simülasyon için kullanılan ikinci örnek, [162]'den alınan ve Eşitlik 3.5 ile tanımlanan küçük açılarla salınan basit bir sarkaça aittir. Sistemin ayrık zaman tanımı aşağıdaki gibidir:

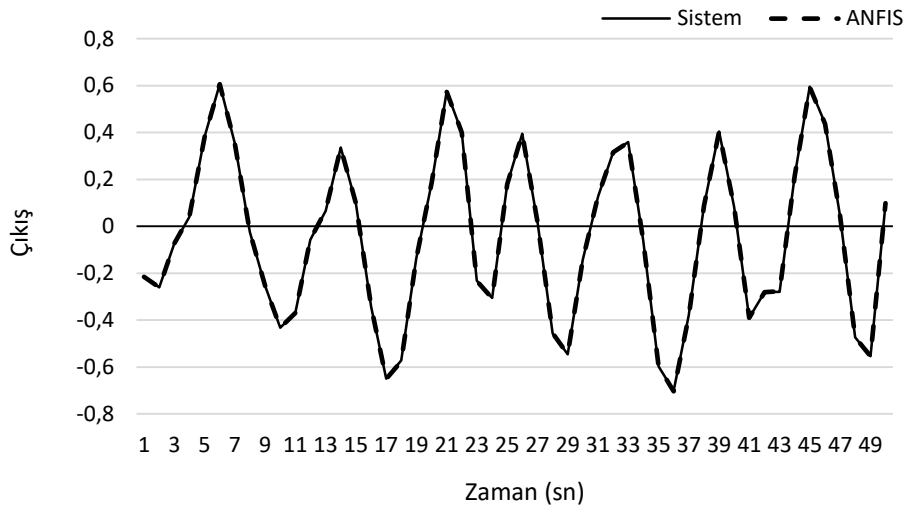
$$y(k) = A_1 y(k-1) + A_2 y(k-2) + A_3 y^3(k-2) + B_1 u(k-2) \tag{3.5}$$

Bu eşitlikteki katsayılar $A_1 = 1.040000$, $A_2 = -0.824000$, $A_3 = 0.130667$ ve $B_1 = -0.160000$ olarak belirlenmiştir.

Sistem 2 için ise simülasyon çalışmaları sonucunda elde edilen *RMSE* hata değerleri ve literatürden alınan farklı yöntemlere ait hata değerleri Tablo 3.9'da verilmiştir. Tablo 3.9'da verilen sonuçlardan bu çalışmada önerilen yaklaşımın performansının diğer yöntemlere göre oldukça yüksek başarıma sahip olduğu açıkça görülmektedir. ANFIS ağının GA ile eğitilmesi durumunda geleneksel BP algoritmasına göre hatanın % 13,38 iyileştiği, SA ile eğitilmesi durumunda ise hata değerinin % 65,00 daha düşük olduğu görülmektedir. ANFIS ağının SA algoritması kullanılarak eğitilmesi durumunda elde edilen hata değerinin ise Elman Ağı için literatürdeki en başarılı sonucu veren uygulamaya göre % 88,10 oranında daha başarılı olduğu da görülmektedir. ANFIS ağının SA algoritması ile eğitilmesine yönelik olarak gerçekleştirilen 15 farklı denemede elde edilen hata değerlerinin eğitim ve test işlemleri için standart sapması sırasıyla 0,0032 ve 0,0022 olarak bulunmuştur. ANFIS ağının SA algoritması ile eğitilmesine yönelik gerçekleştirilen denemelerden birisine ait sistem ve ANFIS ağının cevapları Şekil 3.17'de verilmiştir.

Tablo 3.9. Sistem 2 için sonuçların karşılaştırılması

Model	**Eğitim Hatası**	**Test Hatası**
Elman - BP [162]	-	2,6182E-01
Elman - TS [162]	-	1,1990E-01
Elman - SA [162]	-	2,9430E-02
Elman - PTS [163]	-	2,8318E-02
ANFIS - BP	7,9911E-03	9,6074E-03
ANFIS - GA	8,6912E-03	8,3221E-03
ANFIS - SA	**4,3542E-03**	**3,3691E-03**



Şekil 3.17. Sistem 2 için, sistemin ve SA algoritması ile eğitilen ANFIS modelinin çıkışları (*RMSE* hata değeri = 1,3216E-03)

3.3.4.3. Dinamik Sistem 3

Üçüncü model Eşitlik 3.6'da ayrık zaman ifadesi verilen ve Kaynak [164]' dan alınan doğrusal olmayan sistemdir.

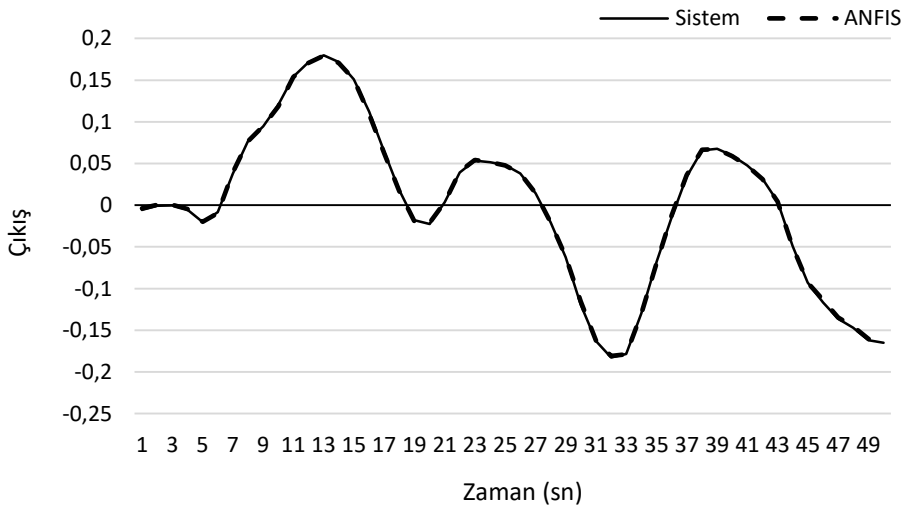
$$y(k) = \frac{A_1 y(k-1) + A_2 y(k-2) + B_1 u(k-1) + B_2 u(k-2)}{1 + y^2(k-2)} \tag{3.6}$$

Bu eşitlikteki katsayılar $A_1 = 1.752821$, $A_2 = -0.818731$, $B_1 = 0.011698$ ve $B_2 = 0.010942$ olarak belirlenmiştir. Sistem 3 için simülasyon çalışmaları sonucunda elde edilen *RMSE* hata değerleri ve literatürden alınan farklı yöntemlere ait hata değerleri Tablo 3.10'da verilmiştir. Tablo 3.10'da verilen sonuçlardan bu çalışmada önerilen yaklaşımın performansının diğer yöntemlere göre oldukça yüksek başarıma sahip olduğu açıkça görülmektedir. ANFIS ağının GA ile eğitilmesi durumunda geleneksel BP algoritmasına göre hatanın % 46,00 iyileştiği, SA ile eğitilmesi durumunda ise hata değerinin % 72,22 daha düşük olduğu görülmektedir. ANFIS ağının SA algoritması kullanılarak eğitilmesi durumunda elde edilen hata değerinin ise Elman Ağı için literatürdeki en başarılı sonucu veren uygulamaya göre % 79,81 oranında daha başarılı olduğu da görülmektedir. ANFIS ağının SA algoritması ile eğitilmesine yönelik olarak

gerçekleştirilen 15 farklı denemede elde edilen hata değerlerinin eğitim ve test işlemleri için standart sapması sırasıyla 0,0018 ve 0,0017 olarak bulunmuştur. ANFIS ağının SA algoritması ile eğitilmesine yönelik gerçekleştirilen denemelerden birisine ait sistem ve ANFIS ağının cevapları Şekil 3.18'de verilmiştir.

Tablo 3.10. Sistem 3 için sonuçların karşılaştırılması

Model	**Eğitim Hatası**	**Test Hatası**
Elman - BP [162]	-	1,2863E-01
Elman - TS [162]	-	6,1720E-02
Elman - SA [162]	-	2,1760E-02
ANFIS - BP	1,9711E-02	1,5819E-02
ANFIS - GA	9,5121E-03	8,5310E-03
ANFIS - SA	**4,2370E-03**	**4,3930E-03**



Şekil 3.18. Sistem 3 için, sistemin ve SA algoritması ile eğitilen ANFIS modelinin çıkışları (*RMSE* hata değeri = 2,1518E-03)

3.3.5. Genel Değerlendirme ve Sonuçlar

Bu bölümde, SA algoritmasının dinamik sistemlerin kimliklendirilmesi amacıyla ANFIS parametrelerinin eğitilmesi için kullanılmasına yönelik yeni bir yaklaşım önerilmiştir.

Doğrusal ve doğrusal olmayan dinamik sistemlerin kimliklendirilmesine yönelik yapılan simülasyon çalışmaları sonucunda, önerilen yaklaşımın dikkate alınan BP, TS ve GA algoritmalarından çok daha başarılı sonuçlar verdiği görülmüştür. Elde edilen sonuçlardan SA algoritmasının ANFIS ağının parametre değerlerinin belirlenmesi amacıyla başarılı bir şekilde kullanılabileceği kanaatine varılmıştır. Ayrıca, SA algoritmasının türeve dayalı algoritmalar gibi kısıtlamalar içermemesi ve problemlere kolaylıkla uygulanabilir olması nedeniyle, ANFIS ağının farklı problemlere yönelik uygulamalarında da kullanılabileceği değerlendirilmiştir.

Elde edilen sonuçların farklı yöntemlere ait sonuçlar ile karşılaştırılması neticesinde en iyi eğitim ve test hata değerlerinin önerilen yöntem ile bulunduğu görülmüştür. Yapılan çalışmalar sonucunda diğer yöntemlerden daha düşük etkin hata değerlerinin bulunması, doğrusal ve doğrusal olmayan sistemlerin kimliklendirilmesi için SA algoritması kullanılarak yapılan ANFIS eğitiminin oldukça etkili olduğunu göstermektedir. Ayrıca önerilen yöntem kullanılarak yapılan çalışmalarda düşük standart sapma değerine sahip sonuçlar elde edilmiştir. Standart sapma değerinin düşük olması, farklı başlangıç çözümleriyle başlanan eğitimler sonucunda elde edilen *RMSE* hata değerlerinin birbirine yakın, yani sonuçların güvenilir olduğunu göstermektedir.

4. BÖLÜM

BULANIK C-ORTALAMALAR TABANLI ANFIS AĞININ BENZETİLMİŞ TAVLAMA ALGORİTMASI İLE EĞİTİLMESİ

4.1. Mikrodizi Gen İfade Kanser Verilerinin Sınıflandırılması

4.1.1. Gen Ekspresyonu (İfade)

İnsan vücudundaki hücrelerin hepsi aynı genetik materyali içerse de her hücrede aynı genler etkin değildir. Hangi genin aktif olup hangisinin olmadığı bilgisi biyologlara, normalde bu hücrelerin nasıl işlediği ve bazı genler doğru çalışmadığı zaman hücrenin bundan nasıl etkileneceği bilgisini vermektedir. Bu aktiflik bilgisine hücrenin gen ifadesi denmektedir [165].

Her hücre aynı genetik bilgiyi içermektedir. Ancak, deri hücreleri, böbrek, ciğer, kan, beyin hücreleri birbirinden farklıdır. Bu farklılıklar genlerin farklı hücrelerde farklı farklı ifade edilmelerinden kaynaklanmaktadır [166].

Geçmişte, biyologlar birkaç genin gen ifade verisini aynı anda ölçebilirken, DNA (Deoksiribonükleik asit) mikroçip teknolojisinin gelişimiyle binlerce genin gen ifade verisi eşzamanlı olarak ölçülebilmeye başlanmıştır [165]. Böylelikle aynı dokudan alınan sağlıklı ve hastalıklı hücrelerin gen aktiviteleri birbirleriyle karşılaştırılabilmektedir.

İnsan genom projesiyle birlikte ortaya çıkarılan gen dizilerinden sonra, yeni amaç; bu genlerin nasıl ifade gösterdiklerini bulmak, yani mRNA profillerini çıkararak diğer genlerle beraber nasıl bir ilişki içinde olduklarını göstermek ve böylece belirli hastalıklarda hangi genlerin rol oynadığını ortaya çıkarmak olmuştur. Bir hücrenin tipi veya içinde bulunduğu evre o hücrenin mRNA ifadesi ile ilgilidir. Daha önceden tanımlanmamış genlerin ifade düzeyleri incelenerek ve diğer bilinen genlerin mRNA

ifadeleri ile karşılaştırılarak o genlerin işlevleri hakkında bilgiler edinilmeye çalışılmaktadır [166].

4.1.2. Mikrodizi Teknolojisi

Bilgisayar teknolojisinin moleküler biyolojiye paralel olarak hızla gelişmesi, iki disiplini birbirine yaklaştırmıştır. Böylece, biyoteknolojinin kavramsal olarak ulaşabileceği son noktalardan biri olan mikrodizi (gen çip) ortaya çıkmıştır. Mikrodizi tekniğinin ilk girişimleri Shalon ve Schena tarafından gerçekleştirilmiştir [167]. Mikrodizi teknolojisi, geleneksel yöntemlerin aksine aynı anda pek çok genin araştırılmasına olanak sağlayan yeni analitiksel metotlar sunmaktadır. Bu teknik özellikle tıp alanı olmakla birlikte; biyoloji, mikrobiyoloji, genetik gibi pek çok alanda da kullanılmaktadır. Farklı amaçlara uygun pek çok mikrodizi tipi mevcuttur [168]. DNA mikrodizi teknolojisi; binlerce genin ifade düzeylerini (ekspresyon seviyesi) ve DNA değişimlerini eşzamanlı olarak inceleme yöntemidir [169].

Moleküler biyolojideki geleneksel metotlarda genellikle “bir deneyde bir gen” ilkesi geçerlidir. Yani; aynı çalışmada gen fonksiyonlarının tamamını görmek geleneksel yöntemlerle zordur. Gen çip teknolojisi olarak da adlandırılan yeni yöntemler ile bütün genomun görüntülenmesi sağlanır ve bu metot aynı anda binlerce genin birbirleriyle olan etkileşimlerinin görülmesine olanak tanır. Binlerce genin tek bir çalışmada incelenmesi esasına dayanan mikrodizi teknolojilerinin ilk girişimleri 1990’ların başında Schena tarafından gerçekleştirilmiştir. Bu teknoloji tek seferde çok sayıda genin ekspresyon düzeyinin incelenmesine olanak tanıyan bir teknolojidir ve binlerce DNA’nın aynı anda analiz edilebilmesini sağlamaktadır [167].

DNA mikrodizi teknolojisi, birçok genin aktivitesinin aynı zamanda izlenebilmesi; hızlı bir yöntem olması; hasta ve sağlıklı hücrelerdeki genlerin aktivitelerinin karşılaştırılmasını sağlaması ve hastalıkları alt-gruplar halinde kategorize edebilme gibi avantajlarının yanı sıra, tek bir seferde çok fazla veri analizi yapıldığından, tüm sonuçların analizinin zaman alması; sonuçların yorumlamak için çok kompleks olabilmesi; sonuçların yeterince kantitatif olmaması ve oldukça pahalı bir teknoloji olması gibi bazı dezavantajlara da sahiptir [170].

Gen ekspresyonunun mikrodiziler kullanılarak ölçülmesi biyoloji ve tıbbın pek çok alanında uygulanabilirlik arz etmektedir. Örneğin mikrodiziler hastalıklı ve normal bir hücrede gen ekspresyonunun karşılaştırılması suretiyle hastalık ile alakalı genlerin belirlenmesinde kullanılabilmektedir. Diğer bir çalışmada beyin dokusundan alınan postmortem prefrontal kortekslere major depresyon için DNA mikrodizi analizi yapılmıştır. Major depresyon hastaları ve kontrol grubunun gen ekspresyon modelleri karşılaştırılmış ve major depresyonda 99 genin farklı şekilde eksprese olduğu görülmüştür [171].

DNA mikrodizi (gen çipi, DNA çipi ya da biyoçip olarak da bilinir) ekspresyon profili oluşturmak, başka bir ifade ile binlerce genin ekspresyon düzeyini aynı anda izlemek amacı ile cam, plastik veya silikon çip gibi katı bir yüzeye tutturularak sıralı bir şekilde (**array**) oluşturulmuş mikroskobik DNA spotlarıdır [171]. Yöntem ve ortaya çıkış metotlarında bazı küçük farklar olmasına rağmen “DNA array”, “DNA çip” ve “mikroçip” gibi tanımlamalar da benzer uygulamaları ifade etmek için kullanılmaktadır [172].

Sağlık alanında analiz ve görüntüleme teknolojilerdeki gelişmeler nedeniyle hemen her alanda üretilen verilerin boyutunda önemli artışlar ortaya çıkmış ve büyük boyutlu veriler üzerinde veri madenciliği uygulamaları geliştirebilmek için etkili yöntemlere olan ihtiyaç da artmış bulunmaktadır. Genetik araştırmalar kapsamında ortaya çıkan laboratuvar verileri de genellikle büyük boyutlu olup, uygulamada bu verilerin sınıflandırılması uygulamalarında önemli güçlükler bulunmaktadır.

DNA mikrodizi teknolojisi içerisindeki spotlar sayesinde yüzlerce geni bünyesinde barındıran bir gen çipidir [173]. Bu teknoloji sayesinde aynı dokudan alınan sağlıklı ve hastalıklı hücrelerin gen aktiviteleri birbirleriyle karşılaştırılabilmektedir. Böylece kanser gibi hastalıklara neden olan genlerin bulunmasında DNA mikrodizi teknolojisi araştırmacılara çok büyük kolaylıklar sunmaktadır. Bunun yanı sıra genlerle ilişkili kanser hastalıklarının alt türlerini bulmada da araştırmacılara destek olmaktadır [171].

DNA mikrodizi gen ifade (ekspresyon) veri analizi, hastalıklarla ilişkili genleri tanımlayarak herhangi bir bireyin hasta veya sağlıklı olduğunu yüksek bir olasılıkla hesaplayabildiğinden dolayı kanser gibi hastalıkların ortaya çıkmasına neden olan

genlerin belirlenmesinde önemli bir rol oynamaktadır. Bu yüzden mikrodizi gen ifade (ekspresyon) problemlerini çözmek için yüksek performansa sahip sınıflandırma yöntemleri büyük önem teşkil etmektedir.

İnsan genomunda yaklaşık olarak 40.000 genin bulunduğu tahmin edildiğine göre bu kadar geni tek tek analiz etmek mümkün değildir. Fakat gelişen yazılım araçları sayesinde çok fazla sayıdaki genin analizi mümkün hale gelmiştir. Bu analizlerin büyük bir kısmını genleri sınıflandırma ve bu genler içerisinden önemli genleri seçme işlemi oluşturmaktadır [174]. Gen ifade verilerinden hastalıklara sebep olan genlerin seçilmesi ve bu seçilen genler ile hastalıkların sınıflandırılması başlıca veri madenciliği çalışmalarındandır.

4.1.3. Mikrodizi Gen İfade Kanser Verilerini Sınıflandırmak için SA Algoritması Kullanılarak FCM Tabanlı ANFIS Ağının Eğitilmesi

Mikrodizi gen ifade (ekspresyon) verilerindeki gen sayılarının çokluğu ve bu veriler arasında çoğunlukla doğrusal olmayan ilişkiler bulunması gibi nedenlerden dolayı geleneksel yöntemler mikrodizi gen problemlerinin çözümlerinde yetersiz ve sınırlı kalabilmektedir. Bu zorluklar araştırmacıları daha modern ve güçlü yöntemler kullanmaya sevk etmiştir. Bu amaçla, zor problemlerin çözümünde araştırmacıların ilgisi giderek artan bir oranda yapay zekâ yöntemleri üzerine yoğunlaşmaya başlamıştır [175].

Son birkaç on yıldır büyük boyutlu problemlerin çözümünde yapay zekâ tabanlı yöntemlerin kullanımına olan ilgi giderek artmış ve sınıflandırma ve öznitelik seçme problemlerine yönelik birçok uygulama gerçekleştirilmiştir. Pirooznia ve arkadaşları [176], mikrodizi gene ifade (ekspresyon) verileri üzerinde sınıflandırma çalışması yapmışlardır. Bu çalışmada farklı mikrodizi kanser veri setleri üzerinde istatistiksel yöntemlerden Destek Vektör Makineleri (SVM), Bayesian, J48 Karar Ağacı, Random Forest, Id3 ve Bagging ve yapay zekâ tabanlı yöntemlerden ise Çok katmanlı Algılayıcılar ve Radyal tabanlı yapay sinir ağlarını (RTYSA) kullanmışlardır. Elde edilen sonuçlar neticesinde ÇKA, RTYSA ve SVM' nin daha iyi performans sergilediği görülmüştür. Sarhan [177], mikrodizi kanser verileri üzerinde Discrete Cosine transforn (DCT) ve Yapay Sinir Ağı yöntemlerinin başarımlarını karşılaştırmıştır. Loganathan ve Girijia [178], ANFIS parametrelerini Runge Kutta Learning algoritması ile eğitmişler ve

modelin başarımını mikrodizi gen ekspresyon kanser verilerini sınıflandırarak ölçmüşlerdir. Kumar ve Punithavalli [179], ANFIS ile kanser verilerini sınıflandırmış ve elde edilen sonuçları istatistiksel yöntemler ile karşılaştırmışlardır.

Yapay zekânın bir dalı olan ANFIS, örneklerden öğrenebilme ve öğrenmiş olduğu bilgiyi belirli kurallara bağlı olarak ağ üzerinde dağıtabilme yeteneği gibi esnek ve güçlü yapısı sayesinde makinalı öğrenme, sınıflandırma ve optimizasyon gibi alanlarda başarı ile kullanılmaktadır. ANFIS, bulanık sistemlerin ve sinirsel ağların bilinen en iyi özelliklerini birleştiren ağ tabanlı bir yapıdır. Bulanık sistemlerin, giriş bilgisini ağ üzerine dağıtarak araştırma uzayını azaltması, sinirsel ağların ise geri yayılım davranışlarını kullanarak kontrol parametrelerinin doğrusal olmayan problemler için en uygun değerlerinin bulunması ANFIS modeline önemli üstünlükler kazandırmaktadır.

ANFIS modelinin eğitilmesi, temelde modelin başlangıç ve sonuç parametrelerine ait en uygun değerlerin belirlenmesi sürecidir. Bu kapsamda, ANFIS'in eğitimi için BP, LSE, GD, LM ve KF gibi türeve dayalı algoritmalar yaygın olarak kullanılmaktadır. Ancak, türeve dayalı algoritmalarda her adımda eğim hesaplanması nispeten zordur ve kullanılması gereken zincir kuralı çoğunlukla yerel minimuma takılma problemine yol açmaktadır. Ayrıca türeve dayalı algoritmalarda parametrelerin yakınsaması oldukça yavaş ve çözümün başarımı büyük oranda başlangıç değerlerine bağlı olmaktadır. Bu nedenle türeve dayalı geleneksel algoritmalar ile ANFIS'in parametrelerinin eğitilmesi ve güncellenmesi temel sorunlardan biridir. Bu nedenle, ANFIS'in başarılı bir şekilde eğitilmesi ve parametrelerin güncellenmesi konusunda farklı yöntemlerin uygulanmasına ihtiyaç olup, bu konu araştırmacıların ilgisini çekmiştir. Bu kapsamda, son yıllarda çeşitli yöntemler önerilmiştir. Bu yöntemlerden bazıları yapay zekâ optimizasyon algoritmalarından olan GA, PSO, ABC ve ACO gibi sezgisel algoritmalardır [2].

GA, PSO, ACO ve ABC gibi algoritmalar popülasyon tabanlıdır. Popülasyon temelli algoritmaların çözümlerin bir seti ile çalışması genellikle onların küresel optimumun bulunduğu bölgeye hızla erişmesine olanak sağlar. Ancak, bu algoritmaların yüksek oranda olasılık tabanlı arama stratejilerine sahip olması nedeniyle bulundukları bölgedeki optimum çözümü bulabilmeleri genellikle uzun süreler gerektirebilmektedir. SA gibi algoritmalar ise iteratif tabanlıdır ve genellikle rastgele seçilen bir çözüm adayını iyileştirmek suretiyle küresel optimumu bulmaya çalışırlar. Bu tür algoritmalar genel

olarak bulundukları bölgenin optimumunu popülasyon tabanlı algoritmalara göre daha kısa sürelerde bulabilmektedirler.

Tez çalışmasının bu bölümünde, mikrodizi gen ifade kanser verilerini sınıflandırmak için iteratif tabanlı bir algoritma olan SA algoritması kullanılarak ANFIS ağının eğitilmesine yönelik yeni bir yaklaşım önerilmiştir. Önerilen yaklaşımın performansı 5 farklı kanser türü için test edilerek, farklı yapay zekâ algoritmaları ve istatistiksel yöntemlerin performanslarıyla karşılaştırılmıştır. Çalışma kapsamında, mikrodizi kanser verilerini başarılı bir şekilde sınıflandırmak için BP, HB, GA, PSO ve SA Algoritmaları kullanılarak ANFIS ağı eğitilmiştir. Ayrıca istatistiksel yöntemlerden Bayes Ağları (BayesNet), Destek Vektör Makineleri (SVM) ve J48 Karar Ağacı (J48) yöntemlerini kullanarak da mikrodizi gen ifade kanser verilerini sınıflandırma çalışmaları yapılmış ve yapılan simülasyon çalışmaları neticesinde elde edilen sonuçlar performans analizlerinde kullanılmıştır. Elde edilen sonuçlardan SA algoritması ile ANFIS ağını eğiterek sınıflandırma yönteminin diğer yöntemlerden daha başarılı olduğu görülmüştür.

4.1.4. Deneysel Veri Setleri

Bu tez çalışması kapsamında veri kaynağı olarak 5 adet mikrodizi gen ifade veri seti kullanılmıştır. Mikrodizi kanser verilerini sınıflandırmada yaygın bir şekilde kullanılan bu veri setleri Rutgers Üniversitesinin Biyoinformatik Laboratuvarlarında bulunan veri tabanlarından elde edilmiştir [180].

4.1.4.1. Akciğer Kanseri (*Gordon-2002*)

Gordon ve arkadaşları [181] tarafından 2002 yılında düzenlenen bu veri seti 12533 gen ifadesinden oluşmaktadır.Bu veri setinin 31 adeti Malignant Pleural Mesothelioma (MPM) ve 150 adeti Adenocarcinoma (AD) olmak üzere iki tip akciğer kanserini kapsamaktadır. Veri önişleme sonrasında geriye 1626 gen ifadesi kalmıştır. Bu veri setinin 1626 gen ifade ölçümlerinin kaynağı genel kullanıma açıktır [182].

4.1.4.2. Beyin Kanseri (*Nutt-2003*)

Nutt ve arkadaşları [183] tarafından 2003 yılında düzenlenen beyin kanseri veri seti, beyin dokusundan alınan ve bu kanserin alt türleri olan Classic Glioblastomas (CG) ve Nonclassic Glioblastomas (NG) türlerine ait olan verileri bulundurur. Toplamda 28 adet örnek bulunduran veri setinde örneklerin 14 tanesi CG türüne ve 14 tanesi ise NG türüne aittir. Toplamda beyin dokularından 28 örnek alınmış ve bu örneklerin her birinden 12625 gen ifadesi çıkarılmıştır. Fakat bu genlerin içerisinde hastalığa etkisi olmayan veya hastalığa etkisi önemsenmeyecek kadar az olan genler elenmiş ve geriye sadece bu genlerden 1070 tanesi bu veri setini temsil etmek için kullanılmıştır [184].

4.1.4.3. Merkezi Sinir Sistemi Kanseri (*Pomeroy-2002*)

25 adet Classic Medulloblastomas (CMD) ve 9 adet Desmoplastic Medulloblastomas (DMD) olmak üzere toplamda 34 örnek bulunduran iki sınıflı bu kanser verisi Pomeroy ve arkadaşları [185] tarafından düzenlenmiştir. Orijinal veri seti 7129 gen ifadesinden oluşmaktadır. Yapılan öznitelik azaltma çalışmaları neticesinde hastalıkta etkisi olduğu düşünülen 857 geni içeren veri seti genel kullanıma açılmıştır [186].

4.1.4.4. Rahim Kanseri (*Risinger -2003*)

Risinger ve arkadaşları [187] tarafından 2003 yılında düzenlenen rahim kanseri veri seti, hastalıklı ve normal kişilerin rahim dokularından alınmış 42 adet örnekten oluşmaktadır. Bu veri setindeki her örnek 8872 mikrodizi gen ifadesi ile temsil edilmektedir. Sınıflandırma yöntemlerinin performansını arttırmak için bu gen diziliminden bazı genler elenmiş ve geriye 1771 mikrodizi gen ifadesi kalmıştır. Bu veri seti, 13 adet Serous Papillary (SP), 3 adet Clear Cell (CC), 19 adet Endometrioid Kanseri (E) ve 7 adet Normal hasta (N) olmak üzere 42 örnek ve 4 sınıf etiketinden oluşmaktadır. Bu veri setinin 1771 gen ifade ölçümlerinin kaynağı genel kullanıma açıktır [188].

4.1.4.5. Prostat Kanseri (*Singh-2003*)

2002 yılında, klinik ortamda Singh ve arkadaşları [189] tarafından çalışılan veri seti, sağlıklı ve hastalıklı (tümörlü) dokulardan alınan hücrelerin gen ifadelerinden oluşmaktadır. Bu veri seti, 50 adet sağlıklı prostat dokusu ve 52 adet hastalıklı prostat

dokusunun 12600 gen ifadesini bulundurmaktadır. Veri seti, Normal (N) ve Prostat kanseri (PR) şeklinde iki farklı sınıf etiketi içermektedir. Prostat kanseri veri setindeki her örnek yapılan veri ön işleme çalışmaları sonucunda 339 gen dizilimi ile ifade edilirken toplamda 102 örneğin mikrodizi gen ifade profili genel kullanıma açılmıştır [190]. Tablo 4.1'de farklı dokulardan alınan örnekler ve kanser türleri özetlenmiştir.

Tablo 4.1. Çalışmada kullanılan veri setleri

Veri Seti	**Doku**	**Sınıf**	**Gen**	**Örnek**
Gordon-2002	Akciğer	2	1626	181
Nutt-2003	Beyin	2	1070	28
Pomeroy-2002	Beyin	2	857	34
Risinger-2003	Rahim	4	1771	42
Singh-2002	Prostat	2	339	102

Mikrodizi veri analizlerinin maliyeti çok yüksek olduğundan dolayı bu analizler çok fazla örnek üzerinde yapılamamaktadır. Bu nedenle mikrodizi gen ifade düzeyine sahip çok az örnek bulunmaktadır. En az örnek beyin kanserinde bulunurken en çok ise akciğer kanserinde bulunmaktadır. Mikrodizi veri setlerinde örnek sayısı az olmasından dolayı sınıflandırma yöntemlerinin performansları da olumsuz etkilenebilmektedir.

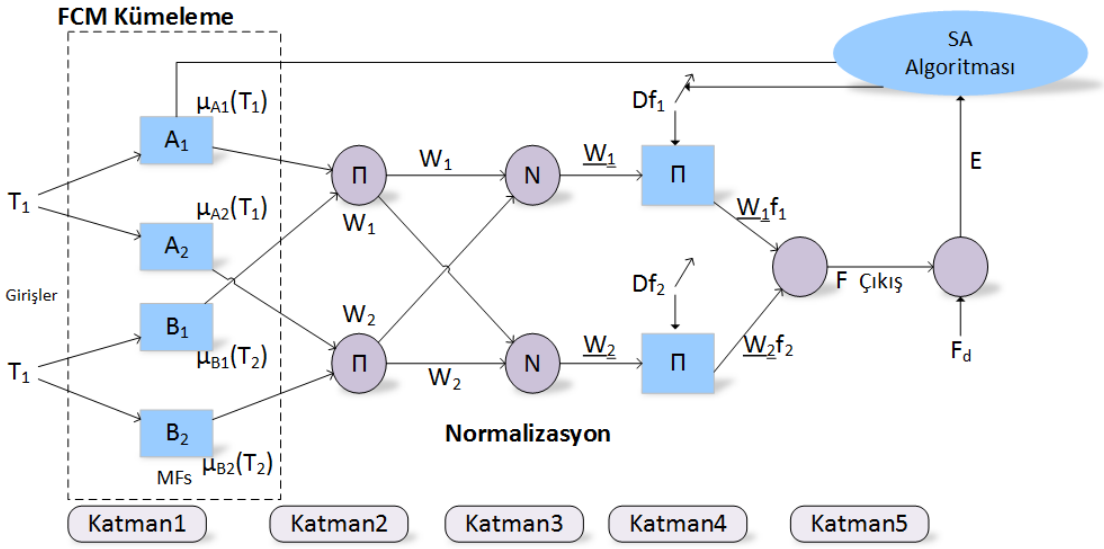
4.1.5. Deneysel Çalışmalar ve Analiz

Bu çalışmada mikrodizi gen ifade (ekspresyon) kanser veri setleri kullanılarak sınıflandırma yapılmıştır. Tüm veri setleri, sınıflandırma işleminden önce [0-1] aralığına normalize edilmiş ve veri setlerine öznitelik seçme yöntemi uygulanmıştır. Birçok öznitelik seçme yöntemi incelenmiş ve başarısından dolayı veri setlerinin bir alt setini oluşturmak için Korelasyon tabanlı öznitelik seçimi kullanılmıştır. Öznitelik seçme çalışmalarından sonra bu genlerin içerisinde hastalığa etkisi olmayan veya hastalığa etkisi önemsenmeyecek kadar az olan genler elenmiştir. Böylece yüksek boyutlu veri setleri yerine bu veri setlerini en iyi şekilde ifade eden alt veri setleri ile çalışarak, sınıflandırma yöntemlerinin performanslarını arttırmak ve sınıflandırmada daha başarılı sonuçlar almak mümkün hale gelmiştir. Boyutları azaltılan veri setlerinin referans alındığı kaynaklar ve gen sayıları Tablo 4.2'de verilmiştir.

Tablo 4.2. Veri setleri ve gen sayıları

Veri Seti	Kaynak	Gen Sayısı
Akciğer Kanseri *(Gordon-2002)*	[181]	92
Beyin Kanseri *(Nutt-2003)*	[183]	36
Merkezi Sinir Sistemi Kanseri *(Pomeroy-2002)*	[185]	30
Rahim Kanseri *(Risinger-2003)*	[187]	53
Prostat Kanseri *(Singh-2002)*	[189]	24

5 farklı mikrodizi gen ifade kanser veri setlerini sınıflandırmak için SA algoritması kullanılarak ANFIS modelinin eğitilmesine yönelik uygulamalar yapılmıştır. Veri setleri üzerinde yöntemlerin başarımlarını ölçmek için verilerin % 70'i eğitim, % 30'u ise test için kullanılmıştır. Yapılan simülasyon çalışmalarında kullanılan modelin oluşturulmasını ve oluşturulan modelin parametrelerinin optimize edilmesini gösteren blok diyagram Şekil 4.1'de verilmiştir. Kullanılan veri setlerindeki parametre sayılarının çokluğundan dolayı, bulanık kural ve parametre sayılarını azaltarak daha performanslı ANFIS modelleri oluşturabilmek için bulanıklaştırma katmanında ANFIS ile FCM kümeleme yönteminin entegre olduğu bir bulanık kural oluşturma tekniği kullanılmıştır. FCM uygulanarak oluşan ANFIS modellerindeki giriş sayısı her bir veri setindeki gen sayısına eşit olup her bir giriş için üyelik fonksiyonu türü, üyelik fonksiyonu sayısı ve bulanık kural sayıları Tablo 4.3'de verilmiştir.



Şekil 4.1. FCM ile modellenmiş temel ANFIS mimarisi

Tablo 4.3. Optimize edilen ANFIS modellerine ait bilgiler

Veri Seti	**Giriş Sayısı**	**MFs Türü**	**MFs Sayısı**	**Kural Sayısı**
Akciğer Kanseri *(Gordon-2002)*	92	*gaussmf*	10	10
Beyin Kanseri *(Nutt-2003)*	36	*gaussmf*	20	20
Merkezi Sinir Sistemi Kanseri *(Pomeroy-2002)*	30	*gaussmf*	24	24
Rahim Kanseri *(Risinger-2003)*	53	*gaussmf*	29	29
Prostat Kanseri *(Singh-2002)*	24	*gaussmf*	10	10

ANFIS'in eğitilmesi gereken iki parametre türü vardır. Bunlar başlangıç parametreleri ve sonuç parametreleridir. Başlangıç parametreleri Eşitlik 2.5'de $\{a_i, c_i\}$ olarak verilen *gauss* üyelik fonksiyonlarına aittir. Buradaki a_i üyelik fonksiyonlarının varyansı, c_i üyelik fonksiyonlarının merkezidir. Başlangıç parametrelerinin toplam sayısı tüm üyelik fonksiyonlarındaki parametrelerinin toplamına eşittir. Sonuç parametreleri ise durulaştırma katmanında kullanılan Eşitlik 2.8'de $\{p_i, q_i, r_i\}$ olarak gösterilen parametrelerdir. Tasarlanan ANFIS modelinin tüm parametrelerinin optimize edilmesi için SA algoritması kullanılmıştır.

Şekil 4.1'de verilen ANFIS modelinin 1. ve 4. katmanındaki başlangıç ve sonuç parametrelerinin başlangıç değerleri için FCM kümeleme yöntemi ile bir çözüm uzayı oluşturulmuş ve sonrasında parametreler her bir iterasyonda SA algoritması kullanılarak eğitilmiştir. Eşitlik 4.1 ile elde edilen *RMSE* hata fonksiyonu, sistemin hata değerinin belirlenmesinde kullanılmıştır. Bu hata fonksiyonunda kullanılan F ve F_d sırasıyla ANFIS tarafından elde edilen çıkış ve veri setinin gerçek çıkışıdır. N ise uygulamada kullanılan örnek sayısını göstermektedir.

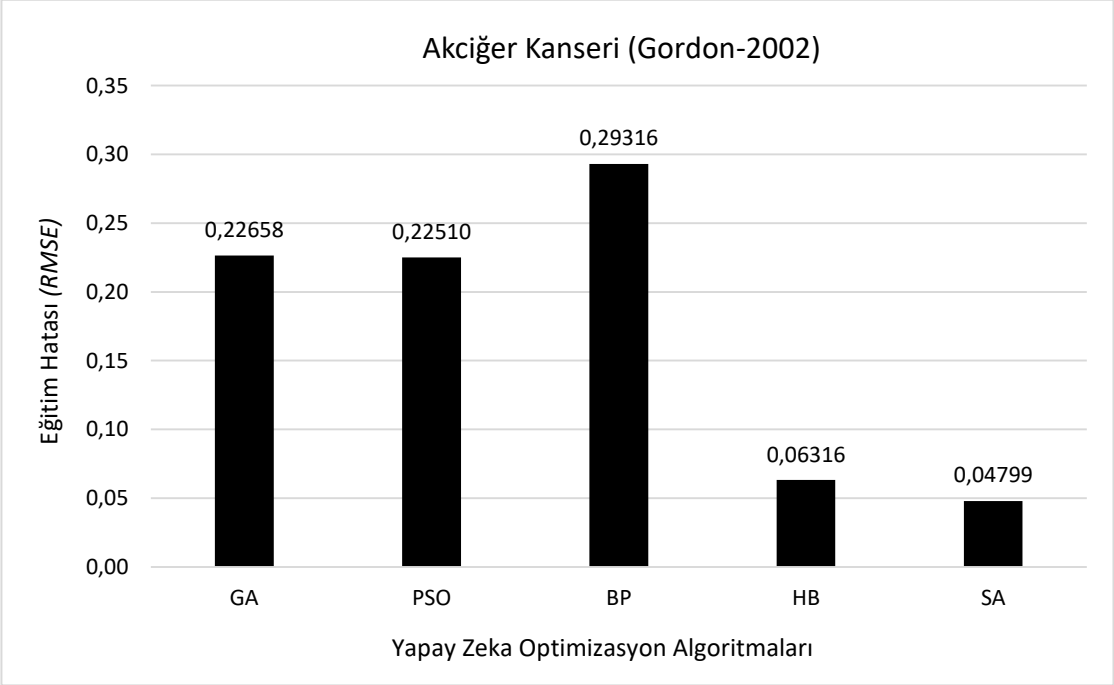
$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(F(i) - F_d(i)\right)^2}{N}} \tag{4.1}$$

Yapay zekâ optimizasyon algoritmalarının performansı algoritmaların kontrol parametrelerine oldukça bağlıdır ve genellikle kontrol parametreleri için hangi değerlerin kullanılması gerektiğine yönelik kesin kural ve yöntemler bulunmamaktadır. Bu kapsamda, parametreler için araştırmacıların yaygın olarak önerdiği aralıklarda değerler

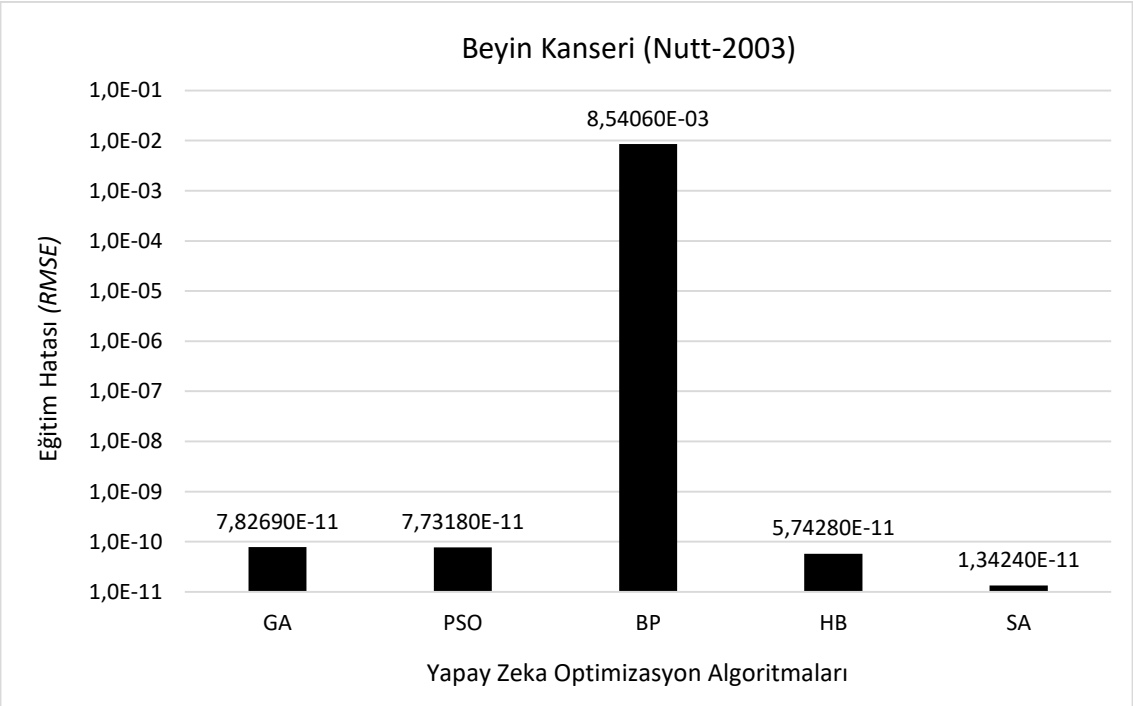
tayin edilmesi veya çok sayıda deneme yapılarak optimal parametre değerlerinin belirlenmesi yaklaşımı kabul görmektedir. Çalışmada, SA algoritmasının parametre değerlerine karar vermek için birçok deneme gerçekleştirilmiş ve denemeler sonucunda sıcaklık noktalarının sayısı 10, her bir sıcaklık noktasında gerçekleştirilen iterasyon sayısı 2 ve sıcaklık azalma parametresi 0,9 olarak alınmıştır.

Önerilen yaklaşımın performansının farklı yöntemlerle karşılaştırılması için ANFIS ağı dikkate alınan problemlerin sınıflandırılması amacıyla GA, PSO, BP ve HB algoritmaları ile de eğitilmiştir. Bu amaçla, GA modeli için iterasyon sayısı 100, popülasyon büyüklüğü 25, çaprazlama oranı 0,8 ve mutasyon oranı 0,01 olarak seçilmiştir. PSO algoritması için iterasyon sayısı 100 ve popülasyon büyüklüğü olarak belirlenmiştir. Kullanılan BP algoritması için öğrenme katsayısı 0,2 ve momentum oranı 0,4 olarak seçilmiştir. Hibrid algoritma olarak da en küçük kareler tahmini ve BP algoritmasının birleştirilmesin oluşan bir yöntem kullanılmıştır. Ayrıca bu iki algoritma için iterasyon sayısı 100 olarak belirlenmiştir.

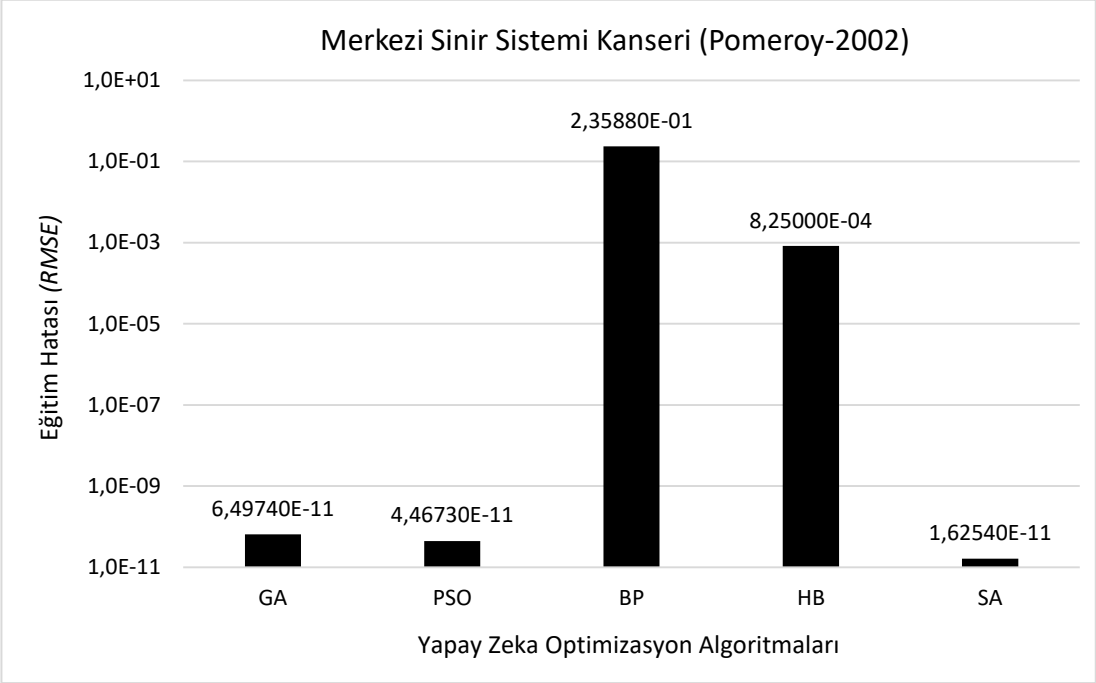
Sonuçların doğruluğunu ve güvenilirliğini göstermek için her algoritma ile 15 farklı simülasyon çalışması yapılmıştır. Ayrıca her model için elde edilen *RMSE* hata değerlerinin ortalaması alınarak, ortalama *RMSE* ($RMSE_{AVG}$) değeri bulunmuştur. Ayrıca, 5 farklı mikrodizi gen ifade kanser verisinin sınıflandırılmasında kullanılan ANFIS ağını optimize eden algoritmalar ile elde edilen eğitim hata değerleri (*RMSE*) Şekil 4.2 – 4.6'da verilmiştir. Şekil 4.7 - 4.11'de ise ANFIS modellerine giriş olarak verilen test veri setinin gerçek çıkışı ile ANFIS tarafından elde edilen çıkış arasındaki farkı gösteren test hata değerleri (*RMSE*) verilmiştir. Verilen grafiklerden de görüldüğü üzere en küçük eğitim ve test hataları ANFIS ağının SA algoritması ile optimize edilmesi sonucunda elde edilmiştir. Sadece prostat kanseri için Şekil 4.6 incelendiğinde elde edilen en küçük eğitim hata değeri az farkla HB algoritmasına aittir. Ancak, Şekil 4.11'e bakıldığında prostat kanseri verisi için elde edilen en küçük test hata değeri yine SA algoritması ile elde edilmiştir. Eğitim ve test hata değerleri arasındaki farkın çok olması aşırı öğrenme olduğunu ve eğitilen modelin uygulanabilir olmadığını ortaya koymaktadır. Eğitim ve test hata değerlerinin ise birbirine yakın olması daha güçlü ve başarımı yüksek modeller ortaya koymaktadır. Bu sebeple yapılan çalışmalarda SA algoritmasının eğitim ve test hata değerlerinin de birbirine yakın olması önerilen yöntemin uygulanabilir ve güvenilir olduğunu göstermektedir.



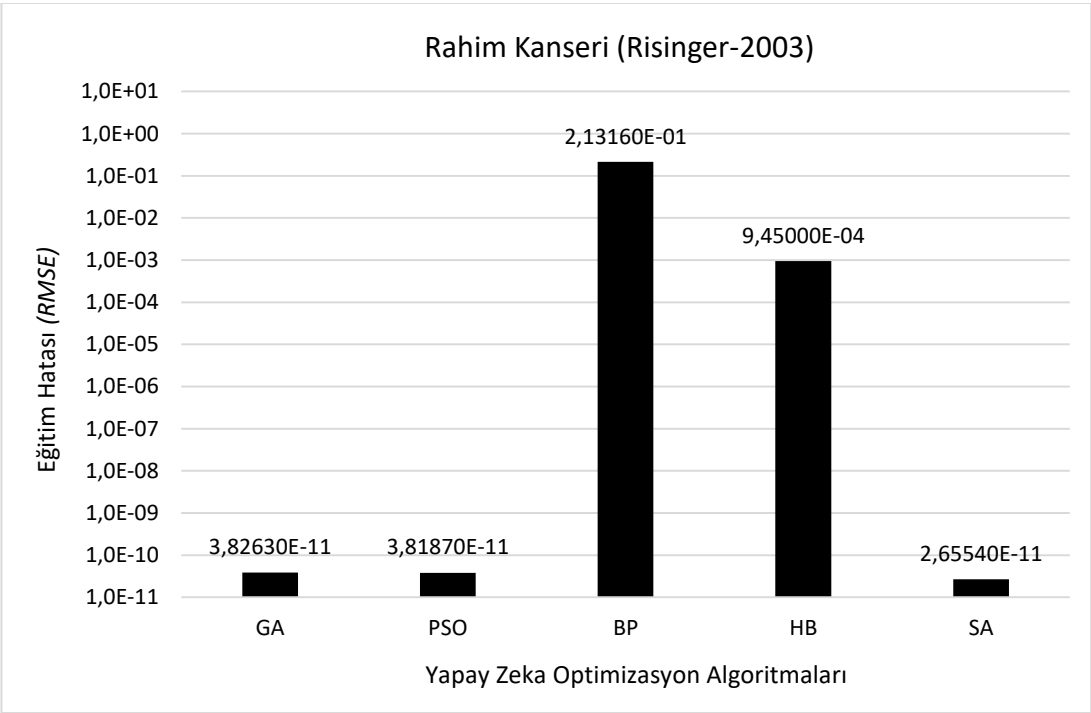
Şekil 4.2. Akciğer kanseri verisinin sınıflandırılması ile elde edilen eğitim (*RMSE*) hata değerleri



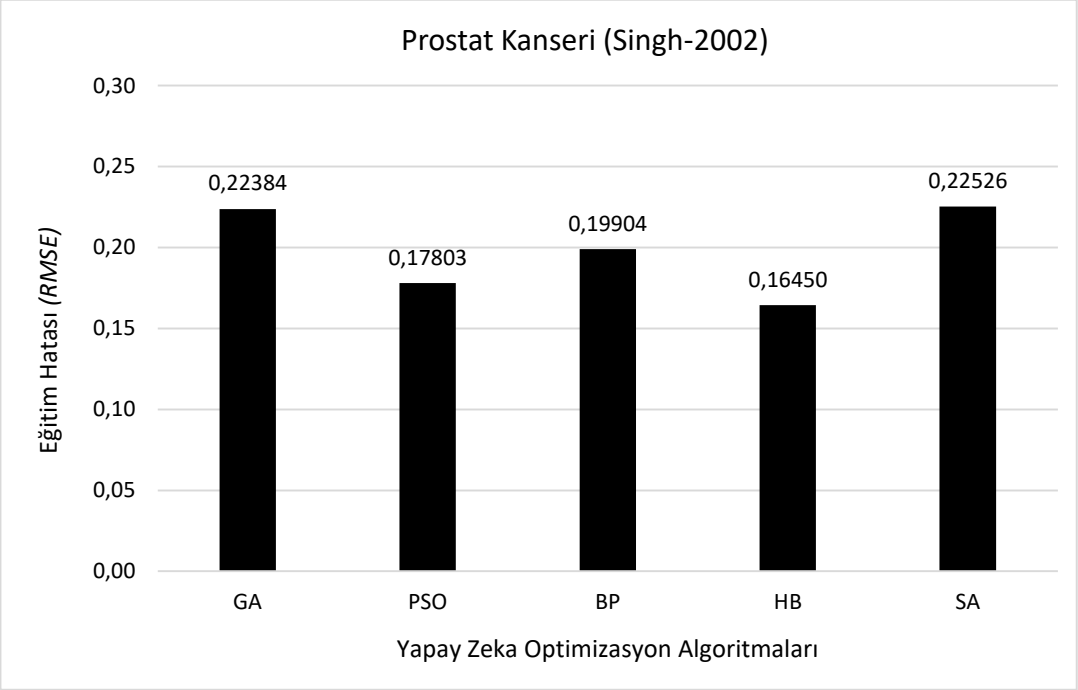
Şekil 4.3. Beyin kanseri verisinin sınıflandırılması ile elde edilen eğitim (*RMSE*) hata değerleri



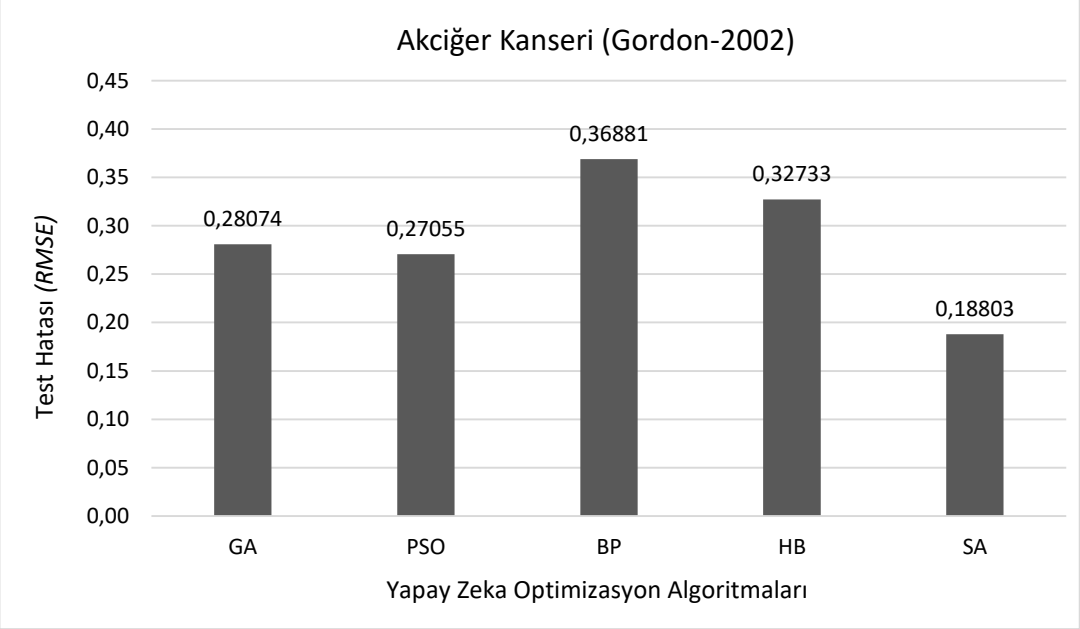
Şekil 4.4. Merkezi Sinir Sistemi kanseri verisinin sınıflandırılması ile elde edilen eğitim (*RMSE*) hata değerleri



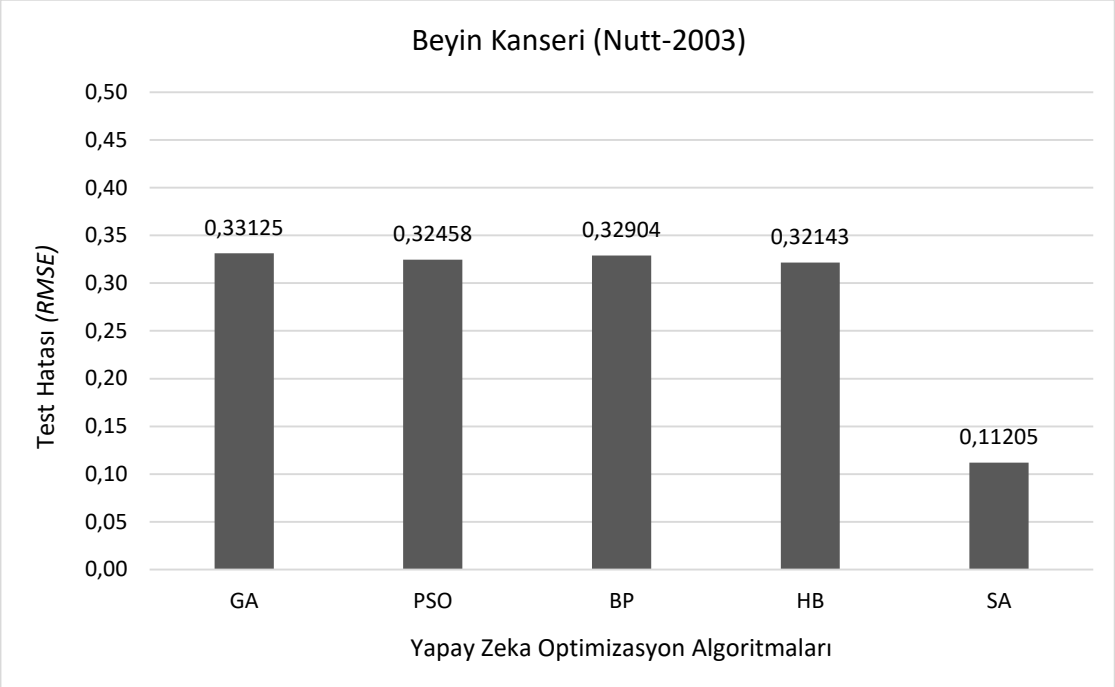
Şekil 4.5. Rahim kanseri verisinin sınıflandırılması ile elde edilen eğitim (*RMSE*) hata değerleri



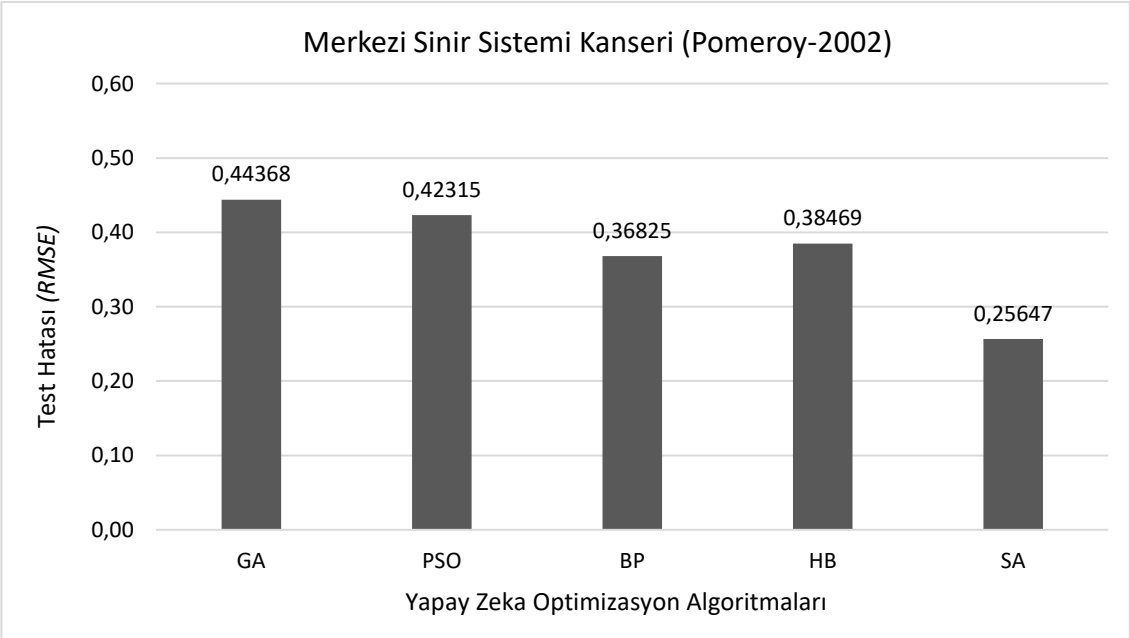
Şekil 4.6. Prostat kanseri verisinin sınıflandırılmasında elde edilen eğitim (*RMSE*) hata değerleri



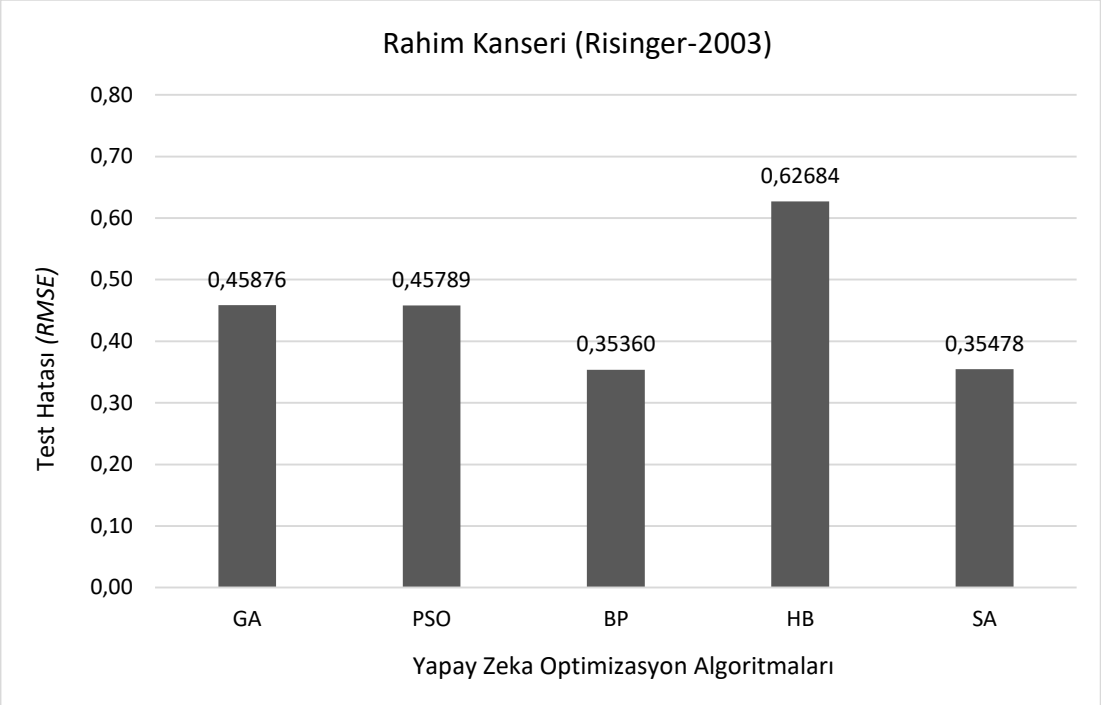
Şekil 4.7. Akciğer kanseri verisinin sınıflandırılmasında elde edilen test (*RMSE*) hata değerleri



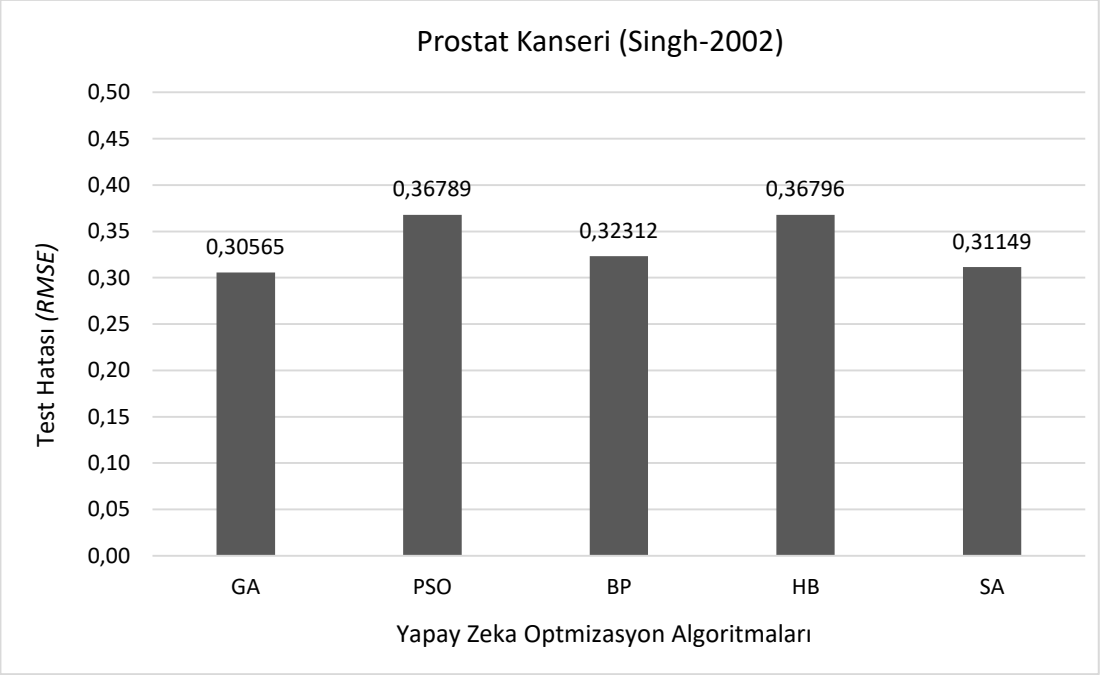
Şekil 4.8**.** Beyin kanseri verisinin sınıflandırılmasında elde edilen test (*RMSE*) hata değerleri



Şekil 4.9**.** Merkezi Sinir Sistemi kanseri verisinin sınıflandırılmasında elde edilen test (*RMSE*) hata değerleri



Şekil 4.10**.** Rahim kanseri verisinin sınıflandırılmasında elde edilen test (*RMSE*) hata değerleri



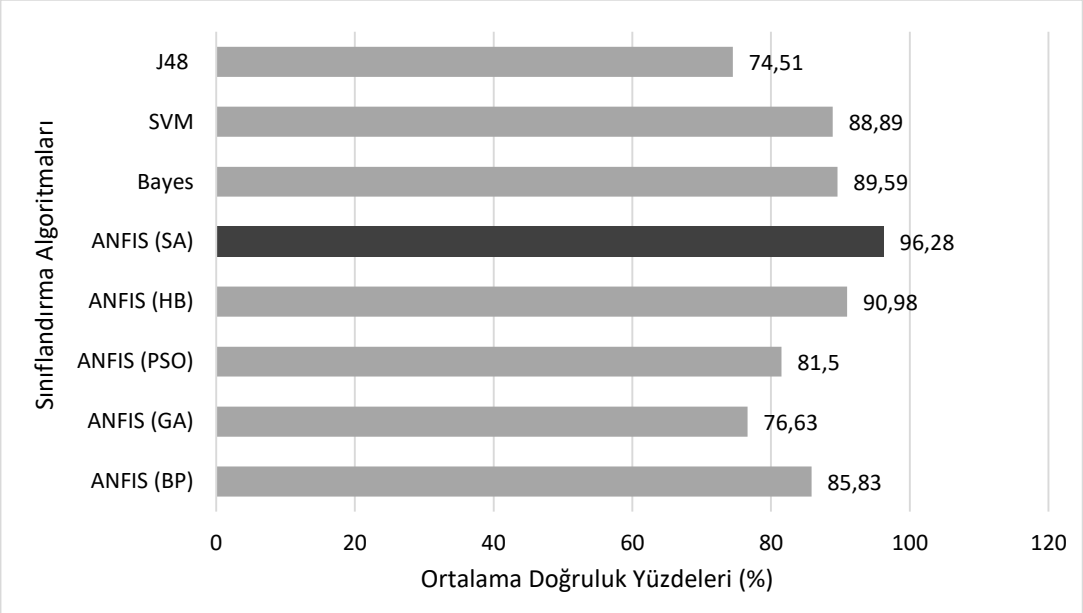
Şekil 4.11**.** Prostat kanseri verisinin sınıflandırılmasında elde edilen test (*RMSE*) hata değerleri

Mikrodizi gen ekspresyon (ifade) verilerinde bulunan gen sayılarının çokluğu ve bu veriler arasında çoğunlukla doğrusal olmayan ilişkiler bulunması gibi nedenlerden dolayı geleneksel sınıflandırma algoritmalarının başarımları sınırlı kalabilmektedir. Bu sebeple bu bölümde önerilen yöntemin performansı, BayesNet, SVM ve J48 gibi istatistiksel yöntemlerin kanser verilerini sınıflandırma performansları ile karşılaştırılmıştır. SVM yönteminde C-Style Soft Margin SVM (C-SVM) ve radyal tabanlı çekirdek (kernel) fonksiyonu kullanılmış ve cost parametresi 0,5 olarak optimize edilmiştir. BayesNet yöntemi için K2 arama algoritması tercih edilmiştir. J48 Karar Ağacı yönteminde ise budama için kullanılan güven faktörü 0,25 seçilmiş ve yaprak başına düşen örnek sayısı 2 olarak alınmıştır.

Tüm veri setleri için oluşturulan ANFIS modelleri GA, PSO, BP, HB ve SA algoritmaları ile eğitilerek eğitim ve test hata değerleri yukarıda verilmiştir. Bu sonuçlara göre yöntemlerin sınıflandırma performansları istatistiksel yöntemlerin performansları ile de karşılaştırılmıştır. Yapılan simülasyon çalışmaları neticesinde optimal parametre değerleri ile elde edilen, mikrodizi kanser veri setlerinin sınıflandırılma performanslarına ait doğruluk yüzdesi bilgileri Tablo 4.4'de ve tüm veri setlerinden elde edilen ortalama doğruluk yüzdeleri ise Şekil 4.12'de verilmiştir. Görüldüğü üzere en başarılı sınıflandırma performansı ANFIS-SA modeline aittir. Dolayısıyla önerilen yöntem, geleneksel yöntemlerin sınırlı kalan performans problemlerine de çözüm getirmiştir.

Tablo 4.4. Sınıflandırma algoritmalarının doğruluk yüzdeleri (%)

Veri Seti	**ANFIS (BP)**	**ANFIS (GA)**	**ANFIS (PSO)**	**ANFIS (HB)**	**ANFIS (SA)**	**Bayes**	**SVM**	**J48**
Akciğer Kanseri *(Gordon-2002)*	88,88	90,74	92,59	94,44	**100**	100	96,30	88,89
Beyin Kanseri *(Nutt-2003)*	87,50	75,00	87,50	100	**100**	87,5	100	75,00
Merkezi Sinir Sistemi Kanseri *(Pomeroy-2002)*	90,00	70,00	80,00	90,00	**100**	90,00	70,00	60,00
Rahim Kanseri *(Risinger-2003)*	69,23	53,85	53,85	76,92	**84,62**	76,92	84,62	61,54
Prostat Kanseri *(Singh-2002)*	93,55	93,55	93,55	93,55	**96,77**	93,55	93,55	87,10



Şekil 4.12**.** Tüm mikrodizi gen ifadesi profilleri üzerindeki sınıflandırma algoritmalarının ortalama doğruluk yüzdeleri (%)

Şekil 4.12'de görüldüğü gibi ANFIS - SA yöntemi, % 96,28 ortalama oranıyla bütün kanser veri setlerini sınıflandırmada en yüksek performansı sergilemektedir. Ayrıca sınıflandırma yöntemlerinin ortalama başarımları ve birbirlerine göre gelişim yüzdeleri de aşamalı olarak Tablo 4.5'de verilmiştir. ANFIS ağının önerilen SA algoritması ile eğitilmesi durumunda sınıflandırma başarımının geleneksel BP algoritmasına göre %10,80 iyileştiği, istatistiksel yöntemlerden J48 yöntemine göre ise % 22,61 daha başarılı sonuçlar sağladığı görülmektedir. Gelişim yüzdeleri SA algoritmasının sınıflandırma problemlerinde diğer yöntemlere göre oldukça başarılı olduğunu göstermektedir.

Tablo 4.5. Metotların ortalama başarımları ve gelişim yüzdeleri

Yöntem	**Ortalama Performans (%)**	**Gelişim Yüzdeleri (%)**
J48 Karar Ağacı	74,51	-
ANFIS – GA	76,63	2,76
ANFIS - PSO	81,50	5,97
ANFIS – BP	85,83	5,04
SVM	88,89	3,44
BayesNet	89,59	0,78
ANFIS - HB	90,98	1,52
ANFIS - SA	**96,28**	**5,50**

Model başarımını değerlendirirken kullanılan temel kavramlar hata oranı, kesinlik (Precision), duyarlılık (Recall) ve F-ölçütüdür (F-Measure). Modelin başarısı, doğru sınıfa atanan örnek sayısı ve yanlış sınıfa atılan örnek sayısı nicelikleriyle alakalıdır. Kesinlik, sınıfı 1 olarak tahminlenmiş True Pozitif örnek sayısının, sınıfı 1 olarak tahminlenmiş tüm örnek sayısına oranıdır. Duyarlılık, doğru sınıflandırılmış pozitif örnek sayısının toplam pozitif örnek sayısına oranıdır. Kesinlik ve duyarlılık ölçütleri tek başına anlamlı bir karşılaştırma sonucu çıkarmamıza yeterli değildir. Her iki ölçütü beraber değerlendirmek daha doğru sonuçlar verir. Bunun için f-ölçütü tanımlanmıştır. F-ölçütü, kesinlik ve duyarlılığın harmonik ortalamasıdır. Bu kapsamda Tablo 4.6 – 4.10'da 5 farklı kanser verisinin kullanılan yöntemler ile sınıflandırılmasına ait Doğruluk, Kesinlik, Duyarlılık ve F-ölçütü değerleri verilmiştir.

Tablo 4.6. Akciğer kanseri veri setine ait model başarım ölçütleri

Akciğer Kanseri *(Gordon-2002)*	**Doğruluk (%)**	**Kesinlik**	**Duyarlılık**	**F-Ölçütü**
ANFIS-SA	100,00	1,000	1,000	1,000
ANFIS-GA	90,74	0,926	0,907	0,909
ANFIS-PSO	92,59	0,935	0,912	0,916
ANFIS-BP	88,88	0,902	0,889	0,890
ANFIS-HB	94,44	0,952	0,944	0,945
SVM	96,30	0,965	0,963	0,962
BayesNet	100,00	1,000	1,000	1,000
J48	88,89	0,889	0,889	0,881

Tablo 4.7. Beyin kanseri veri setine ait model başarım ölçütleri

Beyin Kanseri *(Nutt-2003)*	**Doğruluk (%)**	**Kesinlik**	**Duyarlılık**	**F-Ölçütü**
ANFIS-SA	100,00	1,000	1,000	1,000
ANFIS-GA	75,00	0,750	0,750	0,750
ANFIS-PSO	87,50	0,900	0,875	0,873
ANFIS-BP	87,50	0,900	0,875	0,873
ANFIS-HB	100,00	1,000	1,000	1,000
SVM	100,00	1,000	1,000	1,000
BayesNet	87,50	0,900	0,875	0,873
J48	75,00	0,750	0,750	0,750

Tablo 4.8. Merkezi sinir sistemi kanser veri setine ait model başarım ölçütleri

Merkezi Sinir Sistemi Kanseri *(Pomeroy-2002)*	**Doğruluk (%)**	**Kesinlik**	**Duyarlılık**	**F-Ölçütü**
ANFIS-SA	100,00	1,000	1,000	1,000
ANFIS-GA	70,00	0,720	0,700	0,703
ANFIS-PSO	80,00	0,820	0,800	0,803
ANFIS-BP	90,00	0,920	0,900	0,901
ANFIS-HB	90,00	0,920	0,900	0,901
SVM	70,00	0,490	0,700	0,576
BayesNet	90,00	0,925	0,900	0,903
J48	60,00	0,680	0,600	0,617

Tablo 4.9. Rahim kanseri veri setine ait model başarım ölçütleri

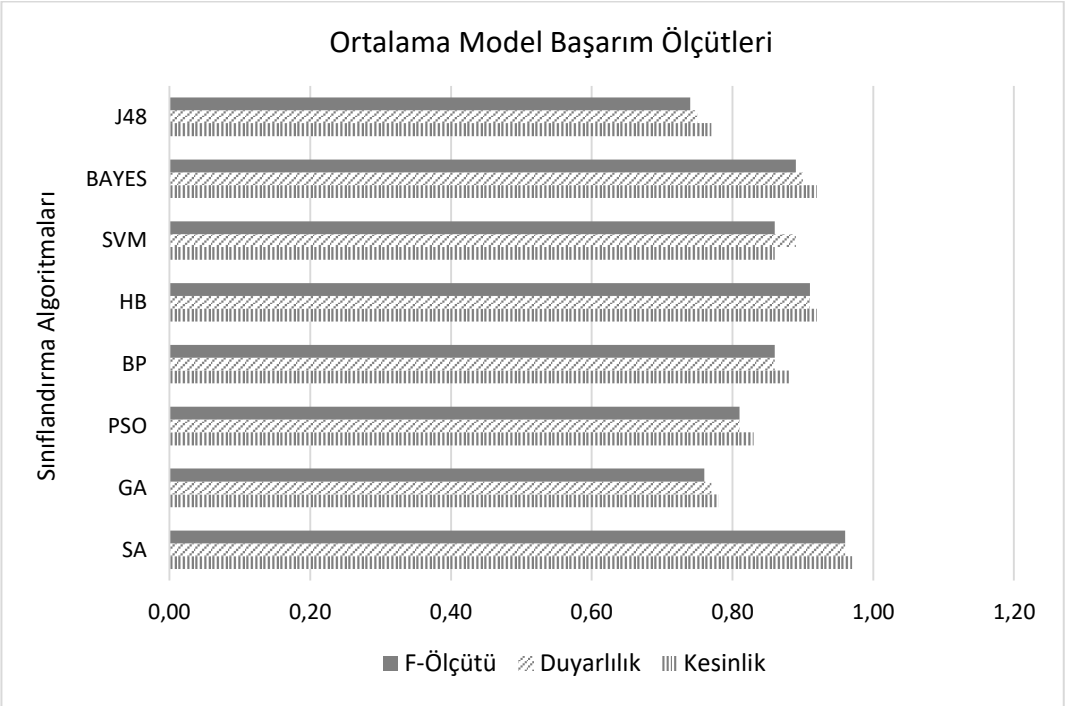
Rahim Kanseri *(Risinger-2003)*	**Doğruluk (%)**	**Kesinlik**	**Duyarlılık**	**F-Ölçütü**
ANFIS-SA	84,62	0,865	0,846	0,846
ANFIS-GA	53,85	0,577	0,538	0,518
ANFIS-PSO	53,85	0,577	0,538	0,518
ANFIS-BP	69,23	0,724	0,692	0,69
ANFIS-HB	76,92	0,785	0,769	0,766
SVM	84,62	0,885	0,846	0,846
BayesNet	76,92	0,849	0,769	0,756
J48	61,54	0,647	0,615	0,596

Tablo 4.10. Prostat kanseri veri setine ait model başarım ölçütleri

Prostat Kanseri *(Singh-2002)*	**Doğruluk (%)**	**Kesinlik**	**Duyarlılık**	**F-Ölçütü**
ANFIS-SA	96,77	0,97	0,968	0,968
ANFIS-GA	93,55	0,944	0,935	0,936
ANFIS-PSO	93,55	0,935	0,935	0,935
ANFIS-BP	93,55	0,935	0,935	0,935
ANFIS-HB	93,55	0,935	0,935	0,935
SVM	93,55	0,943	0,935	0,935
BayesNet	93,55	0,943	0,935	0,935
J48	87,10	0,878	0,871	0,871

Tablo 4.11. Yöntemlere ait genel model başarım ölçütleri

Yöntemler	Doğruluk (%)	Kesinlik	Duyarlılık	F-Ölçütü
ANFIS-SA	96,28	0,97	0,96	0,96
ANFIS-GA	76,63	0,78	0,77	0,76
ANFIS-PSO	81,50	0,83	0,81	0,81
ANFIS-BP	85,83	0,88	0,86	0,86
ANFIS-HB	90,98	0,92	0,91	0,91
SVM	88,89	0,86	0,89	0,86
BayesNet	89,59	0,92	0,90	0,89
J48	74,51	0,77	0,75	0,74



Şekil 4.13**.** Tüm mikrodizi gen ifadesi profilleri üzerindeki sınıflandırma algoritmalarının ortalama Kesinlik, Duyarlılık ve F-ölçütü değerleri

Genel olarak sonuçları değerlendirdiğimizde ise tüm veri setlerinin ortalama istatiksel model başarım ölçütü bilgileri Tablo 4.11 ve Şekil 4.13'de verilmiştir. Ortalama model başarım ölçütü bilgileri SA algoritmasının sınıflandırma problemlerinde diğer yöntemlere göre oldukça başarılı olduğunu göstermektedir. Doğruluk yüzdeleri ile birlikte verilen model başarım ölçütleri sonuçların güvenilirliğini ölçmek için önemlidir. Elde edilen

sonuçlardan SA algoritmasın tüm ölçütlerde en iyi sonucu vermesi, SA algoritması ile önerilen yöntemin oldukça güçlü ve başarımı yüksek bir yöntem olduğunu ortaya koymaktadır.

4.1.6. Genel Değerlendirme ve Sonuçlar

Tez çalışmasının bu bölümünde, mikrodizi gen ifade kanser verilerinin sınıflandırılması amacıyla ANFIS ve SA algoritmasına dayalı hibrid bir yeni yöntem önerilmiştir. Önerilen yöntemin performansı beş farklı kanser türüne ait gen verileri üzerinde test edilerek çalışmalar yapılmıştır. Ayrıca, ANFIS ağı BP, HB, GA ve PSO algoritmalarıyla da eğitilerek kanser verileri sınıflandırılmıştır. Sınıflandırma performansları önerilen metodun performansı ile karşılaştırılmıştır. Ayrıca, mikrodizi gen ifade (ekspresyon) verilerindeki gen sayılarının çokluğu ve bu veriler arasında çoğunlukla doğrusal olmayan ilişkiler bulunması gibi nedenlerden dolayı geleneksel yöntemler mikrodizi gen problemlerinin çözümlerinde yetersiz ve sınırlı kalabilmektedir. Bu kapsamda, tez çalışmasında kullanılan mikrodizi gen ifade kanser verileri BayesNet, SVM ve J48 gibi istatistiksel yöntemler kullanılarak da sınıflandırmıştır. İstatistiksel yöntemlerden elde edilen başarım performansları da önerilen yöntem ve diğer yaklaşımların performansları ile karşılaştırılmıştır.

Sınıflandırma başarım performanslarının yanı sıra istatistiksel başarım ölçütlerinden Kesinlik, Duyarlılık ve F-ölçütü değerleri de model başarımlarının hesaplanmasında önemli etkenlerdir. Bu kapsamda, sınıflandırma performansı için bu değerler de hesaplanarak tüm yaklaşımlar için karşılaştırmalar yapılmıştır. Elde edilen sonuçlardan yine önerilen yöntemin kullanılan tün yaklaşımlardan daha iyi performans sergilediği görülmüştür. Tüm bu çalışmalar önerilen yönteme olan güvenilirliği artırmıştır. Sonuç olarak, ANFIS ağının SA algoritması ile eğitilmesine dayalı yaklaşımın kanser gen verilerinin sınıflandırılmasında oldukça başarılı olduğu görülmüştür.

Genel olarak, belirlenmesi gereken parametre sayısının fazla olması ve ağın eğitilmesinde kullanılan türev tabanlı algoritmaların bölgesel optimuma takılma riski içermesi gibi hususlar ANFIS'in başarımını olumsuz etkileyebilen temel unsurlar arasında olduğundan, tez çalışmasının bu bölümünde, ANFIS ağının türev tabanlı algoritmalar yerine yapay zekâ optimizasyon algoritmalarından olan SA algoritması kullanılarak eğitilmesi bu

olumsuzlukların üstesinden geldiğini göstermektedir. SA algoritmasının türeve dayalı algoritmalar gibi kısıtlamalar içermemesi ve problemlere kolaylıkla uygulanabilir olması nedeniyle, ANFIS ağının farklı problemlere yönelik uygulamalarında da kullanılabileceği değerlendirilmiştir.

4.2. Dinamik Sistemlerin Kimliklendirilmesi için FCM Tabanlı ANFIS Ağının SA Algoritması ile Eğitilmesi

ANFIS modelinin oluşturulması ve yapılandırılması süreci çözülmesi gereken temel problemlerdendir. Giriş değişkenlerinin seçilmesi, giriş uzayının bölümlenmesi, girişler için üyelik fonksiyonlarının türünün ve sayısının seçilmesi, kural katmanının oluşturularak kural sayısının belirlenmesi ve üyelik fonksiyonları parametrelerinin başlangıç değerlerinin belirlenmesi gibi tüm bu işlemler ANFIS modelinin oluşturulması aşamasını kapsamaktadır. ANFIS modelinin yapılandırılması için genellikle tecrübeye dayalı bir uzman bilgisi kullanılmıştır. Ancak birçok durumda uzmanların bilgisi en doğru tercihleri yaparak ANFIS modeli için en uygun kriterlerin belirlenmesinde yetersiz kalabilmektedir. Farklı denemeler yaparak en uygun özelliklerin seçilmesi süreci zaman almakla birlikte uygunluğu tartışılabilmektedir. Ayrıca, bu yaklaşımın dezavantajı giriş sayısı arttıkça üyelik fonksiyonu sayısı ve kural sayısının da hızlı bir şekilde artmasıdır. Bu sayıların artması, ANFIS ağının eğitilmesini zorlaştırıp, gürültü gibi girişleri ezberlemesine neden olmaktadır. Dolayısıyla kullanılan bilgisayar işlemcisi yükünü ve daha fazla hafıza gereksinimini artırmaktadır. Buna ek olarak, giriş alanını bölümlemek için kümelerin sayısı hakkında bilgiye ihtiyaç vardır. Bu nedenlerden dolayı tecrübeye dayalı yöntemi en uygun ANFIS modelinin oluşturulmasında kullanmak oldukça zordur. Bu yöntem giriş sayısı az olan ve / veya her bir giriş için tahsis edilen üyelik fonksiyonu sayısı az olan problemler için uygundur.

Bu süreci otomatikleştiren ve bulanık modelin performansını ve yorumlanabilirliğini geliştirirken çalışma sürecini önemli ölçüde azaltan çeşitli yöntemler bulunmaktadır. Izgara Bölümleme (Grid Partitioning, GP), Eksiltici Kümeleme (Subtractive Clustering, SC) ve Bulanık C-Ortalamalar Kümeleme (Fuzzy C-Means Clustering, FCM) yöntemleri ise yaygın olarak kullanılan bu yöntemlerden bazılarıdır. Bu yöntemlerin arkasındaki temel fikir ise, giriş-çıkış örnek verisi arasındaki ilişkiyi öğrenerek bulanık kuralları tahmin etmektir. Giriş uzayını etkili bir şekilde bölümlemek, ANFIS'in kural sayısını ve

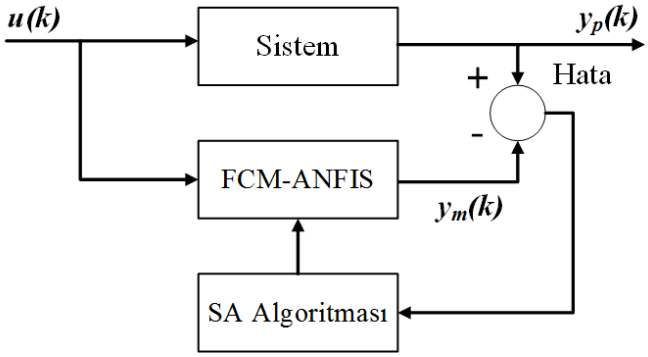
üyelik fonksiyonlarının sayısının azaltarak, modelin hem öğrenme hem de uygulanma aşamasında hızını ve performansını artırmaktadır. Bu çalışmada, bulanık kural sayısını en aza indirgemek amacıyla kullanılan bulanık kural oluşturma tekniklerinden FCM yöntemi kullanılmıştır. FCM, ANFIS ağının kural tabanını ve üyelik fonksiyonlarını otomatik olarak oluşturmaktadır.

Tezin bu bölümünde dinamik sistemlerin kimliklendirilmesi için iteratif tabanlı bir algoritma olan SA algoritması kullanılarak FCM kümeleme yöntemi ile modellenmiş ANFIS ağının (FCM-ANFIS) eğitilmesine yönelik yeni bir yaklaşım önerilmiştir. FCM yöntemine dayalı bu yaklaşım ile öğrenme sürecinin hızlandırılması ve daha iyi sonuçlar elde edilmesi amaçlanmıştır. Önerilen yaklaşımın performansı doğrusal ve doğrusal olmayan sistemler üzerinde test edilmiştir. Simülasyon sonuçları, BP, GA, PSO ve literatürden alınan farklı yöntemler ile karşılaştırılmıştır. Bu çalışma sonunda SA algoritmasının FCM tabanlı ANFIS parametrelerini, BP, GA, PSO ve diğer yöntemlerden daha başarılı optimize ettiği görülmüştür.

4.2.1. Önerilen Yöntem

Bu bölümde, ANFIS modelinin yapılandırılması için tecrübeye dayalı uzman bilgisi yerine, giriş/çıkış veri seti bilgisini kullanarak modelin kural sayısını ve girişlere uygulanan üyelik fonksiyonlarının sayısını otomatik belirleyen FCM kümeleme yöntemi kullanılmıştır. Oluşturulan FCM tabanlı ANFIS ağının (FCM-ANFIS) parametreleri ise SA algoritması ile optimize edilmiştir.

Önerilen yöntem dinamik sistem kimliklendirme problemlerine uygulanmıştır. ANFIS'den elde edilen çıkış ile doğrusal veya doğrusal olmayan dinamik sistemden alınan çıkış arasındaki hata değerinin en az olması için SA algoritması ile ANFIS parametreleri optimize edilerek en iyi modelin oluşturulması sağlanmıştır. Bu kimliklendirme mimarisini gösteren gösteren blok diyagram Şekil.4.14'de verilmiştir.



Şekil 4.14. Sistem kimliklendirme için FCM-ANFIS modelinin SA algoritması ile eğitilmesi planı

Eşitlik 3.3 kullanılarak elde edilen *RMSE* hata fonksiyonu çözümün hata değerinin belirlenmesinde kullanılmıştır. Burada kullanılan $y_m(k)$, k zamanında ANFIS tarafından elde edilen çıkışı ve $y_p(k)$ ise sistemin gerçek çıkışını göstermektedir. *N* ise uygulamada kullanılan örnek sayısını göstermektedir. Amaç, ANFIS ağının giriş ve sonuç parametrelerini ayarlayarak etkin hatayı (*RMSE*) minimize etmektir.

Başlangıçta ANFIS'in tüm parametreleri için FCM kümeleme yöntemi ile çözüm uzayı oluşturulur. Sonrasında her bir iterasyon için SA algoritması kullanılarak parametreler güncellenir. Eğitme aşamasında, önce *u(k) (k=0,1,...)* giriş sinyali dizisinin bir elemanı, mevcut çözümden sağlanan ağırlıklarla tasarlanmış ANFIS ve sistemin her ikisinin girişlerine uygulanır. Sonra ANFIS ve sistem çıkışları arasındaki hata değeri hesaplanır. Giriş sinyali dizisinin tüm elemanları için hesaplanan hata değerleri yardımıyla Eşitlik 3.3 'de verilen etkin hata değeri, modelin mevcut durumunun iyiliğinin ölçüsü olarak kullanılmıştır. Mevcut durumdan sağlanan etkin hata değeri ile yeni üretilen komşu çözümle sağlanan etkin hata değeri karşılaştırılır. Eğer yeni çözüm ile sağlanan etkin hata değeri daha düşük ise yeni çözüm mevcut çözüm olarak atanır ve araştırmaya devam edilir. Aksi takdirde Metropolis kriterine göre seçim yapılır.

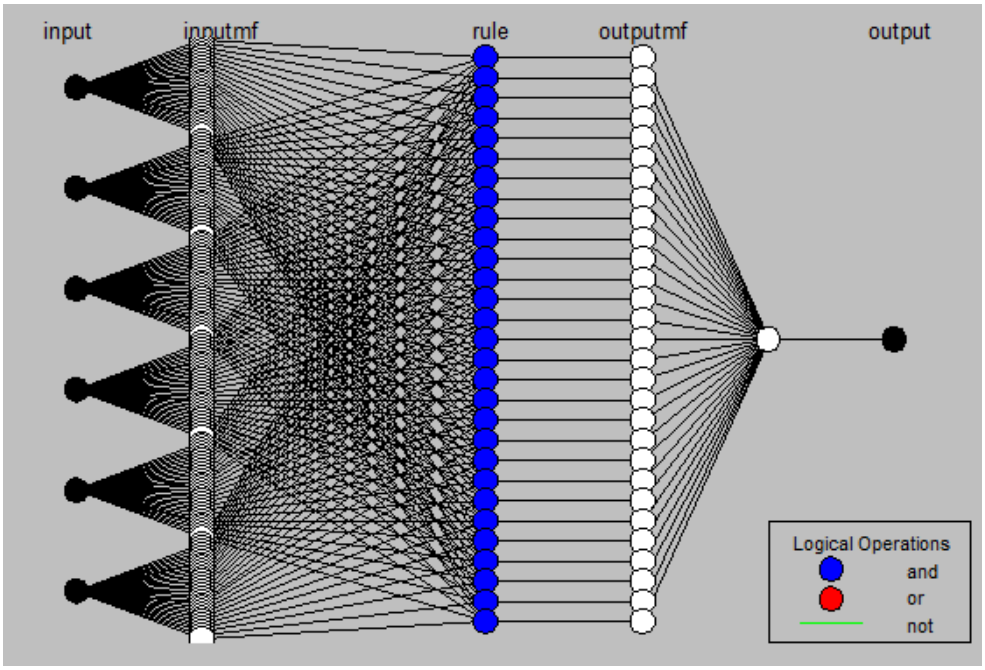
Eğitmede kullanılacak giriş dizisinin seçimi eğitme başarısı üzerinde önemlidir. Giriş dizisinin elemanları genellikle sinüzoidal ve rastgele giriş değerlerinden seçilmektedir [160]. Bu çalışmada gerçekleştirilen simülasyon çalışmalarında düzgün dağılıma sahip rastgele *u(k)* giriş sinyal dizisi kullanılmıştır.

4.2.2. Simülasyon Sonuçları

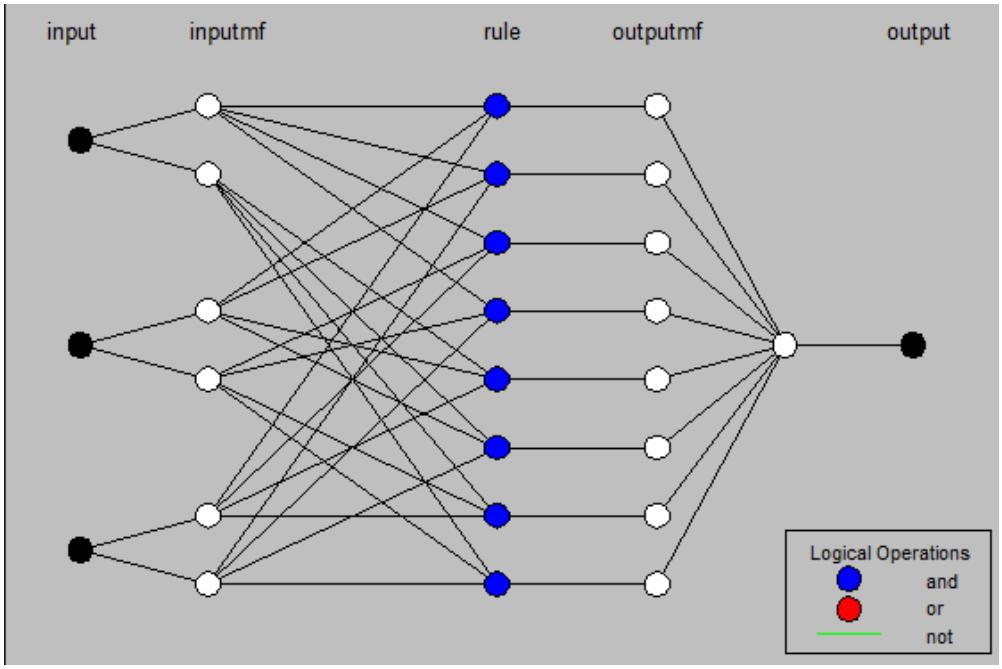
Yeni önerilen FCM-ANFIS modeli SA algoritması ile eğitilerek, performansının incelenmesi amacıyla Bölüm 3'de dikkate alınan sistem kimliklendirme problemleri üzerinde test edilmiştir. Karşılaştırmalar sağlıklı bir şekilde yapılabilmesi amacıyla aynı sistemler ve aynı giriş verileri kullanılarak simülasyonlar gerçekleştirilmiştir.

Kullanılan her bir dinamik sistem için 250 adet giriş sinyali *u(k) (k=0,1, ,.., 250)* rastgele olarak (-2, 2) aralığında üretilmiştir. Üretilen 250 adet giriş sinyali kullanılarak her bir dinamik sistemin gerçek çıkışı $y_p(k)$ hesaplanmıştır. Dolayısıyla 250 adet giriş ve çıkış sinyalinden oluşan bir veri seti hazırlanmıştır. Bu veri setlerinden rastgele seçilen 200 tanesi eğitim için, kalan 50 tanesi ise test için kullanılmıştır. 200 adet giriş sinyali $u(k)$ ANFIS ağına uygulanarak, sistemin gerçek çıkışı $y_p(k)$ ve ANFIS ağının çıkışı $y_m(k)$ arasındaki hata değeri Eşitlik 3.3'e göre hesaplanmıştır. Hesaplanan hatayı minimize etmek için ANFIS ağının parametreleri SA algoritması ile eğitilmiştir. Eğitilen ANFIS modelinin performansı ise kalan 50 adet veri seti ile test edilmiştir.

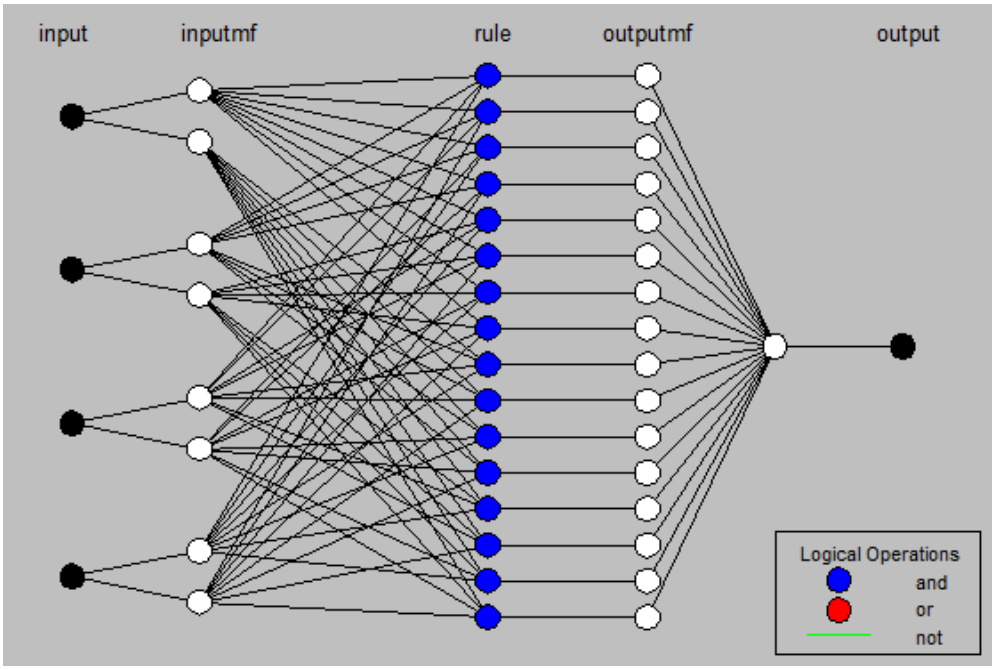
Modellenecek dinamik sistemler için sırasıyla (6,3,4) giriş ve 1 çıkıştan oluşan ANFIS yapıları oluşturulmuştur. ANFIS ağının kural sayısını ve bağlantı sayısını azaltarak, modelin hem öğrenme hem de uygulanma aşamasında hızını ve performansını artırabilmek için FCM kümeleme yöntemi ile ANFIS modelleri oluşturulmuştur. ANFIS ağının girişleri, dinamik sistemlerin i. dereceden önceki tüm giriş ve çıkış değerleridir. Bu yapılar sırasıyla Şekil 4.15, Şekil 4.16 ve Şekil 4.17'de verilmiştir. Oluşturulan ANFIS yapılarının her girişi için 2 parametreli *gauss* türünde üyelik fonksiyonu kullanılmıştır. Bu şekilde birinci sistem için 348 başlangıç parametresi ve 203 sonuç parametresi toplamda 551 parametre ile 29 kuraldan oluşan ANFIS yapısı kullanılarak sistem modellenmiştir. İkinci sistem için 48 başlangıç parametresi ve 32 sonuç parametresi toplamda 80 parametre ile 8 kuraldan oluşan ANFIS yapısı kullanılmıştır. Son olarak üçüncü sistem için 104 başlangıç parametresi ve 65 sonuç parametresi toplamda 169 parametre ile 13 kuraldan oluşan ANFIS yapısı kullanılarak sistem modellenmiştir.



Şekil 4.15. Sistem 1 için 6 girişli ANFIS modeli



Şekil 4.16. Sistem 2 için 3 girişli ANFIS modeli



Şekil 4.17. Sistem 3 için 4 girişli ANFIS modeli

Tezin 3. Bölümünde yapılan dinamik sistemlerin kimliklendirilmesi çalışmalarında tecrübeye dayalı uzman bilgisi ile oluşturulan ANFIS modelleri kullanılmıştır. FCM tabanlı ANFIS ile manuel tecrübeye dayalı yani GP ile modellenen ANFIS karşılaştırıldığında, FCM-ANFIS modellerinin kural katmanında daha az kural ve bağlantı olduğu görülmektedir. Dolayısıyla önerilen yaklaşım ile daha hızlı öğrenme süreci ve daha iyi sonuçlar elde edilmektedir.

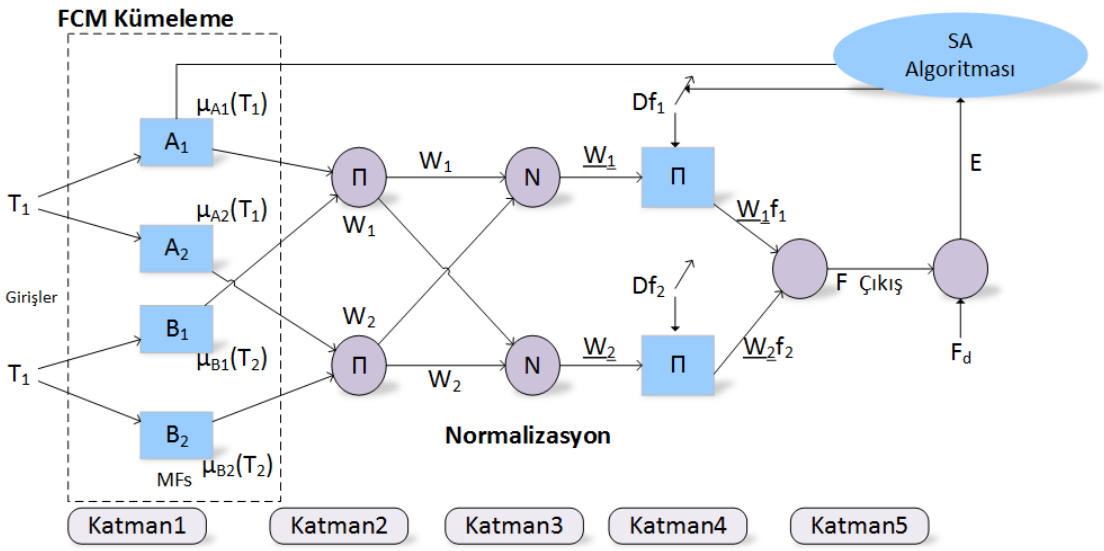
Yapay zekâ optimizasyon algoritmalarının performansı algoritmaların kontrol parametrelerine fazlaca bağlıdır ve genellikle kontrol parametreleri için hangi değerlerin kullanılması gerektiğine yönelik kesin kural ve yöntemler bulunmamaktadır [191]. Bu kapsamda, parametreler için araştırmacıların yaygın olarak önerdiği aralıklarda değerler tayin edilmesi veya çok sayıda deneme yapılarak optimal parametre değerlerinin belirlenmesi yaklaşımı kabul görmektedir. Çalışmada, SA algoritmasının parametre değerlerine karar vermek için birçok deneme gerçekleştirilmiş ve denemeler sonucunda sıcaklık noktalarının sayısı 10, her bir sıcaklık noktasında gerçekleştirilen iterasyon sayısı 8 ve sıcaklık azalma parametresi 0,2 olarak alınmıştır.

Önerilen yaklaşımın performansının farklı yöntemlerle karşılaştırılması için ANFIS ağı dikkate alınan problemlerin sınıflandırılması amacıyla GA, PSO ve BP algoritmaları ile

de eğitilmiştir. Bu amaçla, GA modeli için iterasyon sayısı 100, popülasyon büyüklüğü 25, çaprazlama oranı 0,8 ve mutasyon oranı 0,01 olarak seçilmiştir. PSO algoritması için iterasyon sayısı 100 ve popülasyon büyüklüğü olarak belirlenmiştir. Kullanılan BP algoritması için öğrenme katsayısı 0,2 ve momentum oranı 0,4 olarak seçilmiştir.

Ayrıca, önerilen yaklaşımın performansı dinamik sistemlerin kimliklendirilmesinde yaygın olarak kullanılan Elman geribeslemeli yapay sinir ağına ait literatürden alınan sonuçlarla da karşılaştırılmıştır [162,163]. Dikkate alınan çalışmalarda Kalınlı ve arkadaşları, Elman ağının dinamik sistemlerin kimliklendirilmesi amacıyla Tabu Arama Algoritması (TS), SA algoritması ve Paralel Tabu Arama Algoritması (PTS) kullanılarak eğitilmesine dayalı bir yaklaşım önermiş ve geleneksel BP algoritmasına göre daha başarılı sonuçlar elde etmiştir [162, 163].

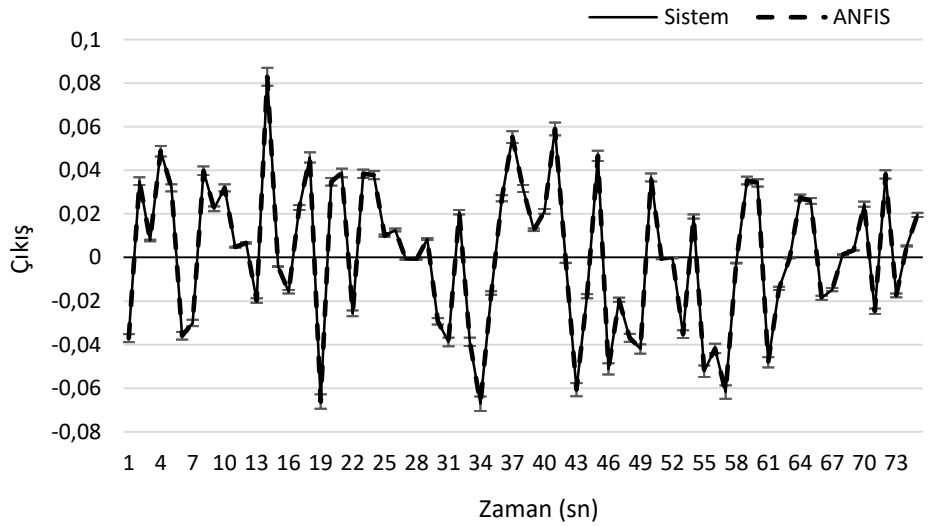
Her sistem için ANFIS 15 kez eğitilmiş ve her seferinde rastgele seçilen çözüm uzayı ile eğitime başlanmıştır. Ayrıca her örnek için elde edilen *RMSE* hata değerlerinin ortalaması alınarak, ortalama *RMSE* ($RMSE_{AVG}$) değeri bulunmuştur. Veri setleri üzerinde yöntemlerin başarımlarını ölçmek için verilerin % 70'i eğitim, % 30'u ise test için kullanılmıştır. Yapılan simülasyon çalışmalarında kullanılan modelin oluşturulmasını ve oluşturulan modelin parametrelerinin optimize edilmesini gösteren blok diyagram Şekil 4.18'de verilmiştir.



Şekil 4.18. FCM ile modellenmiş temel ANFIS mimarisi

4.2.2.1. Dinamik Sistem 1

Simülasyon için kullanılan birinci sistem, Eşitlik 3.4 ile tanımlanan 3. dereceden doğrusal bir sistemdir [162]. Sistem 1 için simülasyon çalışmaları sonucunda elde edilen *RMSE* hata değerleri ve literatürden alınan farklı yöntemlere ait hata değerleri Tablo 4.12'de verilmiştir. Tablo 4.12'de verilen sonuçlardan bu çalışmada önerilen yaklaşımın performansının diğer yöntemlere göre oldukça yüksek başarıma sahip olduğu açıkça görülmektedir. FCM-ANFIS ağının GA ile eğitilmesi durumunda PSO algoritmasına göre hatanın % 5,17 iyileştiği, SA ile eğitilmesi durumunda ise hata değerinin % 33,70 daha düşük olduğu görülmektedir. FCM-ANFIS ağının SA algoritması kullanılarak eğitilmesi durumunda elde edilen hata değerinin ise Elman Ağı için literatürdeki en başarılı sonucu veren uygulamaya göre % 99,97 oranında daha başarılı olduğu da görülmektedir. Ayrıca 3. Bölümde önerilen ANFIS-SA yönteminden FCM_ANFIS-SA yaklaşımı ile % 99,94 daha düşük hata bulunmuştur. ANFIS ağının SA algoritması ile eğitilmesine yönelik olarak gerçekleştirilen 15 farklı denemede elde edilen hata değerlerinin eğitim ve test işlemleri için standart sapması sırasıyla 2,0246E-06 ve 2,2321E-06 olarak bulunmuştur. ANFIS ağının SA algoritması ile eğitilmesine yönelik gerçekleştirilen denemelerden birisine ait sistem ve ANFIS ağının cevapları Şekil 4.19'da verilmiştir.



Şekil 4.19. Sistem 1 için, sistemin ve SA algoritması ile eğitilen ANFIS modelinin cevapları (*RMSE* hata değeri = 9,2087E-12)

Tablo 4.12. Doğrusal sistem için sonuçların karşılaştırılması

Model	Parametre	Kural Sayısı	Eğitim Hatası	Test Hatası
Elman - BP [162]	54	-	-	7,6754E-02
Elman - TS [162]	54	-	-	8,1664E-03
Elman - PTS [163]	54	-	-	6,4693E-03
Elman - SA [162]	54	-	-	4,9653E-03
ANFIS - BP	472	64	5,8892E-03	6,7061E-03
ANFIS - GA	472	64	3,9306E-03	4,7279E-03
FCM_ANFIS – BP	513	29	2,2568E-03	3,0688E-03
ANFIS - SA	**472**	**64**	**2,1126E-03**	**2,6009E-03**
FCM_ANFIS - PSO	513	29	1,9526E-06	2,1798E-06
FCM_ANFIS - GA	513	29	1,7421E-06	2,0671E-06
FCM_ANFIS - SA	**513**	**29**	**1,6655E-06**	**1,4452E-06**

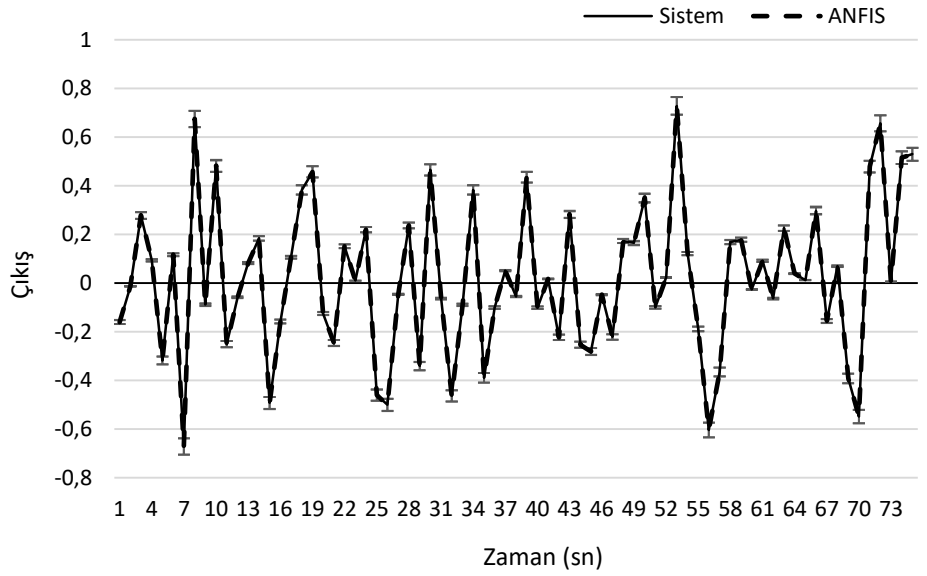
4.2.2.2. Dinamik Sistem 2

Simülasyon için kullanılan ikinci sistem, [162]'den alınan ve Eşitlik 3.5 ile tanımlanan küçük açılarla salınan basit bir sarkaça aittir. Sistem 2 için ise simülasyon çalışmaları sonucunda elde edilen *RMSE* hata değerleri ve literatürden alınan farklı yöntemlere ait hata değerleri Tablo 4.13'de verilmiştir. Tablo 4.13'de verilen sonuçlardan bu çalışmada önerilen yaklaşımın performansının diğer yöntemlere göre oldukça yüksek başarıma sahip olduğu açıkça görülmektedir. FCM-ANFIS ağının GA ile eğitilmesi durumunda PSO algoritmasına göre hatanın % 18,64 iyileştiği, SA ile eğitilmesi durumunda ise hata değerinin % 63,13 daha düşük olduğu görülmektedir. FCM-ANFIS ağının SA algoritması kullanılarak eğitilmesi durumunda elde edilen hata değerinin ise Elman Ağı için literatürdeki en başarılı sonucu veren uygulamaya göre % 94,26 oranında daha başarılı olduğu da görülmektedir. Ayrıca 3. Bölümde önerilen ANFIS-SA yönteminden FCM_ANFIS-SA yaklaşımı ile % 51,76 daha düşük hata bulunmuştur. ANFIS ağının SA algoritması ile eğitilmesine yönelik olarak gerçekleştirilen 15 farklı denemede elde edilen hata değerlerinin eğitim ve test işlemleri için standart sapması sırasıyla 1,7582E-04 ve 8,0535E-04 olarak bulunmuştur. ANFIS ağının SA algoritması ile eğitilmesine yönelik

gerçekleştirilen denemelerden birisine ait sistem ve ANFIS ağının cevapları Şekil 4.20'de verilmiştir.

Tablo 4.13. Sistem 2 için sonuçların karşılaştırılması

Model	Parametre	Kural Sayısı	Eğitim Hatası	Test Hatası
Elman - BP [162]	54	-	-	2,6182E-01
Elman - TS [162]	54	-	-	1,1990E-01
Elman - SA [162]	54	-	-	2,9430E-02
Elman - PTS [163]	54	-	-	2,8318E-02
ANFIS - BP	44	8	7,9911E-03	9,6074E-03
ANFIS - GA	44	8	8,6912E-03	8,3221E-03
FCM_ANFIS – BP	80	8	5,2125E-03	5,6135E-03
FCM_ANFIS - PSO	80	8	4,5072E-03	4,4083E-03
FCM_ANFIS - GA	80	8	5,1024E-03	3,5864E-03
ANFIS - SA	**44**	**8**	**4,3542E-03**	**3,3691E-03**
FCM_ANFIS - SA	**80**	**8**	**7,4850E-04**	**1,6253E-03**



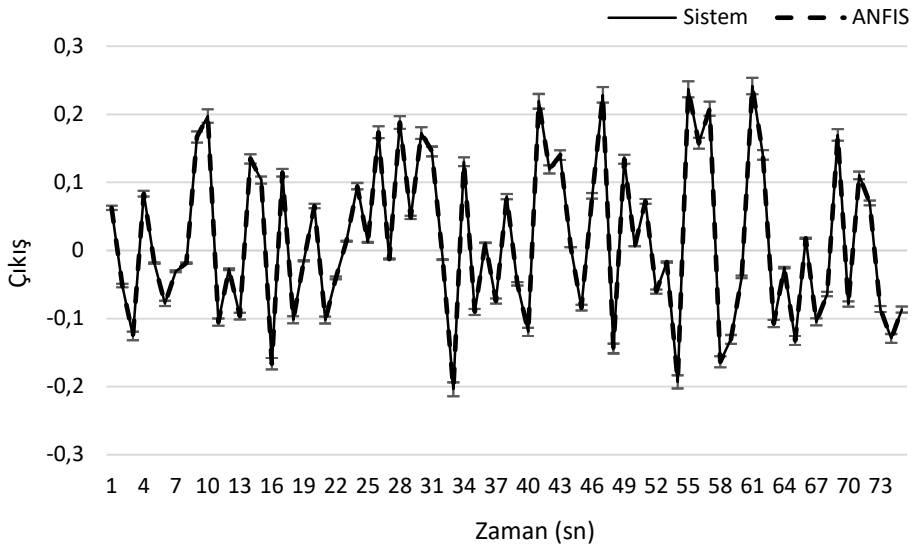
Şekil 4.20. Sistem 2 için, sistemin ve SA algoritması ile eğitilen ANFIS modelinin cevapları (*RMSE* hata değeri = 8,7368E-04)

4.2.2.3. Dinamik Sistem 3

Üçüncü sistem Eşitlik 3.6'da ayrık zaman ifadesi verilen ve Kaynak [164]' den alınan doğrusal olmayan sistemdir. Sistem 3 için simülasyon çalışmaları sonucunda elde edilen *RMSE* hata değerleri ve literatürden alınan farklı yöntemlere ait hata değerleri Tablo 4.14'de verilmiştir. Tablo 4.14'de verilen sonuçlardan bu çalışmada önerilen yaklaşımın performansının diğer yöntemlere göre oldukça yüksek başarıma sahip olduğu açıkça görülmektedir. FCM-ANFIS ağının PSO ile eğitilmesi durumunda GA algoritmasına göre hatanın % 11,48 iyileştiği, SA ile eğitilmesi durumunda ise hata değerinin % 64,38 daha düşük olduğu görülmektedir. FCM-ANFIS ağının SA algoritması kullanılarak eğitilmesi durumunda elde edilen hata değerinin ise Elman Ağı için literatürdeki en başarılı sonucu veren uygulamaya göre % 96,04 oranında daha başarılı olduğu da görülmektedir. Ayrıca 3. Bölümde önerilen ANFIS-SA yönteminden FCM_ANFIS-SA yaklaşımı ile % 80,37 daha düşük hata bulunmuştur. ANFIS ağının SA algoritması ile eğitilmesine yönelik olarak gerçekleştirilen 15 farklı denemede elde edilen hata değerlerinin eğitim ve test işlemleri için standart sapması sırasıyla 6,9267E-05 ve 4,3773E-04 olarak bulunmuştur. ANFIS ağının SA algoritması ile eğitilmesine yönelik gerçekleştirilen denemelerden birisine ait sistem ve ANFIS ağının cevapları Şekil 4.21'de verilmiştir.

Tablo 4.14. Sistem 3 için sonuçların karşılaştırılması

Model	Parametre	Kural Sayısı	Eğitim Hatası	Test Hatası
Elman - BP [162]	54	-	-	1,2863E-01
Elman - TS [162]	54	-	-	6,1720E-02
Elman - SA [162]	54	-	-	2,1760E-02
ANFIS - BP	96	16	1,9711E-02	1,5819E-02
ANFIS - GA	96	16	9,5121E-03	8,5310E-03
FCM_ANFIS – BP	169	13	4,5643E-03	4,9685E-03
ANFIS - SA	**96**	**16**	**4,2370E-03**	**4,3930E-03**
FCM_ANFIS - GA	169	13	2,0582E-03	2,4211E-03
FCM_ANFIS - PSO	169	13	1,5875E-03	2,1432E-03
FCM_ANFIS - SA	**169**	**13**	**3,3650E-04**	**8,6247E-04**



Şekil 4.21. Sistem 3 için, sistemin ve SA algoritması ile eğitilen ANFIS modelinin cevapları (*RMSE* hata değeri = 8,6247E-04)

4.2.3. Genel Değerlendirme ve Sonuçlar

Tez çalışmasının bu bölümünde, SA algoritmasının dinamik sistemlerin kimliklendirilmesi amacıyla FCM-ANFIS parametrelerinin eğitilmesi için kullanılmasına yönelik yeni bir yaklaşım önerilmiştir. Doğrusal ve doğrusal olmayan dinamik sistemlerin kimliklendirilmesine yönelik yapılan simülasyon çalışmaları sonucunda, önerilen yaklaşımın dikkate alınan BP, TS, GA ve PSO algoritmalarından çok daha başarılı sonuçlar verdiği görülmüştür. Bununla birlikte tecrübeye dayalı ANFIS modelleme yerine FCM tabanlı ANFIS oluşturma, modelin hem öğrenme hem de uygulanma aşamasında hızını ve performansını artırmıştır. Elde edilen sonuçlardan SA algoritmasının ANFIS ağının parametre değerlerinin belirlenmesi amacıyla başarılı bir şekilde kullanılabileceği kanaatine varılmıştır. Ayrıca, SA algoritmasının türeve dayalı algoritmalar gibi kısıtlamalar içermemesi ve problemlere kolaylıkla uygulanabilir olması nedeniyle, ANFIS ağının farklı problemlere yönelik uygulamalarında da kullanılabileceği değerlendirilmiştir.

Elde edilen sonuçların farklı yöntemlere ait sonuçlarla karşılaştırılması neticesinde en iyi eğitim ve test hata değerlerinin önerilen yöntem ile bulunduğu görülmüştür. Yapılan çalışmalar sonucunda kabul edilebilir *RMSE* hata değerinin bulunması, doğrusal ve

doğrusal olmayan sistemlerin kimliklendirilmesinde SA algoritması kullanılarak yapılan FCM kümeleme tabanlı ANFIS eğitiminin oldukça etkili olduğunu göstermektedir. Ayrıca önerilen yöntemde düşük standart sapma değerine sahip sonuçlar elde edilmiştir. Standart sapma değerinin düşük olması, farklı başlangıç çözümleriyle başlanan eğitimlerin sonucunda elde edilen *RMSE* hata değerlerinin birbirine yakın, yani sonuçların güvenilir olduğunu göstermektedir.

5. BÖLÜM

BENZETİLMİŞ TAVLAMA ALGORİTMASI İÇİN YENİ BİR MODEL

5.1. Giriş

Benzetilmiş tavlama iteratif tabanlı bir algoritmadır ve genellikle rastgele seçilen tek bir çözüm adayı üzerinde çalışır. Algoritma bu seri yapısı ile birçok problem için çok başarılı sonuçlar vermesine rağmen, küresel optimuma erişebilmesi için gereken süre başlangıç çözümüne bağlı olarak değişim göstermektedir [192-194]. Bu durum temel SA algoritmasının dinçlik bakımından performansını düşürmektedir. Ancak bu tür iteratif tabanlı algoritmalar genel olarak bulundukları bölgenin optimumunu popülasyon tabanlı algoritmalara göre daha kısa sürelerde bulabilmektedirler.

Bununla beraber, popülasyon tabanlı algoritmaların çözümlerin bir seti ile çalışması ve rastgele araştırma algoritmalarının bir türü olması, araştırmanın küresel optimumun bulunduğu bölgeye hızla erişmesini sağlar. Ancak, bu algoritmaların yüksek oranda olasılık tabanlı arama stratejilerine sahip olması nedeniyle bulundukları bölgedeki optimum çözümü bulabilmeleri genellikle uzun süreler gerektirebilmektedir [195-197].

Tez çalışması kapsamında, SA algoritmasına popülasyon tabanlı algoritmaların paralel yapı avantajını kazandırarak yukarıda bahsedilen dezavantajını yok edip, performansını artırmak amacıyla yeni bir paralel benzetilmiş tavlama algoritması (PSA) modeli önerilmiştir. İteratif ve popülasyon tabanlı algoritmaların avantajlarını birleştiren bu model, popülasyon tabanlı algoritmaların temel operatörlerinden olan çaprazlama (crossover) operatörü üzerine dayalıdır. Geliştirilen SA algoritmasının performansı doğrusal ve doğrusal olmayan sistem kimliklendirme problemleri üzerinde test edilmiş ve elde edilen sonuçlar farklı yöntemlere ait sonuçlarla karşılaştırılmıştır.

5.2. Önerilen Modelin Genel Yapısı

Bu bölümde popülasyon tabanlı algoritmaların paralel yapı avantajını SA algoritmasına kazandıracak yeni bir model önerilmiştir. Modelin paralel olarak tasarlanmasındaki amaç, araştırmanın birden çok noktadan başlatılarak çözüm uzayının daha geniş bir kısmının taramasını sağlamaktır. Bu modelde farklı bölgelerde bulunan yerel optimal çözümlerden faydalanarak yeni çözümlerin üretilmesi, başka bir ifadeyle, araştırmanın geçmiş adımlarında kazanılmış olan bilgilerin birleştirilmesi suretiyle daha iyi çözümlerin bulunması amacıyla popülasyon tabanlı algoritmaların çaprazlama operatöründen faydalanılmıştır.

Önerilen model blok diyagram olarak Şekil 5.1'de verilmiştir. Yapıda 1'den ***n***' e kadar numaralandırılmış her bir blok, Bölüm 3' de tanımlanan temel SA algoritmasını temsil etmektedir. Birinci kademede paralel olarak çalışmakta olan ***n*** adet temel SA algoritmasının her biri araştırmaya farklı başlangıç çözümleriyle başlar. Her bir algoritma ***k*** değerlendirme sayısı boyunca koşulur ve bulunan ***n*** çözüme çaprazlama işlemi uygulanarak yeni çözümler elde edilir. İkinci kademedeki ***n*** adet temel SA algoritması için gerekli olan başlangıç çözümleri, birinci kademe sonucu bulunan en iyi çözüm ile (***n*-1**) adet çaprazlama işlemi neticesi elde edilen yeni çözümlerden oluşturulur. Bu tezde önerilen model, Kalınlı tarafından önerilen Baskın İndeks Seçimli Paralel Tabu Arama algoritmasından esinlenerek geliştirilmiştir [196].

Çaprazlama işlemi için 3 farklı çaprazlama operatörü kullanılmıştır. İlk çaprazlama işlemiyle bir önceki kademe sonucu bulunan en iyi birinci ve ikinci çözümün aritmetik ortalaması alınarak yeni bir aday çözüm elde edilir. Bu işlem Eşitlik 7.1'de $\alpha = 0.5$ olduğu durum ile ifade edilir. İkinci çaprazlama işlemi ise yine bir önceki kademe sonucu bulunan en iyi çözüm ile ikinci en iyi çözüm arasında gerçekleştirilmekte ve Eşitlik 7.1'de verilen formülde α , 0 ve 1 arasında rastgele üretilen bir sayı olarak alınmaktadır.

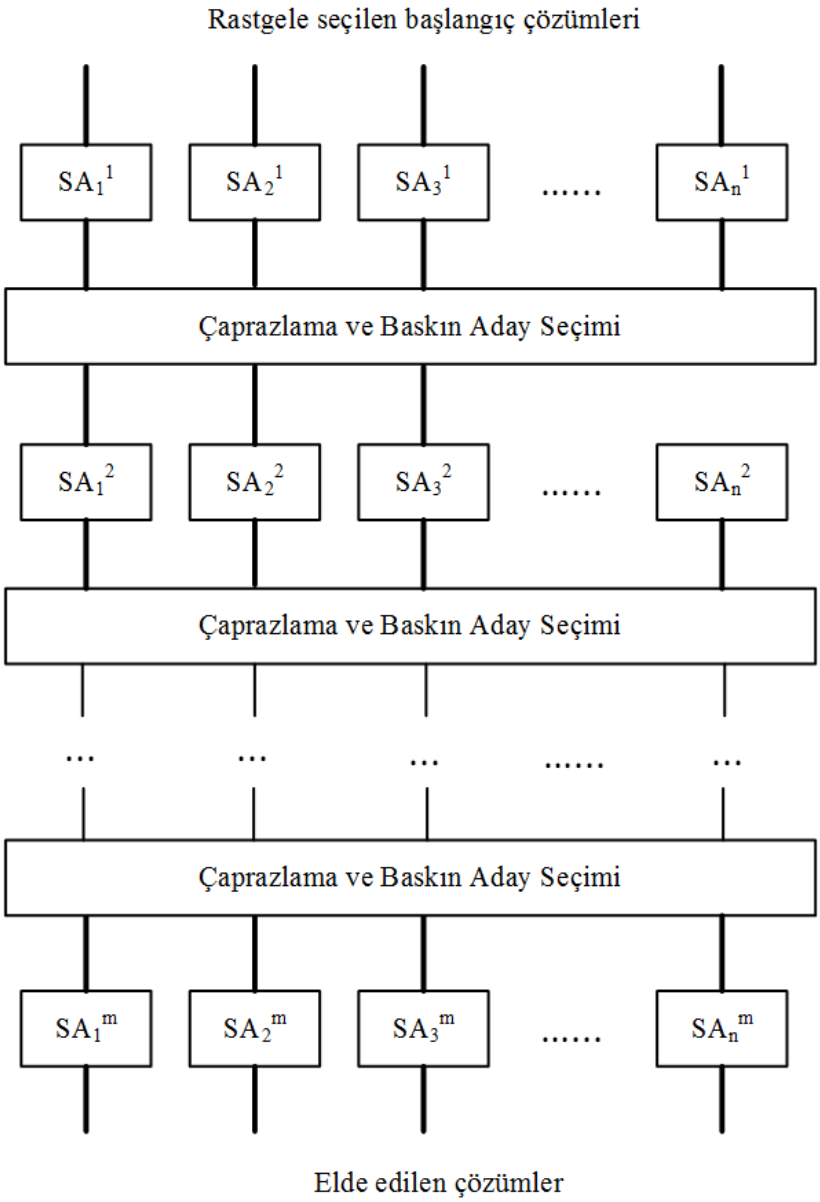
$$S_{yeni} = \alpha S_1 + (1-\alpha) S_2 \tag{7.1}$$

S_1 = Bir önceki kademenin en iyi birinci çözümü

S_2 = Bir önceki kademenin en iyi ikinci çözümü

Üçüncü çaprazlama işlemiyle ise bir önceki kademede bulunan ***n*** adet çözümün aritmetik ortalaması alınarak yeni bir aday çözüm elde edilir. Bu çaprazlama işlemi Eşitlik 7.2'de gösterilmektedir.

$$S_{yeni} = \sum_{i=1}^{n} S_i / n \tag{7.2}$$

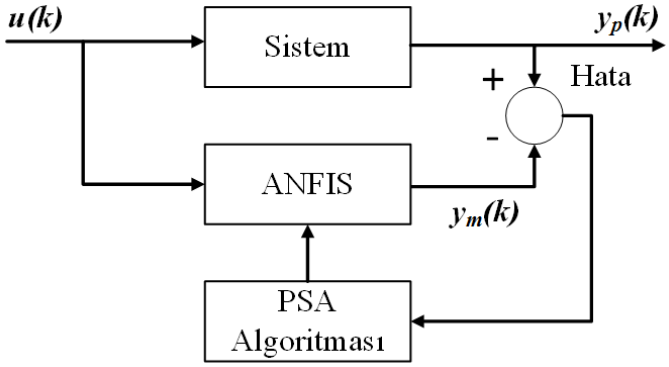


Şekil 5.1. Önerilen paralel benzetilmiş tavlama algoritmasının yapısı

Algoritma yukarıda belirtilen işlemleri "***m***" kademe boyunca tekrarlar ve problem için ***n*** adet yerel optimal çözüm bulunur. Bu yerel optimal çözümlerden birinin veya daha fazlasının küresel optimum olma ihtimali temel SA'nın bulduğu çözüme göre çok daha yüksektir. Değerlendirme sayısı (***k***), paralel çalışan standart SA algoritma sayısı (***n***) ve kademe sayısı (***m***) probleme göre değişebilir. Buradaki kademe sayısı aynı zamanda çaprazlama işlemi sayısını vermektedir. Önerilen algoritmanın kontrol parametreleri olarak nitelendirilebilecek bu parametrelerin değerlerinin, algoritmanın performansı üzerinde etkili olacağı açıktır.

5.3. Sistem Kimliklendirme için Önerilen PSA Modeli Kullanılarak ANFIS Ağının Eğitilmesi

Bölüm 3'de temel SA algoritması ile dinamik sistemlerin kimliklendirilmesi için eğitilen ANFIS ağı, tez çalışmasının bu bölümünde performansının incelenmesi amacıyla yeni önerilen paralel SA algoritması ile de eğitilmiştir. ANFIS'den elde edilen çıkış ile doğrusal veya doğrusal olmayan dinamik sistemlerden alınan çıkış arasındaki hata değerinin en az olması için PSA algoritması ile ANFIS parametreleri optimize edilerek en iyi modelin oluşturulması sağlanmıştır. Bu yapıyı gösteren blok diyagram Şekil.5.2'de verilmiştir.



Şekil 5.2. Sistem kimliklendirme için ANFIS modelinin PSA algoritması ile eğitilmesi planı

Önerilen algoritmanın performansı Bölüm 3'de dikkate alınan sistem kimliklendirme problemleri üzerinde test edilmiştir. Karşılaştırmalar sağlıklı bir şekilde yapılabilmesi

amacıyla aynı sistemler ve aynı giriş verileri kullanılarak simülasyonlar gerçekleştirilmiştir.

Kullanılan her bir dinamik sistem için yine 250 adet giriş sinyali *u(k) (k=0,1, ,.., 250)* rastgele olarak (-2, 2) aralığında üretilmiştir. Üretilen 250 adet giriş sinyali kullanılarak her bir dinamik sistemin gerçek çıkışı $y_p(k)$ hesaplanmıştır. Dolayısıyla 250 adet giriş ve çıkış sinyalinden oluşan bir veri seti hazırlanmıştır. Bu veri setlerinden rastgele seçilen 200 tanesi eğitim için, kalan 50 tanesi ise test için kullanılmıştır. 200 adet giriş sinyali $u(k)$ ANFIS ağına uygulanarak, sistemin gerçek çıkışı $y_p(k)$ ve ANFIS ağının çıkışı $y_m(k)$ arasındaki hata değeri Eşitlik 3.3'e göre hesaplanmıştır. Hesaplanan hatayı minimize etmek için ANFIS ağının parametreleri PSA algoritması ile eğitilmiştir.

Paralel yapıda çözüm araştırması 4 farklı noktadan ($n = 4$) başlamış ve her çözüm noktası 320' şer değerlendirme ($k = 320$) süresince geliştirilmiş, daha sonra çözümlere çaprazlama ve baskın aday seçimi uygulanmıştır. Böylece, 4 paralel SA algoritması bir kademede toplam 320 x 4 = 1280 değerlendirme yapmıştır Toplam 4 kademe ($m = 4$), dolayısıyla toplam 5120 değerlendirme sonunda algoritma durdurulmuştur. Her sistem için yapılan 15 farklı denemede bulunan *RMSE* hata değerlerinin ortalaması alınarak, ortalama *RMSE* ($RMSE_{AVG}$) değeri bulunmuştur.

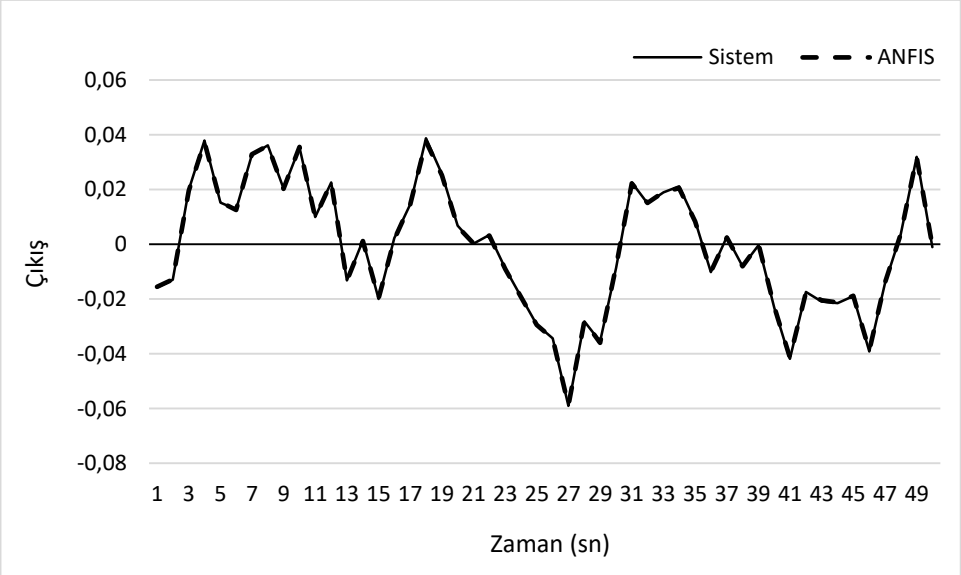
5.3.1. Dinamik Sistem 1

Önerilen yeni paralel yapı Eşitlik 3.4'de verilen doğrusal sisteme uygulanarak elde edilen sonuçlar farklı yöntemlere ait *RMSE* hata değerleri ile karşılaştırmalı olarak Tablo 5.1'de verilmiştir. Tablo 5.1'de verilen sonuçlardan bu çalışmada önerilen yeni yaklaşımın performansının diğer yöntemlere göre daha başarılı olduğu açıkça görülmektedir. ANFIS ağının PSA ile eğitilmesi durumunda elde edilen hata değerinin temel SA algoritması ile elde edilen hata değerinden % 61,46, Elman Ağı için literatürdeki en başarılı sonucu veren uygulamaya göre ise % 79,81 oranında daha başarılı olduğu görülmektedir. ANFIS ağının PSA algoritması ile eğitilmesine yönelik olarak gerçekleştirilen 15 farklı denemede elde edilen hata değerlerinin eğitim ve test işlemleri için standart sapması sırasıyla 3,7416E-04 ve 1,6685E-04 olarak bulunmuştur. ANFIS ağının PSA algoritması ile eğitilmesine yönelik gerçekleştirilen denemelerden 3,1971E-04 en küçük etkin hata değeri için sistem

ve ANFIS ağının cevapları Şekil 5.3'de verilmiştir. Sonuç olarak, yeni modelin kullanılması ile etkin hata değerleri önemli oranda küçültülmüştür. Standart sapma değerlerinin çok küçük olması sistemin kararlılığını da artırmıştır.

Tablo 5.1. Doğrusal sistem için sonuçların karşılaştırılması

Model	Eğitim Hatası	Test Hatası
Elman - BP [162]	-	7,6754E-02
Elman - TS [162]	-	8,1664E-03
Elman - PTS [163]	-	6,4693E-03
Elman - SA [162]	-	4,9653E-03
ANFIS - BP	5,8892E-03	6,7061E-03
ANFIS - GA	3,9306E-03	4,7279E-03
ANFIS - SA	**2,1126E-03**	**2,6009E-03**
ANFIS - PSA	**8,3574E-04**	**1,0025E-03**



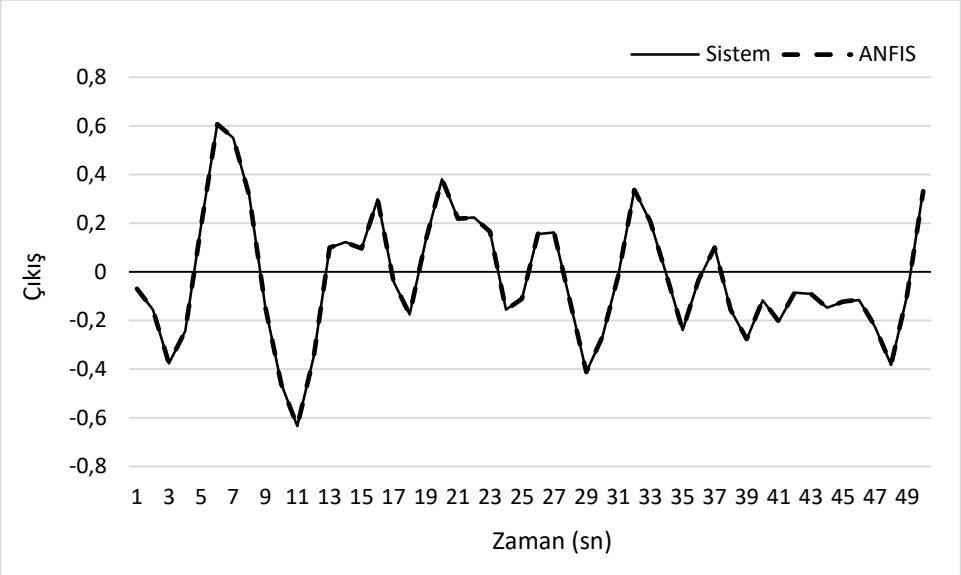
Şekil 5.3. Sistem 1 için, sistemin ve PSA algoritması ile eğitilen ANFIS modelinin cevapları (*RMSE* hata değeri = 3,1971E-04)

5.3.2. Dinamik Sistem 2

Eşitlik 3.5'de verilen doğrusal olmayan Sistem 2 için simülasyon çalışmaları sonucunda elde edilen *RMSE* hata değerleri ve literatürden alınan farklı yöntemlere ait *RMSE* hata değerleri Tablo 5.2'de verilmiştir. Tablo 5.2'de verilen sonuçlardan bu çalışmada önerilen yeni paralel modelin performansının diğer yöntemlere göre daha başarılı olduğu açıkça görülmektedir. ANFIS ağının PSA ile eğitilmesi durumunda elde edilen hata değerinin temel SA algoritması ile elde edilen hata değerinden % 55,09, Elman Ağı için literatürdeki en başarılı sonucu veren uygulamaya göre ise % 94,66 oranında daha iyi olduğu görülmektedir. ANFIS ağının PSA algoritması ile eğitilmesine yönelik olarak gerçekleştirilen 15 farklı denemede elde edilen hata değerlerinin eğitim ve test işlemleri için standart sapması sırasıyla 1,4458E-04 ve 5,1525E-04 olarak bulunmuştur. ANFIS ağının PSA algoritması ile eğitilmesine yönelik gerçekleştirilen denemelerden 8,7335E-04 en küçük etkin hata değeri için sistem ve ANFIS ağının cevapları Şekil 5.4'de verilmiştir. Sonuç olarak, yeni modelin kullanılması ile etkin hata değerleri önemli oranda küçültülmüştür.

Tablo 5.2. Sistem 2 için sonuçların karşılaştırılması

Model	Eğitim Hatası	Test Hatası
Elman - BP [162]	-	2,6182E-01
Elman - TS [162]	-	1,1990E-01
Elman - SA [162]	-	2,9430E-02
Elman - PTS [163]	-	2,8318E-02
ANFIS - BP	7,9911E-03	9,6074E-03
ANFIS - GA	8,6912E-03	8,3221E-03
ANFIS - SA	**4,3542E-03**	**3,3691E-03**
ANFIS - PSA	**1,4213E-03**	**1,5132E-03**



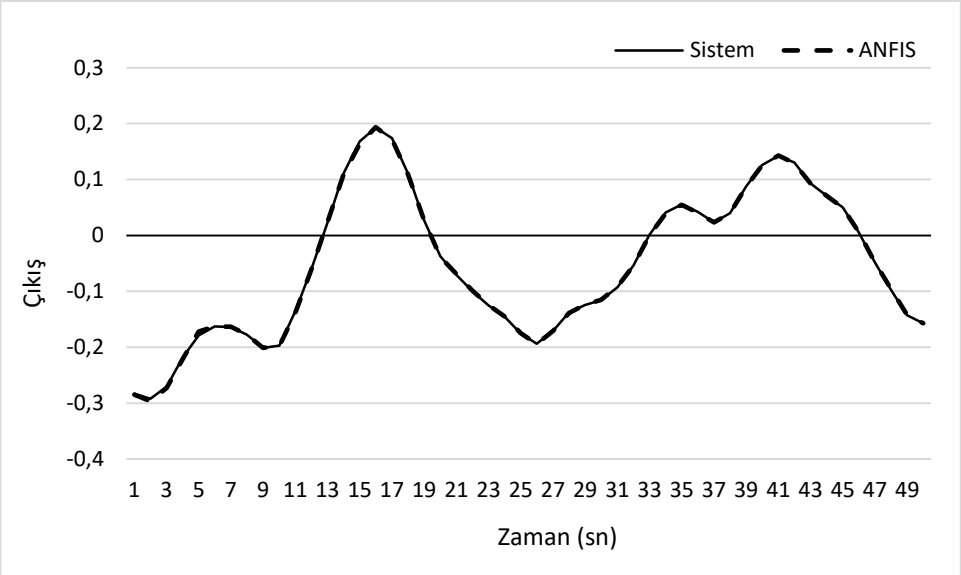
Şekil 5.4. Sistem 2 için, sistemin ve PSA algoritması ile eğitilen ANFIS modelinin cevapları (*RMSE* hata değeri = 8,7335E-04)

5.3.3. Dinamik Sistem 3

Eşitlik 3.6'da verilen doğrusal olmayan Sistem 3 için ise simülasyon çalışmaları sonucunda elde edilen *RMSE* hata değerleri ve literatürden alınan farklı yöntemlere ait *RMSE* hata değerleri Tablo 5.3'de verilmiştir. Tablo 5.3'de verilen sonuçlardan bu çalışmada önerilen yeni yaklaşımın performansının diğer yöntemlere göre daha başarılı olduğu açıkça görülmektedir. ANFIS ağının PSA ile eğitilmesi durumunda elde edilen hata değerinin standart SA algoritması ile elde edilen hata değerinden % 36,23 daha başarılı olduğu görülmektedir. ANFIS-PSA modeli ile edilen hata değerinin Elman Ağı için literatürdeki en başarılı sonucu veren uygulamaya göre ise % 87,13 oranında daha başarılı olduğu da görülmektedir. ANFIS ağının PSA algoritması ile eğitilmesine yönelik olarak gerçekleştirilen 15 farklı denemede elde edilen hata değerlerinin eğitim ve test işlemleri için standart sapması sırasıyla 6,8584E-04 ve 9,1833E-04 olarak bulunmuştur. ANFIS ağının PSA algoritması ile eğitilmesine yönelik gerçekleştirilen denemelerden 1,7015E-03 en küçük etkin hata değeri için sistem ve ANFIS ağının cevapları Şekil 5.5'de verilmiştir. Sonuç olarak, yeni modelin kullanılması ile etkin hata değerleri önemli oranda küçültülmüştür. Standart sapma değerlerinin çok küçük olması sistemin kararlılığını da artırmıştır.

Tablo 5.3. Sistem 3 için sonuçların karşılaştırılması

Model	Eğitim Hatası	Test Hatası
Elman - BP [162]	-	1,2863E-01
Elman - TS [162]	-	6,1720E-02
Elman - SA [162]	-	2,1760E-02
ANFIS - BP	1,9711E-02	1,5819E-02
ANFIS - GA	9,5121E-03	8,5310E-03
ANFIS - SA	**4,2370E-03**	**4,3930E-03**
ANFIS - PSA	**2,0003E-03**	**2,8012E-03**



Şekil 5.5. Sistem 3 için, sistemin ve PSA algoritması ile eğitilen ANFIS modelinin cevapları (*RMSE* hata değeri = 1,7015E-03)

5.4. Genel Değerlendirme ve Sonuçlar

Sonuç olarak, tezin bu bölümünde yapılan çalışmada benzetilmiş tavlama algoritması için popülasyon temelli algoritmaların bir operatörü olan çaprazlama operatörü üzerine dayalı yeni bir model geliştirilmiş ve modelin performansı doğrusal ve doğrusal olmayan dinamik sistemlerin kimliklendirilmesi amacıyla ANFIS parametrelerinin eğitilmesi için kullanılmıştır. Önerilen modelin optimum veya optimuma yakın çözümleri kısa zamanda

bulabilmede temel SA algoritmasına ve diğer yöntemlere göre daha yüksek performansa sahip olduğu görülmüştür. Ayrıca önerilen yöntemde düşük standart sapma değerine sahip sonuçlar elde edilmiştir. Standart sapma değerinin düşük olması da önerilen yöntemin performansıyla birlikte daha güçlü ve dinç olduğunu göstermektedir. Temel SA algoritmasının aynı süreli her koşmada aynı performansı gösterememesi onun başarısız olmasından değildir. Buradaki problem daha öncede açıklandığı gibi zamanla ilgilidir. Temel SA algoritması da her koşmada yüksek kalite değerli çözümleri bulabilecek kabiliyettedir. Fakat bu rastgele tek noktadan çözüme başlanması ve başlangıç çözümünün optimale uzaklığına bağlı olarak farklı zamanlarda gerçekleşebilmektedir. SA algoritmasına paralel çalışma özelliği kazandırılması suretiyle elde edilen sonuçlardan görüldüğü gibi, algoritmanın kısa sürede daha iyi çözümlere ulaşması her koşmada sağlanmış, böylece algoritmanın performansında büyük bir gelişme elde edilmiştir.

Yapılan çalışmalar sonunda Tablo 5.1, 5.2 ve 5.3'de verilen sonuçlardan, paralel benzetilmiş tavlama algoritmasının ANFIS parametrelerini temel benzetilmiş tavlama algoritması ve diğer yöntemlerden daha başarılı optimize ettiği görülmüştür. Önerilen PSA algoritmasının SA algoritmasına göre daha yüksek performans göstermesinin çeşitli nedenleri vardır. Bunlardan birincisi, PSA algoritmasında araştırmaya çözüm uzayının farklı noktalarından paralel olarak başlanması ve her aşamada rastgele çözüm noktalarının araştırmaya katılmasıyla çözüm uzayının daha geniş bir bölgesinin taranmasıdır. İkincisi ise, çaprazlama operatörünün ve baskın aday seçiminin kullanılması ile araştırmanın ilerleyen adımlarında önceki aşamalarda bulunan sonuçların da yönlendirici bir etki yapmasıdır. Burada yeni önerilen benzetilmiş tavlama algoritması, temel benzetilmiş tavlama algoritmasına göre çok daha yüksek bir performansa sahip olmuştur.

6. BÖLÜM

TARTIŞMA, SONUÇ VE ÖNERİLER

6.1. Tartışma ve Sonuç

Büyük boyutlu gerçek dünya problemleri genellikle zor problemler sınıfındadır. Bu problemlerin parametreleri arasında çoğunlukla doğrusal olmayan ilişkiler söz konusu olmaktadır ve yine bu türdeki problemlerin bazılarının matematiksel olarak modellenmesi dahi mümkün değildir. Genel olarak, kabul edilebilir bir zaman dilimi içerisinde en iyi çözümün bulunmasının garanti edilemediği problemler zor problemler olarak tanımlanır. Türeve dayalı geleneksel optimizasyon algoritmaları ile bu tür problemlerin çözümü ya çok uzun süreler gerektirebilmekte ya da mümkün olamamaktadır. Bu nedenle gerçek dünya problemlerinin büyük zorluğu, araştırmacıların ilgisini doğada ve insanda var olan zeki sistematikleri ve davranışları taklit eden yapay zekâ optimizasyon algoritmalarına kaydırmıştır.

Zor problemlerin çözümünde yaygın olarak kullanılan ANFIS, örneklerden öğrenebilme ve öğrenmiş olduğu bilgiyi belirli kurallara bağlı olarak ağ üzerinde dağıtabilme yeteneği gibi esnek ve güçlü yapısı sayesinde makinalı öğrenme, sınıflandırma ve optimizasyon gibi alanlarda başarı ile kullanılabilen bir yapay zekâ yöntemidir. Başlangıç ve sonuç parametrelerinden oluşan ANFIS parametrelerinin başarılı bir şekilde güncellenmesi ve eğitimi, problemlerin çözümünde oldukça önemlidir. Bu kapsamda, ANFIS'in eğitimi için geleneksel algoritmalar yaygın olarak kullanılmaktadır. Ancak, geleneksel algoritmalarda her adımda eğim hesaplanması nispeten zordur ve kullanılması gereken zincir kuralı çoğunlukla yerel minimuma takılma problemine yol açmaktadır. Ayrıca geleneksel algoritmalarda parametrelerin yakınsaması oldukça yavaş ve çözümün başarımı büyük oranda başlangıç değerlerine bağlı olmaktadır. Bu sebeplerden dolayı ANFIS modelinin eğitilmesi temel sorunlardan biridir. Bu konuda yapılmış muhtelif

çalışmalar ve önerilmiş yöntemler bulunmaktadır. Bu yöntemlerden bazıları yapay zekâ optimizasyon algoritmalarından olan Genetik Algoritma, Parçacık Sürüsü Optimizasyon Algoritması, Karınca Koloni Optimizasyon Algoritması ve Diferansiyel Gelişim Algoritması gibi sezgisel algoritmalardır. Bu tez çalışmasında, sınıflandırma ve tahmin amacıyla yaygın olarak kullanılan ANFIS modellerinin eğitilmesinde iteratif tabanlı benzetilmiş tavlama algoritmasının kullanılabilirliği araştırılmıştır. Önerilen yaklaşımın performansı farklı problemlere uygulanarak, geleneksel ve modern optmizasyon yöntemlerinin performanslarıyla karşılaştırılmıştır.

Dördüncü bölümde, dinamik sistemlerin kimliklendirilmesi amacıyla ANFIS modeli SA algoritması kullanılarak eğitilmiştir. Doğrusal ve doğrusal olmayan dinamik sistemlerin kimliklendirilmesine yönelik yapılan simülasyon çalışmaları sonucunda, önerilen yaklaşımın dikkate alınan BP, TS ve GA algoritmalarından ve literatürden alınan yöntemlerden çok daha başarılı sonuçlar verdiği görülmüştür. Elde edilen sonuçlardan SA algoritmasının ANFIS ağının parametre değerlerinin belirlenmesi amacıyla başarılı bir şekilde kullanılabileceği kanaatine varılmıştır. Genel olarak, belirlenmesi gereken parametre sayısının fazla olması ve ağın eğitilmesinde kullanılan türev tabanlı algoritmaların bölgesel optimuma takılma riski içermesi gibi hususlar ANFIS'in başarımını olumsuz etkileyebilen temel unsurlar arasında olduğundan, tez çalışmasının bu bölümünde, ANFIS ağının türev tabanlı algoritmalar yerine yapay zekâ optimizasyon algoritmalarından olan SA algoritması kullanılarak eğitilmesi bu olumsuzlukların üstesinden geldiğini göstermektedir. SA algoritmasının türeve dayalı algoritmalar gibi kısıtlamalar içermemesi ve problemlere kolaylıkla uygulanabilir olması nedeniyle, ANFIS ağının farklı problemlere yönelik uygulamalarında da kullanılabileceği değerlendirilmiştir.

Beşinci bölümde, mikrodizi gen ifade kanser verilerinin sınıflandırılması amacıyla ANFIS ve SA algoritmasına dayalı hibrid bir yeni yöntem sunulmuştur. Önerilen yöntemin performansı beş farklı kanser türüne ait gen verileri üzerinde test edilerek çalışmalar yapılmıştır. Ayrıca, ANFIS ağı BP, HB, GA ve PSO algoritmalarıyla da eğitilerek kanser verileri sınıflandırılmıştır. Sınıflandırma performansları önerilen metodun performansı ile karşılaştırılmıştır. Ayrıca, mikrodizi gen ifade (ekspresyon) verilerindeki gen sayılarının çokluğu ve bu veriler arasında çoğunlukla doğrusal olmayan ilişkiler bulunması gibi nedenlerden dolayı geleneksel yöntemler mikrodizi gen

problemlerinin çözümlerinde yetersiz ve sınırlı kalabilmektedir. Bu kapsamda, tez çalışmasında kullanılan mikrodizi gen ifade kanser verileri BayesNet, SVM ve J48 gibi istatistiksel yöntemler kullanılarak da sınıflandırmıştır. İstatistiksel yöntemlerden elde edilen başarım performansları da önerilen yöntem ve diğer yaklaşımların performansları ile karşılaştırılmıştır. Sınıflandırma başarım performanslarının yanı sıra istatistiksel başarım ölçütlerinden Kesinlik, Duyarlılık ve F-ölçütü değerleri de model başarımlarının hesaplanmasında önemli etkenlerdir. Bu kapsamda, sınıflandırma performansı için bu değerler de hesaplanarak tüm yaklaşımlar için karşılaştırmalar yapılmıştır. Elde edilen sonuçlardan yine önerilen yöntemin kullanılan tün yaklaşımlardan daha iyi performans sergilediği görülmüştür. Tüm bu çalışmalar önerilen yönteme olan güvenilirliği artırmıştır. Sonuç olarak, ANFIS ağının SA algoritması ile eğitilmesine dayalı yaklaşımın kanser gen verilerinin sınıflandırılmasında oldukça başarılı olduğu görülmüştür.

ANFIS modelinin yapılandırılması için genellikle tecrübeye dayalı bir uzman bilgisi kullanılmıştır. Ancak birçok durumda uzmanların bilgisi en doğru tercihleri yaparak ANFIS modeli için en uygun kriterlerin belirlenmesinde yetersiz kalabilmektedir. Farklı denemeler yaparak en uygun özelliklerin seçilmesi süreci zaman almakla birlikte uygunluğu tartışılabilmektedir. Ayrıca, bu yaklaşımın dezavantajı giriş sayısı arttıkça üyelik fonksiyonu sayısı ve kural sayısının da hızlı bir şekilde artmasıdır. Bu sayıların artması, ANFIS ağının eğitilmesini zorlaştırıp, gürültü gibi girişleri ezberlemesine neden olmaktadır. Bu süreci otomatikleştiren ve bulanık modelin performansını ve yorumlanabilirliğini geliştirirken çalışma sürecini önemli ölçüde azaltan çeşitli yöntemler bulunmaktadır. Bu yöntemlerin arkasındaki temel fikir ise, giriş-çıkış örnek verisi arasındaki ilişkiyi öğrenerek bulanık kuralları tahmin etmektir. Giriş uzayını etkili bir şekilde bölümlemek, ANFIS'in kural sayısını ve üyelik fonksiyonlarının sayısının azaltarak, modelin hem öğrenme hem de uygulanma aşamasında hızını ve performansını artırmaktadır. Altıncı bölümde, bulanık kural sayısını en aza indirgemek amacıyla kullanılan bulanık kural oluşturma tekniklerinden FCM yöntemi kullanılmıştır. FCM, ANFIS ağının kural tabanını ve üyelik fonksiyonlarını otomatik olarak oluşturmaktadır. Bu bölümde, Dinamik sistemlerin kimliklendirilmesi amacıyla önerilen FCM-ANFIS modelinin parametrelerinin eğitilmesi için SA algoritmasının kullanılmasına yönelik çalışmalar yapılmıştır. Doğrusal ve doğrusal olmayan dinamik sistemlerin kimliklendirilmesine yönelik yapılan simülasyon çalışmaları sonucunda, önerilen

yaklaşımın dikkate alınan BP, TS, GA ve PSO algoritmalarından çok daha başarılı sonuçlar verdiği görülmüştür. Bununla birlikte tecrübeye dayalı ANFIS modelleme yerine FCM tabanlı ANFIS oluşturma, modelin hem öğrenme hem de uygulanma aşamasında hızını ve performansını artırmıştır.

Yedinci bölümde yapılan çalışmada benzetilmiş tavlama algoritması için popülasyon temelli algoritmaların bir operatörü olan çaprazlama operatörü üzerine dayalı yeni bir model geliştirilmiş ve modelin performansı doğrusal ve doğrusal olmayan dinamik sistemlerin kimliklendirilmesi amacıyla ANFIS parametrelerinin eğitilmesi için kullanılmıştır. Önerilen modelin optimum veya optimuma yakın çözümleri kısa zamanda bulabilmede temel SA algoritmasına ve diğer yöntemlere göre daha yüksek performansa sahip olduğu görülmüştür. Ayrıca önerilen yöntemde düşük standart sapma değerine sahip sonuçlar elde edilmiştir. Standart sapma değerinin düşük olması da önerilen yöntemin performansıyla birlikte daha güçlü ve dinç olduğunu göstermektedir. Temel SA algoritmasının aynı süreli her koşmada aynı performansı gösterememesi onun başarısız olmasından değildir. Buradaki problem daha öncede açıklandığı gibi zamanla ilgilidir. Temel SA algoritması da her koşmada yüksek kalite değerli çözümleri bulabilecek kabiliyettedir. Fakat bu rastgele tek noktadan çözüme başlanması ve başlangıç çözümünün optimale uzaklığına bağlı olarak farklı zamanlarda gerçekleşebilmektedir. SA algoritmasına paralel çalışma özelliği kazandırılması suretiyle elde edilen sonuçlardan görüldüğü gibi, algoritmanın kısa sürede daha iyi çözümlere ulaşması her koşmada sağlanmış, böylece algoritmanın performansında büyük bir gelişme elde edilmiştir. Yapılan çalışmalar sonunda Tablo 5.1, 5.2 ve 5.3'de verilen sonuçlardan, paralel benzetilmiş tavlama algoritmasının ANFIS parametrelerini temel benzetilmiş tavlama algoritması ve diğer yöntemlerden daha başarılı optimize ettiği görülmüştür. Önerilen PSA algoritmasının SA algoritmasına göre daha yüksek performans göstermesinin çeşitli nedenleri vardır. Bunlardan birincisi, PSA algoritmasında araştırmaya çözüm uzayının farklı noktalarından paralel olarak başlanması ve her aşamada rastgele çözüm noktalarının araştırmaya katılmasıyla çözüm uzayının daha geniş bir bölgesinin taranmasıdır. İkincisi ise, çaprazlama operatörünün ve baskın aday seçiminin kullanılması ile araştırmanın ilerleyen adımlarında önceki aşamalarda bulunan sonuçların da yönlendirici bir etki yapmasıdır. Burada yeni önerilen benzetilmiş tavlama

algoritması, temel benzetilmiş tavlama algoritmasına göre çok daha yüksek bir performansa sahip olmuştur.

6.2. Gelecekte Yapılacak Çalışmalar

Temel SA algoritmasına paralel çalışma özelliği kazandırılması suretiyle geliştirilen paralel benzetilmiş tavlama algoritması ile algoritmanın kısa sürede daha iyi çözümlere ulaşması sağlanmış, böylece algoritmanın performansında büyük bir gelişme elde edilmiştir. Önerilen paralel modelin analiz edilmesi, başarımını artırmaya yönelik yeni yaklaşımların araştırılması ve farklı mühendislik problemleri üzerindeki performansının incelenmesi gelecek çalışmalar olarak düşünülmektedir.

KAYNAKLAR

1. Civalek, Ö., Ülker, M., 2004. Dikdörtgen plakların doğrusal olmayan analizinde yapay sinir ağı yaklaşımı. **IMO Teknik Dergi, 213: 3171-3190.**
2. Simon, D., 2002. Training fuzzy systems with the extended Kalman Filter. **Fuzzy Sets and Systems, 132 (2): 189–199.**
3. Surmann, H., 1996. Genetic optimization of a fuzzy system for charging batteries. **IEEE Transactions on Industrial Electronics, 43** (5): 541–548.
4. Ghomsheh, V. S., Shoorehdeli, M. A., Teshnehlab, M., 2007. Training ANFIS structure with modified PSO algorithm, pp. 1-6. *In Control and Automation, Med'07. Mediterranean Conference on IEEE*.
5. Shoorehdeli, M. A., Teshnehlab, M., Sedigh, A., 2009. Training ANFIS as an identifier with intelligent hybrid stable learning algorithm based on particle swarm optimization and extended Kalman filter. **Fuzzy Sets and Systems, 160**: 922–948.
6. Turki, M., Bouzaida, S., Sakly, S., M'Sahli, F., 2012. Adaptive control of nonlinear system using neuro-fuzzy learning by PSO algorithm, pp. 519-523. *In Electrotechnical Conference (MELECON), 16th IEEE Mediterranean, March 2012.*
7. Carrano, E. G., Takahashi, R. H. C., Caminhas, W. M., Neto, O. M., 2008. A genetic algorithm for multiobjective training of ANFIS fuzzy networks, pp. 3259–3265. *IEEE Congress on Evolutionary Computation.*
8. Cus, F., Balic, J., Zuperl, U., 2009. Hybrid ANFIS-ants system based optimisation of turning parameters. **Journal of Achievements in Materials, 36** (1): 79-86.
9. Karaboga, D., Kaya, E., 2013. Training ANFIS using artificial bee colony algorithm, pp. 1-5. *Innovations in Intelligent Systems and Applications (INISTA), IEEE International Symposium.*
10. Juang, C., 2010. Combination of Particle Swarm and Ant Colony Optimization Algorithms for Fuzzy Systems Design, pp. 216. *In:* Fuzzy Systems (Eds. A. T., Azar). INTECH, Crotia.
11. Liu, P., Leng, W., Fang, W., 2013. Training ANFIS model with an Improved Quantum-Behaved Particle Swarm Optimization Algorithm. **Mathematical Problems in Engineering, 2013** (2013): 10 pp.

12. Loganathan, C., Girija, K. V., 2014. Investigations on Hybrid Learning in ANFIS. **Journal of Engineering Research and Applications, 4** (10): 31–37.
13. Behmanesh, M., Mohammadi, M., Naeini, V. S., 2014. Chaotic time series prediction using improved ANFIS with Imperialist Competitive Learning Algorithm. **International Journal of Soft Computing and Engineering, 4** (4), 25–33.
14. Salleh, N. M., Hussain, K., 2016. Accelerated Mine Blast Algorithm for ANFIS training for solving classification problems. **International Journal of Software Engineering and Its Applications, 10** (6), 161–168.
15. Marzi, H., Darwish, A. H., Helfawi, H., 2016. Training ANFIS using The Enhanced Bees Algorithm and Least Squares Estimation. **Intelligent Automation & Soft Computing.**
16. Cárdenas, J. J., García, A., Romeral, J. L., Kampouropoulos, K., 2011. Evolutive ANFIS training for energy load profile forecast for an IEMS in an automated factory. *IEEE International Conference on Emerging Technologies and Factory Automation, ETFA.*
17. Abdulshahed, A.M., Longstaff, A.P., Fletcher, S., Myers, A., 2015. Thermal error modelling of machine tools based on ANFIS with fuzzy c-means clustering using a thermal imaging camera. **Applied Mathematical Modelling, 39** (7): 1837-1852.
18. Al-Himyari, B. A., Yasin, A., Gitano, H., 2014. Adaptive neuro-fuzzy model with fuzzy clustering for nonlinear prediction and control. **Asian Journal of Applied Sciences, 2** (3).
19. Qintero, O.L., Uribe, M.G., Vasquez, A.P., 2014. Colombian Energy Market: An approach of Anfis and Clustering Techniques to an Optimal Portfolio.
20. Mehrabi, M., Sharifpur, M., Meyer, J.P., 2012. Application of the FCM-based neuro-fuzzy inference system and genetic algorithm-polynomial neural network approaches to modelling the thermal conductivity of alumina-water nanofluids. **International Communications in Heat and Mass Transfer, 39** (7): 971-977.
21. Karahoca, A., Karahoca, D., 2011. GSM churn management by using fuzzy c-means clustering and adaptive neuro fuzzy inference system. **Expert Systems with Applications, 38** (3): 1814-1822

22. Pinpin, L.K., Gamarra, D.F.T., Laschi, C., Dario, P., 2008. Topology based fuzzy clustering for robust ANFIS creation, pp. 1-6. *2008 7th IEEE International Conference on Cybernetic Intelligent Systems*.

23. Rantala, J., Koivisto, H., 2002. Optimised subtractive clustering for neuro-fuzzy models.

24. Tawafan, A.H., Sulaiman, M., Ibrahim, Z., 2012. Adaptive neural subtractive clustering fuzzy inference system for the detection of high impedance fault on distribution power system. **International Journal of Artificial Intelligence, 1** (2): 63-72.

25. Berneti, S.M., 2011. Design of fuzzy subtractive clustering model using particle swarm optimization for the permeability prediction of the reservoir. **International Journal of Computer Applications, 29** (11)

26. Gogoi, R., Sarma, K.K., 2012. ANFIS based information extraction using K-means clustering for application in satellite images. **International Journal of Computer Applications, 50** (7)

27. Yılmaz, M., Arslan, E., 2005. Bulanık mantığın jeodezik problemlerin çözümünde kullanılması, pp. 512-522. *2. Mühendislik Ölçmeleri Sempozyumu*.

28. Melek, W.W., Sadeghian, A., 2009. A theoretic framework for intelligent expert systems in medical encounter evaluation. **Exoert Systems, 26** (1): 82-99.

29. Türkbey, O., 2003. Çok amaçlı makine sıralama problemi için bir bulanık güçlü metot. **Dokuz Eylül Üniversitesi Fen ve Mühendislik Dergisi, 5** (3): 81-98.

30. Sarı, N., Arslan, E., 2007. Geoit yüksekliğinin ANFIS ile adım adım hesaplanması. **Jeodezi, Jeoinformasyon ve Arazi Yönetimi Dergisi, 96**: 31-38.

31. Zadeh, L. A., 1965. Fuzzy Sets. **Information and Control, 8**: 338-352.

32. Jang, J.S.R., Sun, C.T., Mizutani, E., 1997. Neuro-Fuzzy and Soft Computing. PrenticeHall, 607 pp.

33. Mamdani, E.H., Assilian, S., 1975. An experiment in linguistic synthesis with a fuzzy logic controller. **Int. Journal of Man-Machine Studies, 7** (1): 1-13.

34. Akyılmaz, O., 2005. Esnek Hesaplama Yöntemlerinin Jeodezide Uygulamaları. İ.T.Ü. Fen Bilimleri Enstitüsü, Doktora tezi, İstanbul.

35. Palancıoğlu, H. M., Beşdok, E., 2005. Hareketli objelerin ulaşım ağındaki hareketlerinin bulanık mantık metotları ile modellenmesi, *2. Mühendislik Ölçmeleri Sempozyumu*, 23-25 Kasım, 2005, İTÜ- İstanbul.

36. Özçalık, H., Uygur, A., 2003. Dinamik sistemlerin uyumlu sinirsel – bulanık ağ yapısına dayalı etkin modellenmesi. **KSÜ Fen ve Mühendislik Dergisi, 6** (1): 36-46.

37. Avcı, E., Akpolat, Z. H., 2002. Uyarlamalı ağ tabanlı bulanık çıkarım sistemi ile doğru akım motorlarının hız denetimi, pp. 193-196. *ELECO'2002 Elektrik-Elektronik-Bilgisayar Mühendisliği Sempozyumu*, Aralık, 2002, Bursa.

38. Özgan, E., Kap, T., Beycioğlu, A., Emiroğlu, M., 2009. Asfalt betonunda marshall stabilitesinin uyarlamalı sinirsel bulanık mantık yaklaşımı ile tahmini, *Uluslararası İleri Teknolojiler Sempozyumu,* Karabük.

39. Güney, K., Sarıkaya, N., 2008. Dairesel mikroşerit antenlerin yama yarıçapının çeşitli algoritmalarla optimize edilen bulanık mantık sistemine dayalı uyarlanır ağlar ile hesaplanması. *ELECO'2008 Elektrik-Elektronik-Bilgisayar Mühendisliği Sempozyumu ve Fuarı*, 26-30 Kasım, 2008, Bursa.

40. Jang, J. S. R., 1993. ANFIS: Adaptive network-based fuzzy inference systems. **IEEE Trans. Syst., Man. and Cybern, 23**: 665-685.

41. Franklin, G. F., Powell, J. D., Workman, M. L., 1990. Digital Control of Dynamic Systems. Addison Wesley, 132 pp.

42. Haznedar, B., Kalinli, A., 2016. Training ANFIS using genetic algorithm for dynamic systems identification, pp. 22-26. *International Conference on Advanced Technology & Sciences (ICAT'16)*, 1-3 Eylül, 2016, Konya, Turkey.

43. Hímer, Z., Kovács, J., Benyó, I., Kortela, U., 2004. Neuro- fuzzy modelıng and genetic algoritms optimization for flue gas oxygen contol, *2nd IFAC Workshop on Advanced Fuzzy/Neural Control*, University of Oulu.

44. Jang, J. S. R., Sun, C. T., 1995. Neuro-Fuzzy modeling and control. **proceedings of the IEEE, 83**: 3.

45. Karaboga, D., Kaya, E., 2016. An adaptive and hybrid artificial bee colony algorithm (aABC) for ANFIS training. **Applied Soft Computing, 49**: 423-436.

46. Hocaoğlu, F., Kurban, M., 2005. Adaptif ağ tabanlı bulanık mantık çıkarım sistemi ile Eskişehir bölgesi için güneşlenme süreleri tahmini, *Elektrik-Elektronik-Bilgisayar Mühendisliği 11. Ulusal Kongresi ve Fuarı*, 22-25 Eylül, 2005, İstanbul.

47. Fırat, M., Yurdusev, M., Mermer, M., 2008. Uyarlamalı sinirsel bulanık yaklaşımı ile aylık su tüketiminin tahmini. **Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi, 23** (2): 449-457.

48. Günay, E., Alçı, M., Parmaksızoğlu, S., 2006. Çoklu kaotik çeker üreten DK-HSA devresinde parçalı – doğrusal fonksiyon karakteristiğinin bulanık sinir ağı kontrolü ile elde edilmesi, *URSI-Türkiye 2006 Bilimsel Kongresi ve Ulusal Genel Kurul Toplantısı*, 6-8 Eylül, 2006, Hacettepe Üniversitesi, Ankara.

49. Civelekoğlu, G., 2006. Arıma Proseslerinin Yapay Zeka ve Çoklu İstatiksel Yöntemler İle Modellenmesi, Çevre Mühendisliği Anabilim Dalı, Doktora Tezi, Isparta.

50. Haznedar, B., Kalinli, A., 2016. Detection of the relationship between thrombophilia disease with genetic disorders by adaptive neuro-fuzzy inference system (ANFIS), **Sakarya University Journal Of Science, 20** (1): 13-21.

51. Akyılmaz, O., Ayan, T., Özlüdemir, M. T., 2003. Geoid surface approximation by using Adaptive Network based Fuzzy Inference Systems. **AVN, 8** (9): 22-27.

52. Chen, M.-S., Wang, S.-W., 1999. Fuzzy clustering analysis for optimizing fuzzy membership functions. **Fuzzy Sets and Systems, 103** (2): 239–254.

53. Hamdan, H., 2013. An Exploration of the Adaptive Neuro-Fuzzy Inference System (ANFIS) in Modelling Survival, The University of Nottingham, Doktora Tezi.

54. Jang, J.S.R., Sun, C.T., 1995. Neuro-fuzzy modeling and control. **In The Proceeding of the IEEE, 83**: 378–406.

55. Chiu, S.L., 1994. Fuzzy model identification based on cluster estimation. **J. Intell. Fuzzy Syst., 2** : 267-278.

56. Chiu, S.L., 1997. Extracting Fuzzy Rules from Data for Function Approximation and Pattern Classification, pp. 149-162. *In:* Chapter 9 in Fuzzy Information Engineering: a Guide Tour of Applications, John Wiley & Sons, New York.

57. Chen, M.Y., Linkens, D.A., 2001. A systematic neuro fuzzy modeling framework with application to material property prediction. **IEEE Transactions on Systems, Man, and Cybernetics, Part B, 31**: 781–790.

58. Nauck, D., Kruse, R., 1999. Obtaining interpretable fuzzy classification rules from medical data. **Artificial Intelligence in Medicine, 16**: 149 – 169.

59. Bezdek, J.C., 1981. Pattern Recognition With Fuzzy Objective Function Algorithms, Plenum Press, New York.

60. Shoorehdeli, M.A., Teshnehlab, M., Sedigh, A.K., 2006. A novel training algorithm in ANFIS structure. *American Control Conf.*

61. Shoorehdeli, M.A., Mahdi, Teshnehlab, M., Sedigh, A.K., 2009. Identification using ANFIS with intelligent hybrid stable learning algorithm approaches. **Neural Computing and Applications, 18** (2): 157-174.

62. Shoorehdeli, M.A., Teshnehlab, M., Sedigh, A.K., 2007. A novel hybrid learning algorithm for tuning ANFIS parameters using adaptive weighted PSO. *IEEE Int Fuzzy Sys Conf.*

63. Rumelhart, D., et al., 1986. Learning representations by backpropagation errors**.** **Nature, 323** (9): 533-536.

64. Öztürk, C., 2011. Yapay Sinir Ağlarının Yapay Arı Koloni Algoritması ile Eğitilmesi, Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, Kayseri.

65. Nayak, P.C., et al, 2004. A Neuro Fuzzy computing technique for modeling hydrological time series. **Journal of Hydr., 29**: 52–66.

66. Firat, M., Güngör, M., 2007. River Flow Estimation using Adaptive Neuro-Fuzzy inference System. **Mathematics and Computers in Simulation, 75** (3-4): 87-96.

67. Mitchell, M., 1998. An Introduction to Genetic Algorithms. MIT Press.

68. Kervan, M., 2009. Çoklu Sensör Konumlandırma Probleminin Genetik Algoritmalar ve Gen Havuzu Tabanlı Genetik Agoritmalar ile Çözülmesi, Hava Harp Okulu Komutanlığı, 124s.

69. Haznedar, B., 2010. Yapay Zeka Teknikleri Kullanılarak Erkek Kısır Hastalarda Genetik Bozukluk Varlığının Tespiti, Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 105s.

70. Kennedy, J., Eberhart, R., 1995. Particle swarm optimization, pp. 1942-1948. *In Proceedings of IEEE International Conference on Neural Networks*, Piscataway, NJ.

71. Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by Simulated Annealing, **Science, 220**: 671-680.

72. Avcı, Y., 2008. Sonlu Eleman Modeli Güncellemesi Tekniğinde Benzetilmiş Tavlama Algoritması Kullanılarak Mekanik Sistemlerde Hasar Tespiti. İTÜ Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, İstanbul, 64s.

73. Dowsland, K.A., 1995. Simulated Annealing. *In:* Modern Heuristic Techniques For Combinatorial Problems. (Editor, Reeves, C.R.), McGraw-Hill.

74. Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., Teller, E., 1953. Equation of the state calculation by fast computing machines. **Journal of Chemical Physics, 21**: 1087-1092.

75. Eglese, R.W., 1990. Simulated Annealing: A Tool for Operational Research. **European Journal of Operational Research, 34**: 600-612.

76. Jhonson, S., Aragon, C., Mccgeoch, L., Schevon, C., 1989. Optimization By Simulated Annealing: An Experimental Evaluation, Part-I, Graph Partitioning, **Operations Research, 7** (6): 865-892.

77. Kalinli, A., 2003. Training Elman network using simulated annealing algorithm. **Journal of the Institute of Science and Technology of Erciyes University, 19**: 28-37.

78. Ergün, K., 2004. Kesme ve Paketleme Problemleri ve Araştırmaya Yönelik Bir Metot Geliştirilmesi ve Bu Metodun Etkinliğinin Sınanması. Balıkesir Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Balıkesir.

79. Wong, K.P., Steven, R.M., Feng, D., Fulham M.J., 2002. Estimation of input function and kinetic parameters using simulated annealing: application in a flow model. **IEEE Transactions on Nuclear Science, 49**: 707-713.

80. Mottershead, J.E., Friswell, M.I., 1993. Model updating in structural dynamics: a survey. **Journal of Sound and Vibration, 167**: 347-375.

81. Arıöz, U., Tıbbi Bilişim, Hacettepe Üniversitesi Tıp Fakültesi Tıp Bilişimi A.D., **http://www.medinfo.hacettepe.edu.tr/ders/TR/D1/3/5221_files/frame.htm.**

82. Saka, O., 2010. Tıp Bilişimi, Gazi Üniversitesi, Ankara, **http://www.ttb.org.tr/udek/images/stories/file/bilisim_calistay/osaka.ppt.**

83. Saka, O., Tıp Bilişimi Nedir, Tıp Bilişim Derneği, İzmir, **http://www.turkmia.org/files/48.ppt**.

84. Özata, M., Arslan, S., 2004. Klinik Karar Destek Sistemleri ve Örnek Uygulamalar.

85. Lisboa, P.J., Taktak, A.F.G., 2006. The use of artificial neural networks in decision support in cancer: A systematic review. **Elsevier Neural Networks, 19**: 408-415.

86. Zini, G., D'onofrio, G., 2003. Neural network in hematopoietic malignancies. **Elsevier, Clinica Chimica Acta, 333**: 195-201.

87. Belacel, N., Boulassel, M., R., 2000. Multicriteria fuzzy assignment method: a useful tool to assist medical diagnosis. **Elsevier, Artificial Intelligence In Medicine, 581**: 1-7.

88. Arthur, G. L., 2009. Implementation Of A Fuzzy Rule-Based Decision Support System For The Immunhistochemical Diagnosis Of Small B-Cell Lymphomas, PhD Thesis, University of Missouri.

89. Hong, J.H., Cho, S.B., 2004. Lymphoma Cancer Classification Using Genetic Programming with SNR Features, pp. 78-88. M. Keijzer et al.(Eds.):EuroGP, LNCS3003.

90. Vanneschi, L., Archetti, F., Castelli, M., Giordani, I., 2009. Classificatin of oncologic data with genetic programming. **Hindawi Publishing Corporation Journal of Artificial Evolution and Applications,** 13pp.

91. Tung, W.L., Quek, C., 2005. GenSo-FDSS: A neural-fuzzy decision support system for pediatric all cancer subtype identification using gene expression data. **Elsevier, Artificial Intelligence in Medicine, 33**: 61-88.

92. Paul, T.K., Iba, H., 2005. Gene selection for classification of cancers using probabilistic model building genetic algorithm. **Elsevier, BioSystems, 82**: 208-225.

93. Corchado, J.M., Paz, J.F., Rodriquez, S., Bajo, J., 2009. Model of experts for decision support in the diagnosisof leukemia patients. **Elsevier, Artificial Intelligence in Medicine, 46**: 179-200.

94. Silva A., Cortez, P., Santos, M.F., Gomes, L., Neves, J., 2006. Mortality assessment in intensive care units via adverse events using artificial neural networks. **Elsevier, Artificial Intelligence in Medicine, 36**: 223-234.

95. Ramesh, AN., Kambhampati C., Monson, JRT., Drew PJ., 2004. Artificial intelligent in medicine. **The Royal Colllege of Surgeons of England, 86**: 334-338.

96. Muselli, M., Costacurta, M., Ruffino, F., 2009. Evaluating switching neural networks through artificial and real gene expression data. **Elsevier, Artificial Intelligence in Medicine, 45**: 163-171.

97. Razzouk, D., Mari, J.J., Shirakawa, I., Wainer, J., Sigulem, D., 2006. Decision Support System for the Diagnosis of Schizodhrenia Disorders**. Braz J Med Biol Res, 39**: 119-128.

98. Melek, W.W., Sadeghian, A., 2009. A theoretic framework for intelligent expert systems in medical encounter evaluation. **Expert Systems, 26** (1).

99. Matyja, D., 2007. Applications of Data Mining Algorithms to Analysis of Medical Data. Master Thesis, Blekinge Tekniska Högskola.

100. Aftarczuk, K., 2007. Evaluation of Selected Data Mining Algorithms Implemented in Medical Decision Support Systems. Master Thesis, Blekinge Tekniska Högskola.

101. Buisson J.C., 2008. Nutri-Educ, a nutrition software application for balancing meals, using fuzzy arithmetic and heuristic search algorithms. **Elsevier, Artificial Intelligence in Medicine, 42**: 213-227.

102. Vassilakis, K.M., Micheloyannis, S., Marakakis, M., Decision Support Systems (DDS) as Tools to Diagnose Epilepsies in Children, NoE program BIOPATTERN.

103. Chiu, S., Chen C., Lin, T., 2008. Using support vector regression to model the correlation between the clinical metastases time and gene expression profile for breast cancer. **Elsevier, artificial Intelligence in Medicine, 44**: 221-231.

104. Ellenus, J., Groth, T., 2008. Dynamic decision support graph-visualization of ANN-generated diagnostic indications of pathological conditions developing over time. **Elsevier, Artificial Intelligence in Medicine, 42**: 189-198.

105. Chen, Z., Li, J., Wei, L., 2007. A multible kertel support vector machine scheme for feature selection and rule extraction from gene expression data of cancer tissue. **Elsevier, Artificial Intelligence in Medicine, 41**: 161-175.

106. Türker, N., Tokan, F., Yıldırım, T., 2005. kalp hastalığı teşhisinde yapay sinir ağlarının performansının roc analizi ile belirlenmesi. *National Symposium on Biomedical Engineering*, İstanbul.

107. Anagnostou, T., Remzi, M., Lykourinas, M., Djavan, B., 2003. Artificial neural networks for decision-making in urologic oncology. **European Urology, 43**: 596-603.

108. Asl, B.M., Setarehdan, S.K., Mohebbi, M., 2008. Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal**.** **Elsevier, Artificial Intelligence in Medicine, 44**: 51-64.

109. Ceylan, M., Ceylan, R., Özbay, Y., Kara, S., 2008. Application of complex discrete wavelet transform in classification of doppler signals using complex-valued artificial neural network. **Elsevier, Artificial Intelligence in Medicine, 44**: 65-76.

110. Remzi, M., Djavan, B., 2004. Artificial neural networks in urology. **European Urology Supplements, 3**: 33-38.

111. Polak, S., Mendyk, A., 2008. Artificial neural networks based internet hypertension prediction tool development and validation. **Elsevier, Applied Soft Computing, 8**: 734-739.

112. Remzi, M., Waldert, M., Djavan, B., 2005. Preoperative nomograms and artificial neural networks (anns) for ıdentification of surgical candidates. **EAU Update Series, 3**: 63-71.

113. Küçükkaya, RD., Aydın, M., 2006. Trombofili Genetiği, Türk Hematoloji Derneği - Moleküler Hematoloji, Klinik Hematoloji için Pratik Genetik Yaklaşım Kursu, pp. 39-43. 11 Kasım, 2006, Antalya.

114. Makris, M., Rosendaal, F.R., Preston, F.E., 1997. Familial thrombophilia: genetic risk factors and management. **J Int Med, 242**: 9-15.

115. Kolodziej, M., Comp, P.C., 1993. Hypercoagulable states due to natural anticoagulant deficiencies. **Curr Op Hematol, 301**.

116. Rodgers, G.M., 1998. Thrombosis and antithrombotic therapy, pp. 1781-1818. *In*: Wintrobe's Clinical Hematology. 10th Edition. Williams and Wilkins Comp.

117. Bauer, K.A., 1995. The Hypercagulable state, pp. 1531-1549. *In*: Williams Hematology 5th edition. McGraw Hill inc.

118. Lane, D.A., Manucci, P.M., Bauer, K.A., 1996. Inherited thrombophilia: Part-2. **Thrombosis and Haemostasis, 76** (6): 824-834.

119. Bick, R.L, Kaplan, H., 1998. Syndromes of thrombosis and hypercoagulability. **Med Clin North Am, 82** (3): 409-458.

120. Hultin, M.B., 1995. Antithrombin III assays, pp. 1-101. *In*: Williams Hematology 5th edition. McGraw Hill inc.

121. Comp, P.C., 1995. Protein C and protein S, 99pp. *In*: Williams Hematology 5th edition. McGraw Hill inc.

122. Welch, G.N., Loscalzo, J., 1998. Homocysteine and atherothrombosis. **N Engl J Med, 338**: 1042-1050.

123. D'Angelo, A., Selhub, J., 1997. Homocysteine and thrombotic disease. **Blood 90**: 1-11.

124. Koster, T., Blann, A.D., Briet, E., Vandenbrouke, J.P., Rosendaal, F.R., 1995. Role of clotting factor VIII in effect of von Willebrand factor on occurrence of deep venous thrombosis. **Lancet, 345**: 152-5.

125. Kamphuisen, P.W., Houwing-Duistermaat, J.J., van Hauwelingen, J.C., 1998. Familial clustering of factor VIII and von Willebrand factor levels. **Thromb Haemost, 79** (2): 323-7.

126. De Stefano, V., Finazzi, G., Mannucci, M., 1996. Inherited thrombophilia: Pathogenesis, clinical syndromes and management. **Blood, 87**: 3531-3544.

127. Lane, D.A., Manucci, P.M., Bauer, K.A., 1996. Inherited thrombophilia: Part-1. **Thrombosis and Haemostasis**, **76** (5): 651-62.

128. Dahlback, B., Carlsson, M., Svensson, P.J., 1993. Familial thrombophilia due to a previously unrecognised mechanism characterized by poor anticoagulant response to activated protein C: Prediction of a cofactor to activated protein C. **Proc Natl Acad Sci USA, 90** (3): 1004-8.

129. Bertina, R.M., Koeleman, B.P.C., Koster, T., Rosendaal, F.R., Dirven, R.J., de Ronde, H., van der Velden, P.A., Reitsma, P.H., 1994. Mutation in blood coagulation factor V associated with resistance to activated protein C. **Nature, 369**: 64-7.

130. Falcon, C.R., Cattaneo, M., Panzeri, D., Martinelli, I., Manucci, P.M., 1994. High prevalanve of hyperhomocyst(e)inemia in patients with juvenile venous thrombosis. **Arterioscler Thromb, 14**: 1080.

131. Poort, S.R., Rosendaal, F.R., Reitsma, P.H., Bertina, R.M., 1996. A common genetic variation in the 3'-untranslated region of the protrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. **Blood, 88** (10): 3698-703.

132. Rosendaal, F.R., 1999. Venous thrombosis: a multicausal disease. Lancet, 353: 1167-73.

133. Wu, O., Robertson, L., Twaddle, S., Lowe, G., Clark, P., Walker, I., Brenkel, I., Greaves, M., Langhorne, P., Regan, L., Greer, I., 2006. Screening for thrombophilia in high-risk situations: a meta analysis and cost effectiveness analysis. **Br J Haematol, 131**: 80-90.

134. Reitsma, P.H., Rosendaal, F.R., 2007. Past and future of genetic research in thrombosis, pp.264–269. J Thromb Haemost; 5 (Suppl. 1).

135. Yang, Q., Khoury, M.J., Botto, L., Friedman, J.M., Flanders, W.D., 2003. Improving the Prediction of Complex Diseases by Testing for Multiple Disease-Susceptibility Genes. **Am. J. Hum. Genet., 72** (3): 636–649.

136. Ganesh, S.K., Sharma, Y., Dayhoff, J., Fales, H.M., Eyk, J.V., Kickler, T.S., Billings, E.M., Nabel, E.G., 2007. Detection of Venous Thromboembolism by Proteomic Serum Biomarkers. **PLoS ONE, 2** (6).

137. Baglin, T., 2011. Using the laboratory to predict recurrent venous thrombosis. **International Journal of Laboratory Hematology, 33**: 333–342.

138. de Haan, H.G., et al., 2012. Multiple SNP testing improves risk prediction of first venous thrombosis. **Thrombosis And Hemostasis, 120** (3): 656-663.

139. Souto, J.C., Soria, J.M., 2012. Predicting individual risk of venous thrombosis. **Blood, 120** (3): 500-501.

140. Penco, S., et al., 2005. Assessment of the role of genetic polymorphism in venous thrombosis through artificial neural networks. **Ann Hum Genet, 69**: 693-706.

141. Shanthi, D., Sahoo, G., Saravanan, N., 2009. Designing an Artificial Neural Network Model for the Prediction of Thrombo-embolic Stroke. **International Journals of Biometric and Bioinformatics (IJBB), 3** (1).

142. Dahabiah, A., et al., 2006. Comparative neural network based venous thrombosis echogenicity and echostructure characterization using ultrasound images, pp. 992 – 997. *Information and Communication Technologies, ICTTA '06. 2nd*.

143. Markus, H., Cullinane, M., Reid, G., 1999. Improved automated detection of embolic signals using a novel frequency filtering approach. **Stroke, 30** (8): 1610-1615.

144. Keme´ny, V., et al., 1998. Automatic embolus detection by a neural network**.** **Stroke, 30** (4): 807-810.

145. Kurtulmuş, Y., Demircan, E., Tütüncüler, F., Eşsizoğlu, E., Türkün, H., Duman, C., Bakar, C., 2013. Demir düzeyi ve iskemi modifiye albümin. **Smyrna Tıp Dergisi**, 30-35.

146. Uysal, Z., Çulha, V., Cin, Ş., 1998. Demir eksikliği tanısında serum transferrin reseptörü. **Ankara Üniversitesi Tıp Fakültesi Dergisi, 51** (3).

147. Onyesom, I., Opajobi, A.O., Uzuegbu, U.E., Oriero, D., Mordi, J., Awhin, P.E, Timothy, E., 2009. Relationship between placental alkaline phosphatase activity and cord blood glucose, albumin and neonatal birth weight at term. **Invest Clin, 50** (4).

148. Rao, G.M., Morghom, L.O., 1986. Correlation between serum alkaline phosphatase activity and blood glucose levels. **Enzyme, 35** (1).

149. Webb-Robertson, B.M., Wiberg, H.K., Matzke, M., Brown, J.N., Wang, J., McDermott, J.E., Smith, R.D., Rodland, K.D., Metz, T.O., Pounds, J.G., Waters, K.M., 2015. Review, evaluation, and discussion of the challenges of missing value imputation for mass spectrometry-based label-free global proteomics. **Journal of Proteome research, 14** (5): 1993-2001.

150. Cheng, B.Y., Carbonell, J.G., Klein-Seetharaman, J., 2005. Protein classification based on text document classification techniques. **PROTEINS, 58** (4): 955-970.

151. Kuhn, M., 2008. Building predictive models in R using the caret Package. **Journal of Statistical Software, 28** (5).

152. Wu, X., Srihari, R., 2003. New v-Support Vektor Machines and their Sequential Minimal Optimization, *Proceedings of the 20. International Conference on Machine Learning (ICML-2003)*, Washington DC.

153. Perrot, M.H., Cohen, R.J., 1996. An efficient approach to arma modelling of biolagical systems with multiple inputs and delays. **IEEE Transactions on Biomedical Engineering, 43** (1): 1-14.

154. Ljung, L., 1987. System Identification: Theory for The User. Prentice Hall.

155. Candy, J.V., 1989. Signal Processing: The Model-Based Approach. Mc Graw Hill.

156. Ljung, L., 1987. System Identification. Prentice-Hall Englewood Cliffs, NJ.

157. Özçelik, Y., 2007. Farksal Evrim Algoritması Kullanarak Sistem Kimliklendirme. Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Kayseri.

158. Karaboga, D., Kaya, E., 2014. Training ANFIS using artificial bee colony algorithm for nonlinear dynamic system identification, pp. 493-496. SIU.

159. Shoorehdeli, M.A., Teshnehlab, M., Sedigh, A.K., Khanesar, M.A., 2009. Identification using ANFIS with intelligent hybrid stable learning algorithm approaches and stability analysis of training methods. **Applied Soft Computing, 9** (2): 833–850.

160. Rumelhart, D.E., McClelland, J.L., 1986. Explorations in the Micro-Structure of Cognition, Parallel Distrubuted Processing 1. Cambridge, MA: MIT Press.

161. Juang, C.F., 2002. A TSK-type recurrent fuzzy network for dynamic systems processing by neural network and genetic algorithms. **IEEE Transactions on Fuzzy Systems, 10**(2): 155–170.

162. Kalinli, A., 2012. Simulated annealing algorithm-based Elman network for dynamic system identification. **Turkish Journal of Electrical Engineering & Computer Sciences, 20** (4): 569-582.

163. Kalinli, A., Karaboga, D., 2004. Training recurrent neural networks by using parallel tabu search algorithm based on crossover operation. **Engineering Applications of Artificial Intelligence, 17**: 529-542.

164. Pham, D.T., Liu, X., 1992. Dynamic system identification using partially recurrent neural networks. **Journal of Systems Engineering, 2** (2): 90–97.

165. Özcan, G. S., 2014. Bütünleştirici Modül Ağlarıyla Gen Düzenleme Analizi. Başkent Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Ankara, 62 s.

166. Lüleyap, H. Ü., 2008. Moleküler Genetiğin Esasları. Nobel Kitabevi, 437 s.

167. Bal, S. H., Budak, F., (2012). Mikroarray teknolojisi. **Uludağ Üniversitesi Tıp Fakültesi Dergisi, 38** (3): 227-233.

168. Şimşek, D. Ö., 2013. Mikroarray teknolojisi ve diş hekimliğinde kullanımı. **Atatürk Üniversitesi Diş Hekimliği Fakültesi Dergisi, 2013** (7).

169. Shakya, K., Ruskin, H. J., Kerr, G., Crane, M., Becker, J., 2010. Comparison of microarray preprocessing methods. **Advances Experimental Medicine and Biology, 680:** 139-147.

170. Liu, H., Bebu, I., Li, X., 2010. Microarray probes and probe sets. **Frontiers in bioscience (Elite edition), 2**: 325-338.

171. İpekdal, K., 2011. Microarray teknolojisi. (Web sayfası: **http://yunus.hacettepe.edu.tr/~mergen/sunu/s_mikroarrayandecology.pdf**) (Erişim Tarihi: Nisan 2015)

172. Bier, F. F., von Nickisch-Rosenegk, M., Ehrentreich-Förster, E., Reiß, E., Henkel, J., Strehlow, R., Andresen, D., 2008. DNA microarrays, pp. 433-453. *In*: Biosensing for the 21st Century (Eds: F. F., Bier, M., von Nickisch-Rosenegk, E., Ehrentreich-Förster, E., Reiß, J., Henkel, R., Strehlow, D., Andresen). Springer Berlin Heidelberg.

173. Schulze, A., Downward, J., 2001. Navigating gene expression using microarrays. **Nature cell biology, 3** (8): 190-195.

174. Coşgun, E., Kaya, A., Keçeli, A. S., Aksarı, Y., Karaağaoğlu, E., 2011. Gene 3E: Genetik Veri Madenciliği için Yeni Analiz Aracı.

175. Arslan, M.T., 2016. Yapay Zeka Tabanlı Yöntemler Kullanılarak Mikroarray Gen Verilerinin Sınıflandırılması, Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Kayseri, 109s.

176. Pirooznia, M., Yang, J. Y., Yang, M. Q., Deng, Y., 2008. A comparative study of different machine learning methods on microarray gene expression data. **BMC genomics, 9** (1).

177. Sarhan, A. M., 2009. Cancer classification based on microarray gene expression data using DCT and ANN. **Journal of theoretical and applied information technology, 6**(2): 208-216.

178. Loganathan, C., Girija, K. V., 2013. Cancer Classification using Adaptive Neuro Fuzzy Inference System with Runge Kutta Learning. **International Journal of Computer Applications, 79** (4).

179. AnandaKumar, K., Punithavalli, D. M., 2011. Efficient cancer classification using Fast Adaptive Neuro-Fuzzy Inference System (FANFIS) based on statistical techniques. **International Journal of Advanced Computer Science and Applications**, 132-137.

180. Arslan, M.T., 2016. Yapay Zeka Tabanlı Yöntemler Kullanılarak Mikroarray Gen Verilerinin Sınıflandırılması, Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, Kayseri, 126s.

181. Gordon, G.J., Jensen, R.V., Hsiao, L.L., Gullans, S.R., Blumenstock, J.E., Ramaswamy, S., Richards, W.G., Sugarbaker, D.J., Bueno, R., 2002. Translation

of microarray data into clinically relevant cancer diagnostic tests using gene expression ratios in lung cancer and mesothelioma. **Cancer Res. 62**: 4963–4967.

182. (Gordon-2002) Lung Cancer Dataset. 2013. (Web Page: **http://bioinformatics.rutgers.edu/Static/Supplements/CompCancer/Affymetrix/gordon-2002/gordon-2002_database.txt**) (Erişim Tarihi: Temmuz 2016).
183. Nutt, C.L., et al., 2003. Gene expression-based classification of malignant gliomas correlates better with survival than histological classification. **Cancer Res., 63** (7): 1602–1607.
184. (Nutt-2003-v2) Brain Cancer Dataset. 2013. (Web Page: **http://bioinformatics.rutgers.edu/Static/Supplements/CompCancer/Affymetrix/nutt-2003-v2/nutt-2003-v2_database.txt)** (Erişim Tarihi: Temmuz 2016).
185. Pomeroy, S.L., et al., 2002. Gene Expression-Based Classification and Outcome Prediction of Central Nervous System Embryonal Tumors. **Nature, 415** (6870): 436–442.
186. (Pomeroy-2002-v1) Central Nervous System Cancer Dataset. 2013. (Web Page: **http://bioinformatics.rutgers.edu/Static/Supplements/CompCancer/Affymetrix/pomeroy-2002-v1/pomeroy-2002-v1_database.txt**) (Erişim Tarihi: Temmuz 2016).
187. Risinger, J.I., et al., 2003. Advances in Brief Microarray Analysis Reveals Distinct Gene Expression Profiles among Different Histologic Types of Endometrial Cancer. **Cancer, 63** (1): 6 –11.
188. (Risinger-2003) Endometrial Cancer Dataset. 2013. (Web Page: **http://bioinformatics.rutgers.edu/Static/Supplements/CompCancer/CDNA/risinger-2003/risinger-2003_database.txt**) (Erişim Tarihi: Temmuz 2016).
189. Singh, D., 2002. Gene expression correlates of clinical prostate cancer behavior. **Cancer Cell, 1** (2): 203–209.
190. (Singh-2002) Prostate Cancer dataset. 2013. (Web Page: **http://bioinformatics.rutgers.edu/Static/Supplements/CompCancer/Affymetrix/singh-2002/singh-2002_database.txt)** (Erişim Tarihi: Temmuz 2016).
191. Haznedar, B., Kalinli, A., 2016. Training ANFIS using genetic algorithm for dynamic systems identification, **International Journal of Intelligent Systems and Applications in Engineering, 4**: 44-47.

192. Yip, P.P.C., Pao, Y.H., 1995. Combinatorial optimization with use of guided evolutionary simulated annealing, **IEEE Transactions on Neural Networks, 6** (2): 290-294.

193. Reeves, C., 1993. Diversity and diversification in genetic algorithms: Some connections with tabu search, pp. 338-344. *In:* Artificial Neural Nets and Genetic Algorithms - Proceedings of the International Conference in Innsbruck. Springer-Verlag, Austria.

194. Paulli, J., 1993. A computational comparison of simulated annealing and tabu search applied to the quadratic assigment problem, **Annals of Operations Research, 41**: 85-101.

195. Michalewicz, Z., 1992. Genetic Algorithms + Data Structures = Evaluation Programs, pp. 165-191. Springer-Verlag Press, USA.

196. Karaboğa, D., 1995. A modified genetic algorithm for function optimisation, pp. 255-261. *In the Fourth Turkish Symposium on Artificial Intelligence and Neural Networks - TAINN'95*, Türkiye.

197. Kalınlı, A., 1996. Geri Beslemeli Yapay Sinir Ağlarının Genetik Operatörlere Dayalı Tabu Araştırma Algoritması Kullanarak Eğitilmesi, Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, Kayseri, 143s.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı, Soyadı : Bülent HAZNEDAR

Uyruğu : Türkiye (TC)

Doğum Tarihi ve Yeri: 13 Ağustos 1984, Gaziantep

Tel : +90 532 507 18 63

E-mail : bulenthaznedar@gmail.com

Yazışma Adresi : Hasan Kalyoncu Üniversitesi (HKÜ), Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Şahinbey/GAZİANTEP

EĞİTİM

Derece	Kurum	Mez. Yılı
Yüksek Lisans	ERÜ Fen Bil. Ens. Bilgisayar Müh. ABD., Kayseri	2010
Lisans	ERÜ Müh. Fak. Bilgisayar Müh. Bölümü, Kayseri	2007
Lise	Fitnat Nuri Tekerekoğlu Anadolu Lisesi, Gaziantep	2003

İŞ DENEYİMLERİ

Yıl	Kurum	Görev
2013-Halen	Hasan Kalyoncu Üniversitesi, Mühendislik Fakültesi, Bilgisayar Müh. Bölümü, Gaziantep	Öğretim Görevlisi
2011 - 2013	Erciyes Üniversitesi, Kayseri	Uzman
2008 - 2011	Gaziantep Büyükşehir Belediyesi GASKİ Genel Müdürlüğü	Bilgisayar Mühendisi

YABANCI DİL

İngilizce

TEZ İLE İLGİLİ YAYINLAR

1. Haznedar, B., Kalinli, A., *Revizyon gerçekleşti*. Training ANFIS structure using simulated annealing algorithm for dynamic systems identification. **Applied Soft Computing.**
2. Haznedar, B., Arslan, M.T., Kalinli, A., *Revizyon gerçekleşti*. A New Hybrid Approach for Cancer Classification Based on Microarray Gene Expression Data. **Computer Methods and Programs in Biomedicine.**
3. Haznedar, B., Arslan, M.T., Kalinli, A., 2017. Karaciğer Mikroarray Kanser Verisinin Sınıflandırılması için Genetik Algoritma Kullanarak ANFIS'in eğitilmesi, **Sakarya University Journal Of Science, 21** (1): 54-62.
4. Haznedar, B., Kalinli, A., 2016. Training ANFIS using genetic algorithm for dynamic systems identification, **International Journal of Intelligent Systems and Applications in Engineering, 4**: 44-47.
5. Haznedar, B., Kalinli, A., 2016. Training ANFIS using genetic algorithm for dynamic systems identification, pp. 22-26. *International Conference on Advanced Technology & Sciences (ICAT'16)*, 1-3 Eylül, 2016, Konya, Turkey.
6. Haznedar, B., Kalinli, A., 2016. Detection of the relationship between thrombophilia disease with genetic disorders by adaptive neuro-fuzzy inference system (ANFIS), **Sakarya University Journal Of Science, 20** (1): 13-21.