ZAMAN SERİSİ KESTİRİMİ İÇİN PARAMETRİK T-NORM OPERATÖRLÜ YENİ BİR BULANIK ÇIKARIM SİSTEMİ VE TOPLULUK METODU

Yüksek Lisans Tezi

Salih Berkan AYDEMİR

Eskişehir, 2016

ZAMAN SERİSİ KESTİRİMİ İÇİN PARAMETRİK T-NORM OPERATÖRLÜ YENİ BİR BULANIK ÇIKARIM SİSTEMİ VE TOPLULUK METODU

Salih Berkan AYDEMİR

YÜKSEK LİSANS TEZİ

Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Prof. Dr. Yusuf OYSAL

Eskişehir

Anadolu Üniversitesi

Fen Bilimleri Enstitüsü

Haziran, 2016

JÜRİ VE ENSTİTÜ ONAYI

Salih Berkan Aydemir'in **"Zaman Serisi Kestirimi için Parametrik T-norm Operatörlü Yeni bir Bulanık Çıkarım Sistemi ve Topluluk Metodu"** başlıklı tezi 20/06/2016 tarihinde, aşağıdaki jüri tarafından değerlendirilerek "Anadolu Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliği"nin ilgili maddeleri uyarınca, Bilgisayar Mühendisliği Anabilim dalında Yüksek Lisans tezi olarak kabul edilmiştir.

		Adı Soyadı	**İmza**
Üye (Tez Danışmanı)	**:**	**Prof. Dr. Yusuf OYSAL**	**……..**
Üye	**:**	**Doç. Dr. Hüseyin POLAT**	**……..**
Üye	**:**	**Doç. Dr. Ahmet YAZICI**	**……..**

…….……..

Enstitü Müdürü

ÖZET

ZAMAN SERİSİ KESTİRİMİ İÇİN PARAMETRİK T-NORM OPERATÖRLÜ YENİ BİR BULANIK ÇIKARIM SİSTEMİ VE TOPLULUK METODU

Salih Berkan AYDEMİR

Bilgisayar Mühendisliği Anabilim Dalı

Anadolu Üniversitesi, Fen Bilimleri Enstitüsü, Haziran, 2016

Danışman: Prof. Dr. Yusuf OYSAL

2016, 64 sayfa

Tezin iki ana konusu bulunmaktadır. İlki, adaptif dalgacık ağlar ve adaptif sinirsel bulanık çıkarım sistemi üzerinde, Parametrik-Hamacher T-norm operatörünün uygulanmasıdır. Adaptif bulanık çıkarım sistemi yaklaşımlarında, genel olarak çarpım veya minimum T-norm operatörleri kullanılmaktadır. Parametrik-Hamacher T-normun en önemli özelliği ise, parametrik yapısı sebebi ile birçok T-norm operatörünü sağlamasıdır. Sunulan yapılarda, Adaptif bulanık çıkarım sistemi koşul kısmında, Gaussian üyelik fonksiyonu kullanılırken, Adaptif Dalgacık Ağlarda dalgacık fonksiyonu kullanılmaktadır. Her iki modelin, sonuç kısmında, birinci dereceden polinom fonksiyonları yer almaktadır. Sunulan model, kaotik zaman serileri ve sistem tanılama problemi üzerinde denenmiştir. Ayrıca, bilinmeyen parametrelerin en uygun değerlerinin bulunması, yaklaşık Newton yöntemi tabanlı, Broyden-Fletcher-Goldfarb-Shanno algoritması ile gerçekleştirilmiştir. Ayrıca sunulan yapı, literatürdeki diğer çalışmalar ile karşılaştırılmıştır. Sunulan diğer yapı ise önerilen ilk modellerin üzerinde, topluluk yapısının uygulanmasıdır. Topluluk yapısının son aşamasında, ağırlıklı ortalama uygulanmaktadır. En uygun ağırlıkların seçilebilmesi adına, bu kısımda da öğrenme algoritması uygulanmış, daha sonra ağırlıklı ortalama hesaplanmıştır. Bu model de sunulan yapı, finansal zaman serileri üzerinde denenmiş ve Topluluk-Adaptif Dalgacık Ağlar ile Topluluk-Adaptif Bulanık Çıkarım Sistemi birbiriyle karşılaştırılmıştır. Sonuçlara bakıldığında, sunulan yapılar, genel olarak kullanılan çarpım T-normundan daha iyi sonuçlar verdiği görülmektedir.

Anahtar Kelimeler: Adaptif Bulanık Çıkarım Sistemi, Adaptif Dalgacık Ağları, T-operatörler, Zaman Serisi Kestirimi, Topluluk Metodu

ABSTRACT

NEW A FUZZY INFERENCE SYSTEM AND ENSEMBLE METHOD WİTH PARAMETRİC T-NORM OPERATOR FOR TİME SERİES PREDİCTİON

Salih Berkan AYDEMİR

Department Computer Engineering

Anadolu University, Graduate School of Science, June, 2016

Supervisor: Prof. Dr. Yusuf OYSAL

2016, 64 pages

The thesis consists of two main issues. First issue is that Parametric Hamacher T-norm operator is applied to Adaptive Wavelet Networks (AWN) and Adaptive Neuro-Fuzzy Inference System (ANFIS). Product and min. T-norm operators are usually performed in Adaptive Neuro-Fuzzy Inference System. The major feature of parametric Hamacher T-norm is to provide several T-norm operator by means of its parametric structures. In proposed methods while Gaussian membership function is utilized for Adaptive Neuro-Fuzzy Inference System's condition part, wavelet function is used for Adaptive Wavelet Networks for the same manner. First order polynomial functions are available in consequence parts of both these models. Proposed model is tried chaotic time series prediction and system identification problems. Furthermore, optimum values of unknown parameter are determined by Broyden-Fletcher-Goldfarb-Shanno algorithm based on approximate Newton method. Moreover, proposed model is compared with other studies in the literature. The other proposed model is to apply ensemble method on the first proposed model. In the last phase of Ensemble model weighted average technique is applied. In order to determine optimal weights, learning algorithm is implemented and weighted average is calculated. The second proposed model is experimented on financial time series. Ensemble Adaptive Neuro-Fuzzy Inference System and Ensemble Adaptive Wavelet Networks is compared with each other.

Keywords: ANFIS, Adaptive Wavelet Network, T-Operators, Time Series Prediction, Ensemble Method

TEŞEKKÜR

Yüksek lisans tezim boyunca benden yardımlarını esirgemeyen, karşılaştığım sorunlarda bilgi ve deneyimlerini benimle paylaşan değerli hocam ve tez danışmanım Prof. Dr. Yusuf OYSAL'a teşekkürü bir borç bilirim.

Tez süresince desteklerini aldığım, Anadolu Üniversitesi Bilgisayar Mühendisliği Bölümü'nde görev yapmakta olan tüm hocalarıma ve çalışma arkadaşlarıma teşekkürlerimi sunarım.

Çalışma hayatım boyunca her zaman yanımda olan ve desteklerini esirgemeyen, biricik anne ve babama teşekkür ederim.

Salih Berkan AYDEMİR

Haziran, 2016

…/…./2016

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Bu tezin bana ait, özgün bir çalışma olduğunu; çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı; bu çalışma kapsamında elde edilemeyen tüm veri ve bilgiler için kaynak gösterdiğimi ve bu kaynaklara kaynakçada yer verdiğimi; bu çalışmanın Anadolu Üniversitesi tarafından kullanılan "bilimsel intihal tespit programı"yla tarandığını ve hiçbir şekilde "intihal içermediğini" beyan ederim. Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçlara razı olduğumu bildiririm.

…….….

(İmza)

……...

(Adı-Soyadı)

İÇİNDEKİLER

Sayfa

BAŞLIK SAYFASI i

JÜRİ VE ENSTİTÜ ONAYI ii

ÖZET iii

ABSTRACT iv

TEŞEKKÜR v

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ vi

TABLOLAR DİZİNİ ix

ŞEKİLLER DİZİNİ x

SİMGELER VE KISALTMALAR DİZİNİ xi

1. GİRİŞ 1

1.1 T(Üçgensel)-Operatörler 1

1.2 Sinir Ağları 3

1.3 Dalgacık Sinir Ağları 5

1.4 Bulanık Sinir Ağları Sistemi 5

1.5 Topluluk Adaptif Sinirsel-Bulanık Çıkarım Sistemi 6

1.6 Tezin İçeriği 7

2. KURAMSAL İÇERİK 8

2.1 T-Norm ve T-Conorm (Üçgensel Normlar) 8

2.1.1 Parametrik Hamacher T-norm yapısının genelleştirilmesi 15

2.2 Adaptif Sinirsel Bulanık Çıkarım Sistemi 16

2.2.1 Bulanık çıkarım sistemi 16

2.2.2 ANFIS-AWN mimarisi 18

2.3 Tezde Sunulan Modeller 21

3. ANFIS ve AWN YAPILARINDA FARKLI T-NORMLARIN KULLANIMI 23

3.1 ANFIS-Einstein ve ANFIS-PH Yapıları ... 23

3.1.1 ANFIS-Einstein ve ANFIS-PH için eğitim algoritması ... 26

3.1.2 Broyden-Fletcher-Goldfarb-Shanno gradient yöntemi ... 27

3.1.3 ANFIS-PH için parametre gradyanları ... 28

3.2 AWN-Einstein ve AWN-PH Yapıları ... 30

3.2.1 AWN-PH için parametre gradyanları ... 33

4. TOPLULUK ANFIS-PH ve AWN-PH YAPILARI ... 35

5. DENEYSEL ÇALIŞMALAR ... 38

5.1 ANFIS/AWN-PH Deneyleri ... 38

5.1.1 Box-Jenkins zaman serisi tahmini ... 38

5.1.2 Mackey-Glass Zaman Serisi Tahmini ... 41

5.1.3 Sistem tanımlama örneği ... 44

5.2 Ens-ANFIS-PH ve Ens-AWN-PH Deneyleri ... 48

5.2.1 Dow-Jones zaman serisi tahmini ... 49

5.2.2 S&P 500 zaman serisi tahmini ... 52

5.2.3 Euro/Dolar zaman serisi tahmini ... 54

6. SONUÇ ... 56

KAYNAKÇA ... 58

ÖZGEÇMİŞ

TABLOLAR DİZİNİ

Sayfa

Tablo 2.1 Bazı T-norm ve T-conorm yapıları 10

Tablo 5.1 Box-Jenkins zaman serisinin λ değerleri için ANFIS-PH sonuçları..... 39

Tablo 5.2 Box-Jenkins zaman serisinin λ değerleri için AWN-PH sonuçları..... 39

Tablo 5.3 Box-Jenkins zaman serisi için modellerin test sonuçlarının karşılaştırılması 41

Tablo 5.4 Mackey-Glass zaman serisinin λ değerleri için ANFIS-PH sonuçları..... 42

Tablo 5.5 Mackey-Glass zaman serisinin λ değerleri için AWN-PH sonuçları..... 42

Tablo 5.6 Mackey Glass zaman serisi için modellerin test sonuçlarının karşılaştırılması..... 44

Tablo 5.7 Sistem tanılama probleminin λ değerleri için ANFIS-PH sonuçları 46

Tablo 5.8 Sistem tanılama probleminin λ değerleri için AWN-PH sonuçları 46

Tablo 5.9 Sistem tanıma problemi için modellerin test sonuçlarının karşılaştırılması .. 48

Tablo 5.10 Dow-Jones zaman serisinin λ değerleri için Ens-ANFIS-PH sonuçları 51

Tablo 5.11 Dow-Jones zaman serisinin λ değerleri için Ens-AWN-PH sonuçları 51

Tablo 5.12 S&P 500 zaman serisinin λ değerleri için Ens-ANFIS-PH sonuçları..... 52

Tablo 5.13 S&P 500 zaman serisinin λ değerleri için Ens-AWN-PH sonuçları..... 53

Tablo 5.14 Euro/Dolar zaman serisinin λ değerleri için Ens-ANFIS-PH sonuçları 54

Tablo 5.15 Euro/Dolar zaman serisinin λ değerleri için Ens-AWN-PH sonuçları 55

Tablo 6.1 ANFIS–PH ve AWN–PH modellerinin karşılaştırılması (HKOK Değerleri)56

Tablo 6.2 Ens-ANFIS–PH ve Ens-AWN–PH modellerinin karşılaştırılması (OMYH Değerleri) 56

ŞEKİLLER DİZİNİ

Sayfa

Şekil 1.1 Biyolojik Nöron 4
Şekil 1.2 Yapay Nöronun temel elemanları 4
Şekil 2.1 Gaussian ve G-Bell üyelik fonksiyonları 11
Şekil 2.2 Min. ve Çarpım T-normlarının gösterimi 12
Şekil 2.3 Sınırlandırılmış ve Güçlü(drastic) çarpım T-normlarının gösterimi 12
Şekil 2.4 Hamacher ve Einstein çarpım T-normlarının gösterimi 13
Şekil 2.5 Max. ve Çarpım T-conorm yapılarının gösterimi 13
Şekil 2.6 Sınırlandırılmış ve Güçlü(drastic) T-conorm yapılarının gösterimi 14
Şekil 2.7 Hamacher ve Einstein T-conorm yapılarının gösterimi 14
Şekil 2.8 Bulanık Çıkarım Sistemi 17
Şekil 2.9 Sugeno tip bulanık çıkarım sistemi 18
Şekil 2.10 ANFIS-AWN yapısı 19
Şekil 3.1 ANFIS-Parametrik Hamacher yapısı (mavi çizgili kısım) 23
Şekil 3.2 AWN-Parametrik Hamacher yapısı (mavi çizgili kısım Parametrik-Hamacher, kırmızı çizgili kısım dalgacık fonksiyonu) 31
Şekil 4.1 Topluluk AWN-PH yapısı 36
Şekil 5.1 Box-Jenkins zaman serisi için ANFIS-PH (λ=0,035) tahmin hatası 40
Şekil 5.2 Box-Jenkins zaman serisi için ANFIS-PH (λ=0,035) tahmini ve gerçek değerleri 40
Şekil 5.3 Mackey-Glass zaman serisi için AWN-PH (λ=2) tahmin hatası 43
Şekil 5.4 Mackey-Glass serisi için AWN-PH (λ=2) tahmini ve gerçek değerleri 43
Şekil 5.5 ANFIS-PH veya AWN-PH ile Sistem Tanılama Modeli 45
Şekil 5.6 Sistem tanılama için AWN-PH (λ=1,0806) tahmin hatası 47
Şekil 5.7 Sistem tanılama için AWN-PH (λ=1,0806) tahmini ve gerçek değerleri 47
Şekil 5.8 Ens-ANFIS ve Ens-AWN-PH yapısı 50
Şekil 5.9 Dow-Jones zaman serisinin Ens-AWN-PH yapısı için tahmin hatası ve tahmini 51
Şekil 5.10 S&P 500 zaman serisinin Ens-ANFIS yapısı için tahmin hatası ve tahmini. 53
Şekil 5.11 Euro/Dolar zaman serisinin Ens-AWN-PH yapısı için tahmin hatası ve tahmini 55

SİMGELER VE KISALTMALAR DİZİNİ

AWN: Adaptive Wavelet Network

ANFIS: Adaptive Neuro-Fuzzy Inference System

WNN: Wavelet Neural Network

PH: Parametric Hamacher

RMSE: Root Mean Square Error

BFGS: Broyden-Fletcher-Goldfarb-Shanno

S&P: Standard and Poors

MAPE: Mean Absolute Percentage Error

LLWN: Local Linear Wavelet Neural Network

Ens: Ensemble

FWNN: Fuzzy Wavelet Neural Network

1. GİRİŞ

Bulanık sinir ağı mantığı, yapay sinir ağlarının öğrenme yetisi ile bulanık mantığın insan gibi karar verebilme ve uzman bilgisi sağlama faydalarının bütününe dayanmaktadır. Bulanık-Sinir ağı sistemleri üzerinde sistem tanıma, fonksiyon yaklaşımı, zaman serisi kestirimleri gibi birçok alanda uygulanmıştır [1].

Günümüze kadar geliştirilen Bulanık-Sinir ağları sistemlerinin amacı, hem çeşitli zaman serileri hem de çeşitli sistem tanıma problemleri üzerinde, sistemin doğruluğunu arttırmak, yani çeşitli performans ölçütlerinin değerini azaltmaktır. Bu ölçütün sağlanması için de birçok çalışmada, bu sistemin katmanlarında, farklı sürekli fonksiyonlar kullanılmış, hatta çeşitli öğrenme algoritmaları ile hibrit edilmesi sonucunda da birçok Bulanık-Sinir ağı modelleri sunulmuştur [1-6]. Ancak Bulanık-Sinir ağları sistemi üzerinde, üyelik fonksiyonlarından elde edilen değerler yaygın olarak standart çarpım işlemi ile birbirleri ile çarpılmakta veya minimum operatörü kullanılmaktadır. Sistemin bu kısmında farklı T-norm operatörleri kullanılarak sistemde iyileştirme yapmak mümkündür. Tezin iki ana konusu bulunmaktadır. İlki zaman serisi kestirimi için bir bulanık sinir ağı modeli üzerinde genel olarak kullanılan minimum veya çarpım T-normu yerine, parametrik Hamacher (PH) T-norm operatörü ile Einstein T-norm operatörü esas alınmaktadır. Einstein T-norm, PH T-normun özel bir halidir. Sunulan bir diğer yapı ise topluluk (ensemble) Adaptif Sinirsel Bulanık Çıkarım Sistemi (ANFIS)'i üzerinde öğrenme algoritmasının uygulanması ve finansal zaman serileri üzerinde sınanmasıdır.

Tezin giriş kısmında bahsedeceğimiz diğer bölümler şu şekilde sıralanabilir; Bölüm 1.1'de T-normların farklı sistemler üzerindeki etkilerine ve bu konu ile ilgili yapılmış diğer çalışmalara, Bölüm 1.2'de çeşitli Sinir Ağları yapılarına ve bu yapılar hakkındaki çalışmalara, Bölüm 1.3'de Dalgacık-Sinir ağları yapılarına Bölüm 1.4'de Bulanık Sinir Ağları sistemi hakkında yapılan çalışmalara ve Bölüm 1.5'de ise topluluk ANFIS'i üzerine yapılan çalışmalara yer verilecektir. Son olarak Bölüm 1.6'da ise tezin içeriğinden bahsedilecektir.

1.1 T(Üçgensel)-Operatörler

T-operatörler, olasılık metrik uzayları ve çoklu değer mantığın yapısında bulunan

operatörlerdir. Bu çalışmada T-operatörler bulanık kümeler üzerinde değerlendirilecektir. T-operatörler; T-norm, T-conorm ve olumsuzluk (negation) fonksiyonları olmak üzere üçe ayrılmaktadır. Bulanık kümelerdeki karşılıklarından bahsedilmesi gerekirse; T-normlar bulanık kümelerin kesişiminin, üyelik değerlerini hesaplamak için kullanılırken, T-conorm ve olumsuzluk operatörleri sırasıyla; bulanık kümelerin birleşimi ve tümleyen üyelik değerlerini hesaplamak için kullanılır [7]. T-operatörler hakkındaki diğer bilgiler ve tanımlar tezin ikinci bölümünde ayrıntılı şekilde yer verilecektir.

T-operatörler üzerine yapılan çalışmalarda bazı T-norm operatörlerinin matematiksel özellikleri sunulmuştur. Bulanık mantık karar verme mekanizmasında Zadeh'in minimum ve maksimum operatörünün dışında birçok T-norm operatörüne yer verilmektedir [7]. T-norm yapısı, şimdiye kadar birçok sistemde kullanılmıştır. Bu anlamda daha önceden üzerinde çalışmalar yapılmış olan, bakteriyel tip algoritmalarının üzerinde, farklı T-norm operatörleri kullanılmış ve Mamdani çıkarım sistemi ile test edilip, sonuçlar sunulmuştur [8]. Ayrıca bu çalışmada Standart, cebirsel (Algebraic), Hamacher ve Trigonometrik T-normlar karşılaştırılmış ve her bir bakteriyel tip algoritmasının içinde de öğrenme algoritması olarak, gradient tabanlı Levenberg-Marquardt metodu kullanılmıştır. Bulanık kural tabanlı sınıflandırmanın üzerine yapılan bir diğer çalışmada [9] ise dokuz farklı T-norm operatörünün, bu sınıflandırma üzerindeki etkisi incelenmiştir.

Ciaramella ve ark. [10] yaptığı çalışmada bulanık-ilişkisel model sunulmuş, bu modelde bulanık sinir ağları sisteminin IF-THEN kısmında minimum ve çarpım T-normları dışında, sınırlandırılmış çarpım ve Yager T-normunun kullanıldığı görülmüştür. Sonuçlar ise çok katmanlı algılayıcı (perceptron) ve radyal tabanlı fonksiyonlar ile karşılaştırılmıştır. Zhang ve Liao [11] 2014 yılında çok girdili Hamacher T-norm' un ispatları ile genelleştirmesini yapmış ve ANFIS yapısını Benchmark problemlerinden biri olan Box-Jenkins veri seti üzerine uygulamıştır. Ayrıca Zhang ve Liao [12] 2015 yılında Pingan (Çin) bankası üzerinde bu sistem ile stok fiyat tahmini yapmış ve farklı T-normlar ile çalışma açısından umut vadeden sonuçlar sunulmuştur.

Gál ve ark. [13] 2014 yılında yaptığı bir çalışmada bulanık kural tabanlı sistemler üzerinde trigonometrik ve Hamacher T-normlarını kullanmışlardır. Bu çalışmada Hamacher T-normun farklı parametre değerleri için karşılaştırılması yapılmıştır. Çalışmada iki tip bakteriyel memetik algoritma kullanılmıştır. Bunlar gelişmiş bakteriyel

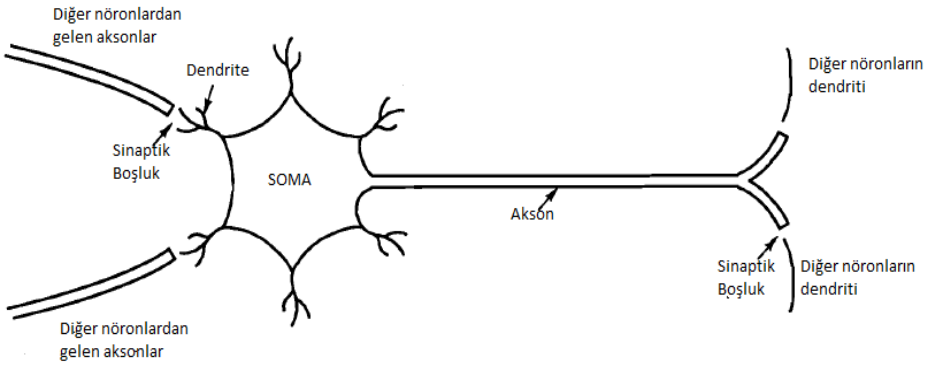
memetik algoritma ve modifiye bakteriyel memetik algoritmalarıdır. Bütün parametre değerleri Levenberg-Marquardt metot ile güncellenmiştir. Ancak diğer parametreler gibi, Hamacher T-norm'un parametresi de, kullanılan veri seti, öğrenme algoritması ve başlangıç parametrelerine göre değişiklik göstermektedir.

Einstein, Hamacher, çarpım gibi T-norm operatörleri metin dosyalarını sınıflandırmak için kullanılan, bulanık benzerlik yöntemlerinde kullanılmış ve çevrimiçi haber derlemeleri üzerinde yapılan sınıflandırmada en iyi performansı Einstein T-normun verdiği görülmüştür [14]. Bahsedildiği gibi Hamacher T-norm ile ilgili birçok çalışma yapılmış, hatta ANFIS yapısında da kullanılmıştır. Ancak hangi parametre değerinin Hamacher T-norm için daha uygun olduğu sorusunun cevabı hala daha açık değildir. Diğer yandan, yapılan çalışmalar incelendiğinde, Einstein T-normun ANFIS gibi uzman sistemlerin hiçbirinde kullanılmadığı, sınıflandırma gibi yapılarda kullanıldığı görülmektedir. Ayrıca benzerlik hesaplama tabanlı işlemlerde ise diğer bazı T-norm yapılarından daha iyi sonuç verdiği görülmüştür. Bunun üzerine bu çalışmada PH T-norm ve Einstein T-normları ile uzman sistemler üzerinde çalışılmıştır.

1.2 Sinir Ağları

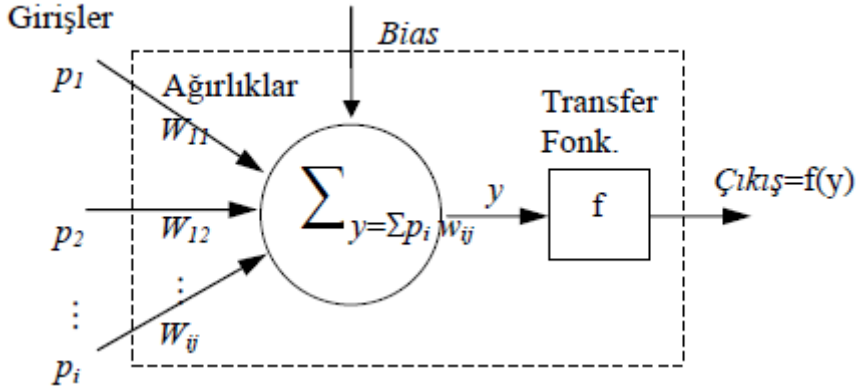
Yapay sinir ağları sisteminin oluşumunda esasen biyolojik sinir ağları sisteminin yapısından esinlenilmektedir. Yapay sinir ağları, sinir biyolojisinin matematiksel modelleri genel bir hale getirilerek, bazı temel hatlar üzerinden sistemler oluşturulmuştur. Bilginin işlemesi, sinir hücreleri (neuron) olarak adlandırılan hücre elemanları üzerinde gerçekleşir. Sinyaller, hücreler arası bağlantılar üzerinden, bir hücreden diğerine geçer. Genel bir sinir ağında, her bir bağlantı bir ağırlığa sahiptir ve bu ağırlık ise bağlantı üzerinden geçen sinyalin bir çarpanıdır. Her hücre kendi çıkışını saptamak için ağ (network) girişlerine, çoğunlukla doğrusal olmayan bir aktivasyon fonksiyonu uygular [15].

Sinir hücresi ile yapay hücre arasında çok benzer bir yapı bulunmaktadır. Biyolojik hücreler üç temel bileşenden oluşmaktadır. Bunlar dendritler, soma (çekirdek), ve akson. Bir biyolojik sinir ağının yapısı şekil 1.1'de gösterilmektedir.



Şekil 1.1 *Biyolojik Nöron*

Yapay sinir ağları genel olarak biyolojik bir sistemin yapısını taklit ederek insan beyninin modellenmesi için tasarlanmış sistemlerdir. Biyolojik nöronlarda; dentrit, hücre gövdesi, aksonlar ve sinapslar, sırasıyla işlem yapan birim, toplama fonksiyonu, transfer fonksiyonu, yapay nöron çıkışı ve ağırlıklara denk gelmektir. Yapay sinir ağlarının en küçük bilgi işleme elemanı olan yapay nöron, Şekil 1.2'de verilmektedir [16].



Şekil 1.2 *Yapay Nöronun temel elemanları*

Bazı sinir ağları türlerinden bahsedilecek olursa; ileri beslemeli (feed-forward), özyinelemeli (recurrent), rekabetçi yapay sinir ağları gibi birçok sinir ağı modeli bulunmaktadır. İleri beslemeli yapay sinir ağları, katmanlardan oluşmaktadır. Bir katmandan elde edilen çıkış, bir sonraki katman için giriş olarak verilir. Giriş katmanına verilen gerçek veriler, direk olarak gizli katmana yönlendirilir ve girdi olarak verilen bilgi, gizli ve çıkış katmanlarında işlenerek sistemin çıktısı belirlenir. Bütün işlemler, girdi kısmından çıktı kısmına doğru ileriye dönük gerçekleşmektedir.

Özyinelemeli ağlar ise aynı zamanda geri beslemeli ağlardır. Sistemdeki bazı çıktı elemanları, girdiler ile ilişkili elemanlardır [17]. Sistemde girdi ve çıktılar arasında en az

1 döngü bulunmaktadır [18]. İleri beslemeli sinir ağına, radyal tabanlı sinir ağları örnek olarak verilebilir. radyal tabanlı sinir ağları, 3 katmandan oluşmaktadır. Girdi katmanı, radyal tabanlı fonksiyona ait bir gizli katman ve çıkış katmanından oluşmaktadır [17]. Rekabetçi yapay sinir ağlarında ise sadece rekabeti kazanan düğümün çıkışı aktive edilir ve sadece rekabeti kazanan düğümün parametreleri güncellenir. Bu çalışmada ileri beslemeli sinir ağlarından dalgacık sinir ağları ve bulanık sinir ağları seçilmiştir.

1.3 Dalgacık Sinir Ağları

Dalgacık teorisi nümerik analiz ve sinyal işleme gibi birçok alanda matematiğin bir dalı olarak bilinmekte hızlı bir şekilde gelişme göstermektedir [19, 20]. Dalgacık sinir ağları (Wavelet Neural Network) modeli ise sinir ağları ile dalgacık fonksiyonunun birleşmesi sonucu ortaya çıkmıştır [21-23]. Sinir ağları öğrenme ve genelleme becerisine sahiptir. Ancak yaklaşım problemleri (approximation problems) için fazla sayıda nörona ihtiyaç duyulmaktadır. Fakat WNN modelinin yakınsaması hızlıdır ve eğitimi kolaydır. WNN modelinin temel noktası ise normalde Gauss ve Sigmoid gibi fonksiyonlar yerine aktivasyon fonksiyonu olarak dalgacık fonksiyonu kullanılmıştır. İki tip WNN yapısı bulunmaktadır. Bunlardan birisi sürekli dalgacık dönüşümlerinden oluşan, genişleme (dilated) ve öteleme (translated) parametreleri ile çıktı katmanındaki ağırlıkları düzenlenebilen yapıdan oluşmaktadır. Diğer yapı ise kesikli dalgacık dönüşümünden oluşan, ancak genişleme (dilated) ve öteleme (translated) parametreleri sabit olan, sadece çıktı katmanındaki ağırlıkları düzenlenen yapılardır [21].

1.4 Bulanık Sinir Ağları Sistemi

Bulanık sinir ağı, yapay sinir ağları ve bulanık sistemlerin birleşmesinden oluşan bir sistemdir. Bulanık mantığın, belirsiz veriler üzerindeki işleme yeteneği ile yapay sinir ağlarının öğrenme yeteneği birleştirilmiştir. Yapay sinir ağı sistemi ise bulanık sistemin parametrelerini belirlemek için kullanılır. Öğrenme tamamlandıktan sonra ise bulanık sistem yapay sinir ağına ihtiyaç duymadan çalışmasına devam etmektedir. Yapı eğitim aşamasında iken, yapay sinir ağlarını, karar verme sürecinde ise bulanık mantığı kullanmaktadır. Literatürde bulanık sinir ağları ile yapılan birçok çalışma mevcuttur [24-37].

Bulanık sinir ağları sistemleri içerisinde en yaygın olanı ve bu tez çalışmasının temelini teşkil eden yapı, ANFIS'tir [1]. ANFIS işlevsel olarak Sugeno çıkarım sistemine eştir. Sugeno bulanık modelinde girdi uzayı bulanık bölgelere ayrılır ve bu bölgeler bulanık kuralların IF kısmındaki (antecedent part) üyelik fonksiyonlarını temsil eder. Kuralların THEN kısmında (consequent part) ise sabit sayılar veya girdilere bağlı lineer fonksiyonlar yer almaktadır [1].

1.5 Topluluk Adaptif Sinirsel-Bulanık Çıkarım Sistemi

Topluluk ANFIS'i farklı özellikteki ya da farklı girdilere sahip, birden fazla ANFIS yapısının, topluluk yapısı için girdi teşkil etmesidir. Her bir ANFIS yapısı sisteme girdi olarak verilir ve çıktılar alınır, ardından gerçek değerler ile tahmin değerler arasındaki hata hesaplanır. Burada "farklı özellikte" anlamından kasıt, girdi olarak alınan ANFIS yapılarının çeşitli katmanlarında farklı üyelik fonksiyonlarının veya farklı sürekli fonksiyonların kullanılmasıdır (AWN'de yer alan dalgacık fonksiyonu gibi). Bu çalışmada her bir ANFIS yapısına farklı girdiler verilmiştir.

Topluluk ANFIS'i kullanılarak son yıllarda kaotik zaman serileri ve finansal zaman serileri üzerine birçok çalışma yapılmıştır. Bu çalışmaları ana hatları ile özetlenmesi gerekirse;

Topluluk ANFIS'i üzerine yapılan ilk çalışmalar arasında 2002 yılında Lima ve ark. [38] yaptıkları çalışma gösterilebilir. Ens-ANFIS ile bulanık sistem yapısı adlı makalede temel topluluk metodu ile genelleştirilmiş topluluk metodu tanımlanmıştır. Çalışmada ilk olarak ANFIS, 8 farklı üyelik fonksiyonu üzerinde uygulanmış ve çıktıları alınmıştır. Ağırlıklandırılmış ortalama hesaplanırken ise amacın ortalama kareler hata (OKH) yönteminin en az indirgeme olması da göz önünde bulundurularak uygun ağırlıkları bulmak için Hashem ve Schmeiser'in [39] sunduğu MSE-OLC (ortalama kareler hata-en uygun lineer birleşimler) yöntemi kullanılmıştır. Daha sonrada Ens-ANFIS'in çıktısı elde edilmiştir.

Chen ve Zhang [40] tarafından yapılan çalışmada eğitim verisi rastgele ve bootstrap metot olmak üzere iki şekilde oluşturulur. Tekdüze (uniform) ve tekdüze olmayan (non-uniform) ağırlıklandırma kullanılarak da her eğitim verisi için oluşturulan ANFIS yapılarını sonucunda da dört Ens-ANFIS (Topluluk ANFIS) oluşur. Çalışmada geçen tekdüze ağırlıklandırma her bir ANFIS'in çıktısının ANFIS sayısına oranı iken tekdüze olmayan ağırlıklandırma hata kareler ortalaması karekökü (HKOK) ile

ilişkilendirmiş bir ağırlıklandırılmış ortalama yöntemi olarak kullanılmıştır. Sistem kaotik zaman serileri ile Pekin'in kentsel kısmındaki trafik lambalarının oluşturduğu zaman serileri üzerinde kullanılmıştır.

Soto ve ark. [41] yaptığı çalışmada 3 farklı üyelik fonksiyonu çeşiti ile 3 ANFIS çıktısı elde edilmiş ve daha sonra test verileri üzerinde bu ANFIS çıktıları üzerinden ortalama ve ağırlıklandırılmış ortalama hesap edilerek elde edilen Ens-ANFIS'i üzerinden RMSE hesaplanmıştır. Ağırlıklandırılmış ortalama hesap edilirken ağırlıklar hatanın daha küçük bulunması açısından denenmiş, 0,5, 0,3 ve 0,2 değerleri seçilmiştir. Bu sistem ise Mackey-Glass zaman serisi üzerinde uygulanmıştır.

2012 yılında Lei ve Wan [42] tarafından yapılan çalışmada Çin'in Makao şehrindeki hava kirlilik indeksi tahmini için yukarıda bahsedilen Chen ve Zhan'ın Ens-ANFIS'i yöntemini kullanmışlardır. Suhartono ve ark. [43] yaptığı bir çalışmada ise ANFIS-ARIMA sistemleri üzerinde Endonezya'nın hava yağışı tahmini yapmışlardır. Ens-ANFIS'nin çıktı kısmındaki katsayıların daha uygun seçebilmesi adına, Breiman'ın [44] istifleme yöntemi kullanılmıştır.

Gabralla ve ark. [45] ham petrol fiyatı tahmini için yaptığı çalışmada, Ens-ANFIS çıktısı oluşturulurken ağırlıklı ortalamalar kısmındaki ağırlık katsayıları, parçacık koloni optimizasyon yöntemi ile güncellenmiştir.

1.6 Tezin İçeriği

Tezin 2. Bölümünde, ANFIS'in yapısına ve yararlanılan öğrenme algoritmasına, Adaptif Dalgacık Ağları'nın yapısına ve T-normlar hakkında ayrıntılı bilgilere yer verilecektir, daha sonra ise tezde sunulan yapı ile ilgili bilgilere yer verilecektir. Tezin 3. ve 4. bölümlerinde sunulan modellere yer verilecektir. 5. bölümde, bu modellerin biri kaotik zaman serileri ve sistem tanıma problemi için test edilirken, diğeri ise finansal zaman serileri üzerinde test edilecektir. Deneylerin sonuçlarına ve diğer çalışmalarla olan karşılaştırmalarına yer verilecektir. 6. bölümde sonuç kısmı olup, sunulan modelin, geliştirilmesi adına kullanılabilecek bazı yöntemlerden bahsedilecektir.

2. KURAMSAL İÇERİK

Bu bölümde, ilk olarak, Bölüm 2.1'de, T-normlar hakkında geniş bilgilere yer verilecektir. Bölüm 2.2 ve 2.3'de sırasıyla, ANFIS'in yapısına ve öğrenme algoritmasına yer verilip, Bölüm 2.4'de ise tezde amaçlanan kısım ve sunulan modeller özetlenecektir.

2.1 T-Norm ve T-Conorm (Üçgensel Normlar)

T-norm ve T-conorm'lar 1942 yılında Merger'in [46] yapmış olduğu "İstatiksel Metrikler" çalışması ile temelleri atılmıştır. T-normlar istatistiksel metrik uzaylarda üçgen eşitsizliğinin genelleştirilmesi için yapılan çalışmalarda ortaya çıkmaktadır. Ardından Schweizer ve Sklar [47] ise " Olasılıksal Metrik uzaylar" adlı kitap ile bu alana katkıda bulunmuşlardır. Çok değerli mantık (Many-valued logic) ile bulanık mantık (fuzzy logic) arasındaki bağlantı T-norm yapısı ile ifade edilmektedir. T-normun tanımına geçmeden önce, gerekli bazı tanımları vermemiz daha uygun olur.

2.1 ***Tanım: Bulanık Küme***

A boş olmayan bir küme olsun. U:A→[0,1] fonksiyonuna A üzerinde bir bulanık küme denir. Ayrıca a ϵ A olacak şekildeki U(a) değerine a'nın U'daki üyelik derecesi (Membership degree) denir. U'ya ise üyelik fonksiyonu (Membership function) denir ve $U = \{(a, \mu_A(a))a\epsilon A\}$ şeklinde gösterilir. Burada 0$\leq \mu_A(a) \leq 1$ 'dir [48].

2.2 ***Tanım: T- Operatörler***

T-norm: Her $x, y, z \epsilon$ [0,1] için aşağıdaki koşulları sağlayan T:[0,1] × [0,1] → [0,1] fonksiyonuna, üçgensel norm ya da kısaca T-norm denir [7].

(T-1) $T(x, y) = T(y, x)$ (simetri özelliği)

(T-2) $T(x, T(y, z)) = T(T(x, y), z)$ (birleşmelilik)

(T-3) $y \leq$ z iken $T(x, y) \leq$ T(x, z) (monotonluk)

(T-4) $T(x, 1) = x$ ve $T(0,0) = 0$ (sınır koşulu)

Bahsedilen T fonksiyonu, bulanık kümelerde kesişim işlemine karşılık gelmektedir. T-norm olmanın ilk şartı, T-norm işlem sonucunun sıralamayla değişmediğini göstermektedir. İkinci şartı, birleşmeli olduğu için, ikiden fazla bulanık kümenin üyelik derecesinin hesaplanabileceğini göstermektedir. Tanımın üçüncü şartı, y

ve z üzerinde üyelik derecesinin azalmasının işlem sonucu oluşacak kümede üyelik fonksiyonu değerinde bir artışa neden olmayacağından bahsetmektedir. Son şartı ise T-norm işleminin klasik kümelere genelleştirmesi sağlanmaktadır.

T-conorm: Her $x, y, z \epsilon$ [0,1] için aşağıdaki koşulları sağlayan T^*:[0,1] × [0,1] → [0,1] fonksiyonuna, üçgensel conorm ya da kısaca T-conorm denir [50].

(T^*-1) $T^*(x, y) = T^*(y, x)$ (simetri özelliği)

(T^*-2) $T^*\big(x, T^*(y, z)\big) = T^*(T^*(x, y), z)$ (birleşmelilik)

(T^*-3) $y \leq$ z iken $T^*(x, y) \leq T^*(x, z)$ (monotonluk)

(T^*-4) $T^*(x, 0) = x$ ve $T^*(1,1) = 1$ (sınır koşulu)

T-conorm yapıları bulanık kümelerde birleşim işlemine karşılık gelmektedir. Ayrıca T-norm ve T-conorm yapıları arasında bir dualite ilişkisi bulunmaktadır. Bu ilişki, $T(x, y) = 1 - T^*(1 - x, 1 - y)$ ile ifade edilebilir. Bu dönüşüm sayesinde T-norm yapısından T-conorm elde edilebilmektedir [7].

T-normlar üzerine yapılan ilk çalışmaların [47] ardından birçok T-norm yapısı sunulmuş, çeşitli alanlarda ise uygulamaları yapılmıştır [7-14]. Tablo 2.1'de verilen T-norm yapıları, en çok bilinen T-norm yapılarıdır. T-norm yapıları sürekli ve kesikli olmak üzere ikiye ayrılmaktadır. Tezde sunulan yapıların öğrenme aşamasında, türev-tabanlı öğrenme algoritması kullanılacağı için, bu kısımda bazı sürekli t-norm yapılarına yer verilmektedir [7].

Tablo 2.1'de bazı T-norm yapıları benzerlik hesaplama, sınıflandırma, tahmin üretme gibi temel alanlarda kullanılmaktadır. T-norm yapılarından birkaçının, diğer bazı T-norm yapılarını kapsadıkları görülebilir. T-norm yapısının anlaşılabilir olması açısından Tablo 2.1'de bahsedilen bazı T-normların grafikleri verilmiştir. İlk olarak Şekil 2.1'de hiçbir T-norm ve T-conorm işlemine girmeden, G-bell ve Gaussian üyelik fonksiyonları verilmiştir.

Tablo 2.1. *Bazı T-norm ve T-conorm yapıları*

Sıra No	$T(x,y)$	$T^*(x,y)$	KAPSAM
1	$min(x,y)$	$max(x,y)$	Zadeh [49]
2	$x.y$	$x+y-xy$	Goguen [50], Bandler ve ark. [51]
3	$max(x+y-1,0)$	$min(x+y,1)$	Giles ve ark. [52]
4	$\begin{cases} y=1\ ise\ x \\ x=1\ ise\ y \\ 0 \quad d.d \end{cases}$	$\begin{cases} y=0\ ise\ x \\ x=0\ ise\ y \\ 1 \quad d.d \end{cases}$	Weber ve ark. [50]
5	$\frac{xy}{x+y-xy}$	$\frac{x+y-2xy}{1-xy}$	Temel Hamacher [50]
6	$\frac{xy}{1+(1-x)(1-y)}$	$\frac{x+y}{1+xy}$	Einstein [50]
7	$max(1-((1-x)^p+(1-y)^p)^{\frac{1}{p}},0)$	$\min((x^p+y^p)^{\frac{1}{p}},0)$	Yager [53] p=1 $\rightarrow$ $T_3\ ve\ T_3^*$
8	$\frac{1}{1+((\frac{1}{x}-1)^{\lambda}+(\frac{1}{y}-1)^{\lambda})^{\frac{1}{\lambda}}}$	$\frac{1}{1+((\frac{1}{x}-1)^{-\lambda}+(\frac{1}{y}-1)^{-\lambda})^{-\frac{1}{\lambda}}}$	Dombi [54] $\lambda=1\rightarrow\ T_5\ ve\ T_5^*$
9	$\frac{xy}{\lambda+(1-\lambda)(x+y-xy)}$	$\frac{\lambda(x+y)+xy(1-2\lambda)}{\lambda+xy(1-\lambda)}$	Hamacher [50] $\lambda=0\rightarrow\ T_5\ ve\ T_5^*$ $\lambda=1\rightarrow\ T_2\ ve\ T_2^*$ $\lambda=2\rightarrow\ T_6\ ve\ T_6^*$
10	$e^{-((-lnx)^a+(-lny)^a)^{\frac{1}{a}}}$	$1-e^{-((-\ln(1-x))^a+(-\ln(1-y))^a)^{\frac{1}{a}}}$	Aczel-Alsina [55]
11	$\frac{2}{\pi}\arcsin(\sin(a\frac{\pi}{2})\sin(b\frac{\pi}{2}))$	$\frac{2}{\pi}\arccos(\cos(a\frac{\pi}{2})\cos(b\frac{\pi}{2}))$	Gal ve ark. [56]
12	$\max((1+\lambda)(x+y-1)-\lambda xy,0)$	$\min(x++y+\lambda xy,1)$	Yu Yandong [57]

T-normların yapısı görselleştirilirken aralarındaki farkların görülebilmesi açısından Gaussian ve G-Bell gibi iki faklı üyelik fonksiyonu seçilmiştir. Bu üyelik fonksiyonları ise parametreleri ile birlikte verilmektedir. İlk olarak Şekil 2.1'de hiçbir T-norm ve T-conorm işlemine girmeden, G-bell ve Gaussian üyelik fonksiyonları verilmiştir.

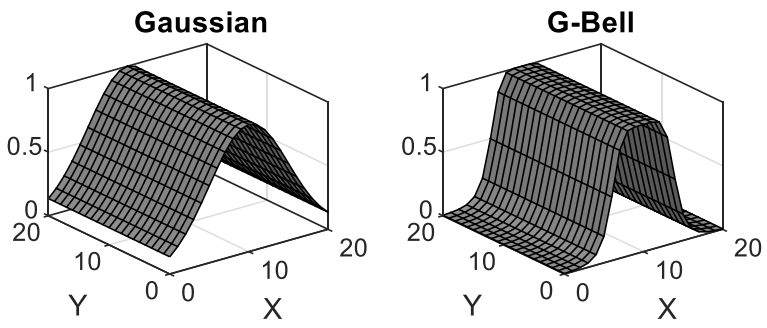
Gaussian$(x; c, \sigma)$ = $e^{-\frac{1}{2}(\frac{x-c}{\sigma})^2}$ (c: merkez, σ: genişlik)

Bell$(x; a, b, c)$ = $\frac{1}{1+\left|\frac{x-c}{a}\right|^{2b}}$ (a: genişlik, b: tepe genişliği, c: merkez)

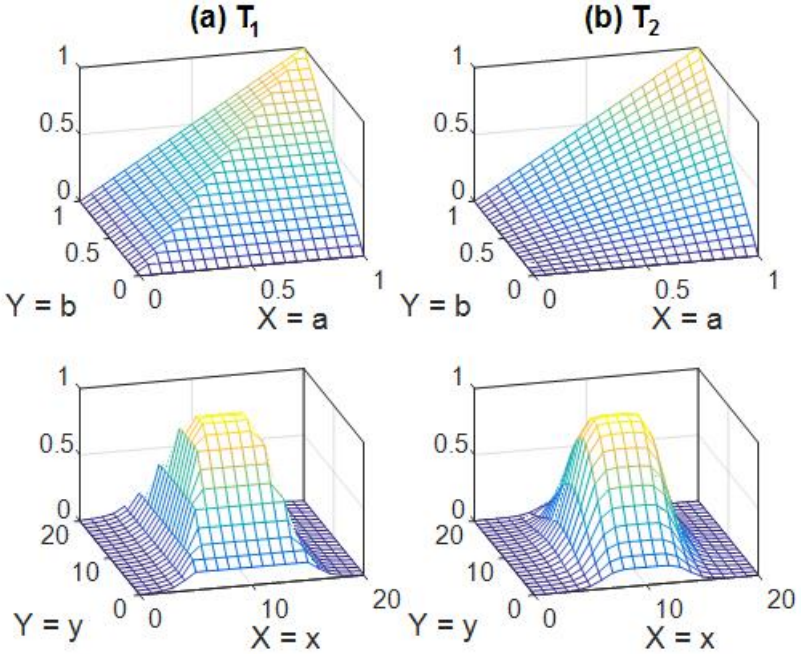
A=gbell_mf(x, [4, 3, 10])

B=gauss_mf(x, [10, 5])

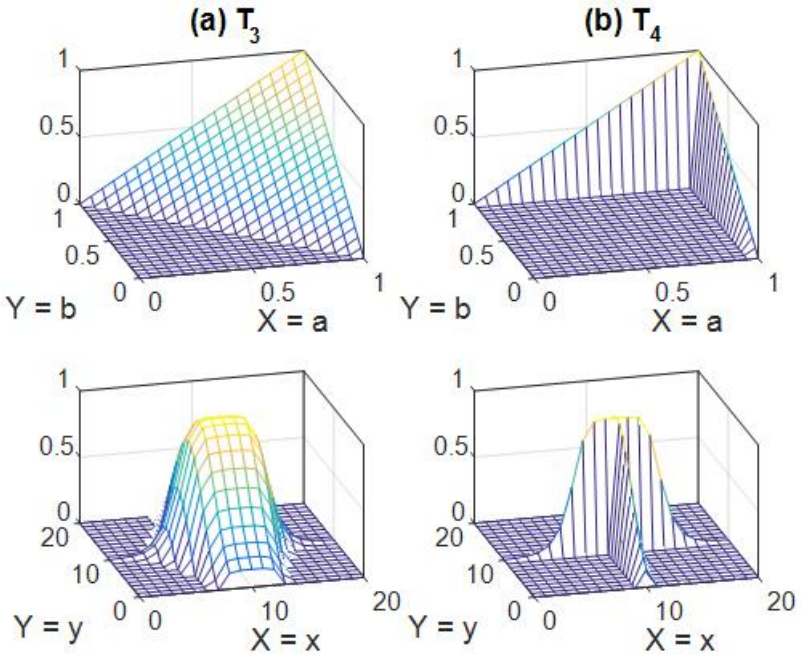
Grafikler için A ve B üyelik fonksiyonları kullanılmıştır.



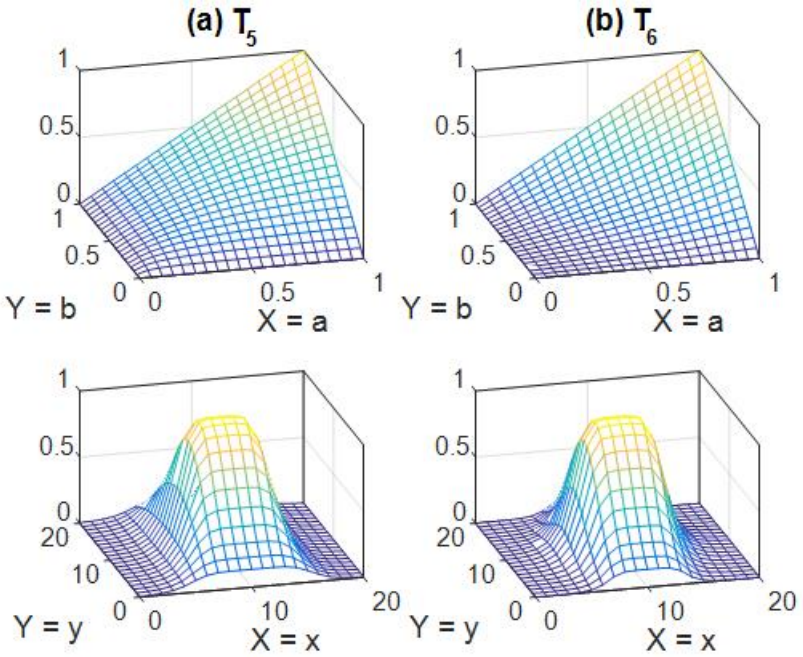
Şekil 2.1. *Gaussian ve G-Bell üyelik fonksiyonları*



Şekil 2.2. *Min. ve Çarpım T-normlarının gösterimi*

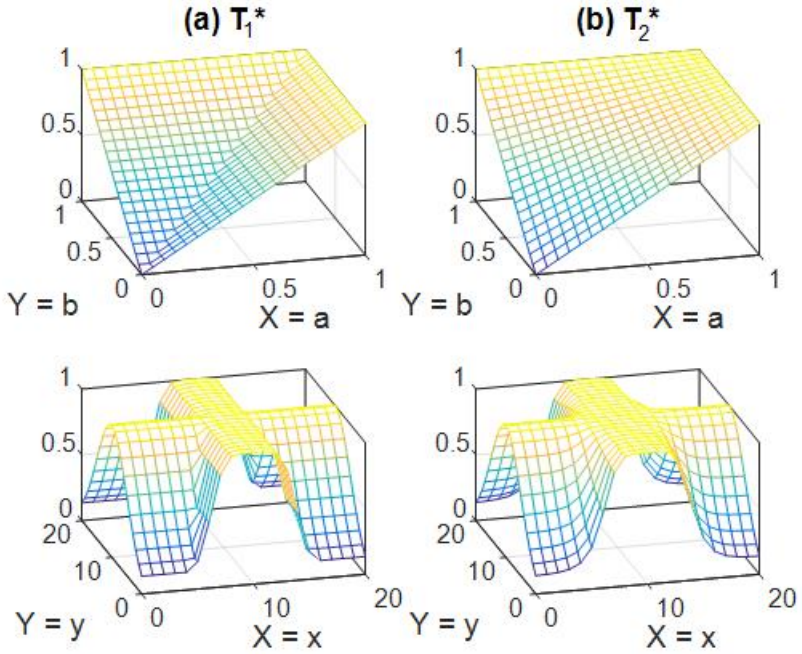


Şekil 2.3. *Sınırlandırılmış ve Güçlü(drastic) çarpım T-normlarının gösterimi*

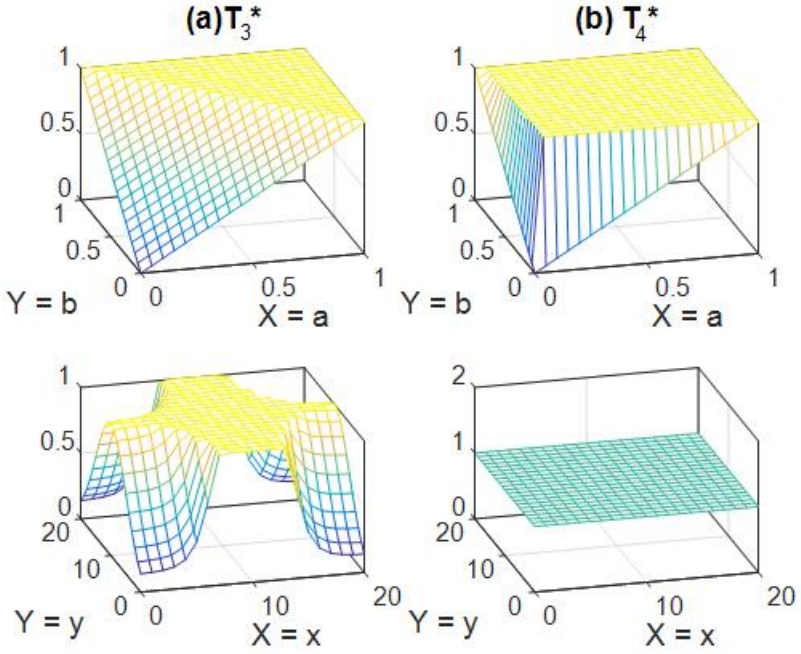


Şekil 2.4. *Hamacher ve Einstein çarpım T-normlarının gösterimi*

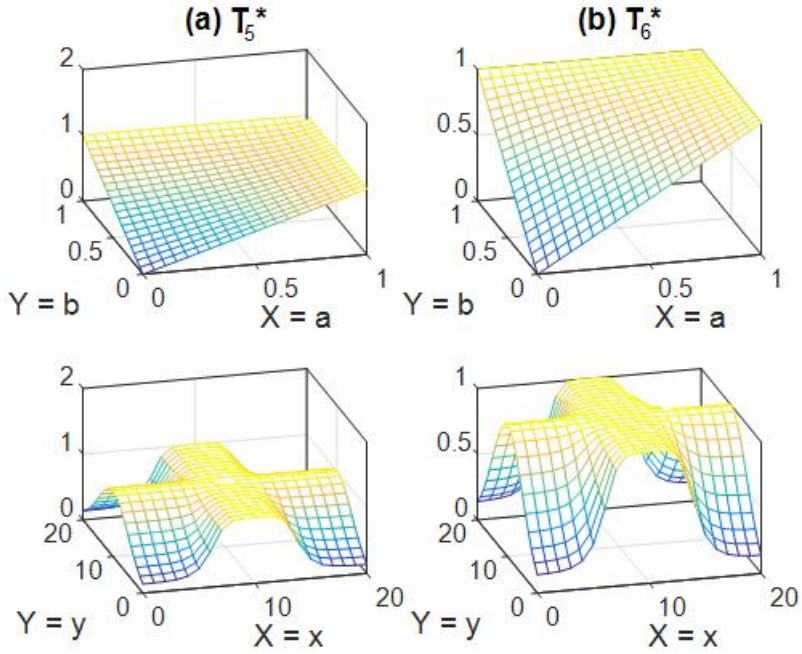
Şekil 2.2, 2.3 ve 2.4'de bulunan ilk satırlar T-normların tanımını ifade ederken (0 ile 1 arasındaki noktasal ilişkileri), ikinci satırlar ise A ve B üyelik fonksiyonlarının faklı T-norm yapılarına göre kesişimini temsil etmektedir. Ayrıca A ve B üyelik fonksiyonlarının T-conorm yapılarının da grafikleri verilmiştir [8].



Şekil 2.5. *Max. ve Çarpım T-conorm yapılarının gösterimi*



Şekil 2.6. *Sınırlandırılmış ve Güçlü(drastic) T-conorm yapılarının gösterimi*



Şekil 2.7. *Hamacher ve Einstein T-conorm yapılarının gösterimi*

Şekil 2.5, 2.6 ve 2.7'de bulunan ilk satırlar T-conormların tanımı ifade ederken (0 ile 1 arasındaki noktasal ilişkileri), ayrıca her şekildeki ilk satırlar 0 ile 1 arasındaki sayıların 0,05 artımla, karşılıklı T-norm veya T-conorm sonuçlarıdır. ikinci satırlar ise A ve B üyelik fonksiyonlarının faklı T-conorm yapılarına göre birleşimini temsil

etmektedir. Tezde, PH ve Einstein T-normları kullanılacağı için, MISO (Multı Input Single Output) -PH ve Einstein yapılarına da yer verilmesi gerekmektedir.

2.1.1 Parametrik Hamacher T-norm yapısının genelleştirilmesi

2014 yılında Zhang ve Liao [11]'in yaptığı çalışmada Hamacher yapısı genelleştirilmiş ve genelleştirilen yapı üzerinde, bazı tanımlar ve önermeler verilmiştir. Bu tanımlar ve önermelerden yola çıkılarak, çok girdili Hamacher yapısının tanımı yapılmaktadır.

2.3 ***Tanım:*** $\mathrm{T}_\lambda(\mathrm{A_n})$, $\mathrm{A_n}$ üzerinde çok girdili bir Hamacher T-norm olsun. Burada,

$\mathrm{A_n} = \{\mathrm{a}_1, \mathrm{a}_2, \dots, \mathrm{a}_n\}$, $1 \leq t \leq n$ olacak şekilde a_t'ler üyelik fonksiyonlarıdır. Ayrıca $\forall i \in N^+, 2 \leq n, 0 \leq a_i \leq 1$, $\mathrm{T}_\lambda(\mathrm{A_n}) = \mathrm{T}_\lambda(\mathrm{T}_\lambda(\mathrm{A_{n-1}}), \mathrm{a_n})$, $n = 2$ olduğu zaman $\mathrm{T}_\lambda(\mathrm{A}_2) = \mathrm{T}_\lambda(\mathrm{a}_1, \mathrm{a}_2)$ elde edilmektedir. PH'in yapısını ifade etmek için, temel teşkil eden fonksiyon verilecek olursa,

$\chi^{\,j}(A_n) = \sum_{c_1,\dots,c_j \in \{1,\dots,n\}, c_1 \neq \dots \neq c_j} (a_{c_1}, a_{c_2}, a_{c_3}, \dots a_{c_j})$ $n \in N^+, n \geq 2$ ve $j \in N$,

$j \leq n$ şeklindedir [11].

Burada PH T-normun önermesi, tanımlanan, $\chi^{\,j}(A_n)$ fonksiyonu üzerinden yapılacaktır [11]. Özel olarak, $\chi^{\,0}(A_n) = 1$ şeklindedir. Ayrıca örnek olarak, $\chi^{\,2}(A_4) = \mathrm{a}_1.\mathrm{a}_2 + \mathrm{a}_1.\mathrm{a}_3 + \mathrm{a}_1.\mathrm{a}_4 + \mathrm{a}_2.\mathrm{a}_3 + \mathrm{a}_2.\mathrm{a}_4 + \mathrm{a}_3.\mathrm{a}_4$ verilebilir.

Zhang ve Liao Tanım 2.3'de çok girdili Hamacher yapısını tanımlayabilmek için, $\chi^{\,j}(A_n)$ yardımcı fonksiyonunu tanımlamışlardır. Fonksiyonu sözel olarak ifade etmek gerekirse; terimleri j'li kombinasyonlarının, n.üyelik fonksiyonuna kadar olan bileşenlerinin toplamı, şeklinde ifade edilebilir.

2.1 ***Sonuç:*** $\frac{\partial \chi^{\,j}(A_n)}{\partial a_k} = \chi^{\,j-1}(A_n \backslash a_k)$, $\boldsymbol{j, k \in N}$ ve $\boldsymbol{j, k \neq n}$.

$\boldsymbol{A_n \backslash a_k} = \{\boldsymbol{a_1, a_2, \dots, a_{k-1}, a_{k+1}, \dots, a_n}\}$ olarak tanımlanmaktadır [11].

Sonuç 2.1, tanım 2.3'de yer alan, χ fonksiyonunun, üyelik fonksiyonlarına göre türevidir. Bu sonuç, PH T-norm'un genelleştirilmesi verildikten sonra, üyelik fonksiyonlarına göre türevini bulmak için kullanılmaktadır.

2.1 ***Önerme:*** $\boldsymbol{T_\lambda(A_n)} = \frac{\chi^{\,n}(A_n)}{\lambda^{n-1} + \sum_{j=1}^{n-1} \lambda^{n-j-1}(1-\lambda)^j \chi^{\,j}(A_n) - \sum_{i=1}^{n-1}(1-\lambda)^i \chi^{\,n}(A_n)}$ $, \boldsymbol{\lambda \in} \mathbb{R}^+$

Önerme 2.1'de PH T-norm'un genelleştirilmiş hali verilmektedir. Çok girdili bir yapıda bu önerme kullanılmaktadır.

2.2 ***Önerme:*** $\frac{\partial T_\lambda(A_n)}{\partial \lambda} = \frac{\chi^n(A_n)R_n}{Q_n^2}$, burada

$$Q_n = \lambda^{n-1} + \sum_{i=1}^{n-1} \lambda^{n-i-1}(1-\lambda)^i \chi^i(A_n) - \sum_{i=1}^{n-1}(1-\lambda)^i \chi^n(A_n) \quad \textbf{(2.1)}$$

$$R_n = (n-1)\lambda^{n-2} + \sum_{i=1}^{n-1} \lambda^{n-i-2}(1-\lambda)^{i-1}[(n-1)-(n-1)\lambda - i]\chi^i(A_n) + \sum_{i=1}^{n-1} i(1-\lambda)^{i-1}\chi^n(A_n) \quad \textbf{(2.2)}$$

Önerme 2.2'de PH T-normun parametresi olan λ'ya göre türevi verilmiştir. 2.1 sonuç, 2.1 önerme ve 2.2 önermenin ispatları Zhang ve Liao'nun çalışmasında, tümevarım yöntemi ile ispatlanmıştır [11]. Sonuç ve önermelerin ışığında, PH'nin üyelik fonksiyonlarına göre türevi bulunabilir.

$$\frac{\partial T_\lambda(A_n)}{\partial a_k} = \frac{(\chi^n(A_n)/a_k)Q_n - (\sum_{j=1}^{n-1} \lambda^{n-j-1}(1-\lambda)^j \chi^{j-1}(A_n \backslash a_k)}{Q_n^2} - \frac{\sum_{i=1}^{n-1}(1-\lambda)^i(\chi^n(A_n)/a_k))\chi^n(A_n)}{Q_n^2} \quad \textbf{(2.3)}$$

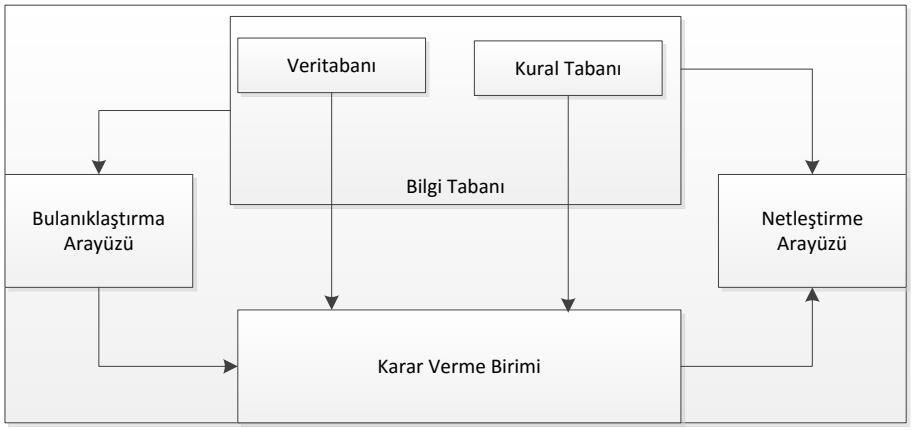
2.2 Adaptif Sinirsel Bulanık Çıkarım Sistemi

Adaptif Sinirsel Bulanık Çıkarım Sistemi, insan bilgi ve deneyimlerini kural tabanlı bir sisteme dönüştürmek için kullanılan bir sistemdir. ANFIS'in yapısında bir yandan bulanık kurallar yardımıyla bulanık çıkarım sistemi çalışırken diğer yandan ise yapay sinir ağları ile öğrenme gerçekleşmektedir. ANFIS'in temelinde Takagi ve Sugeno bulanık çıkarım sistemi bulunmaktadır. Bu yüzden bulanık çıkarım sistemini, ana hatları ile özetlememiz gerekmektedir.

2.2.1 Bulanık çıkarım sistemi

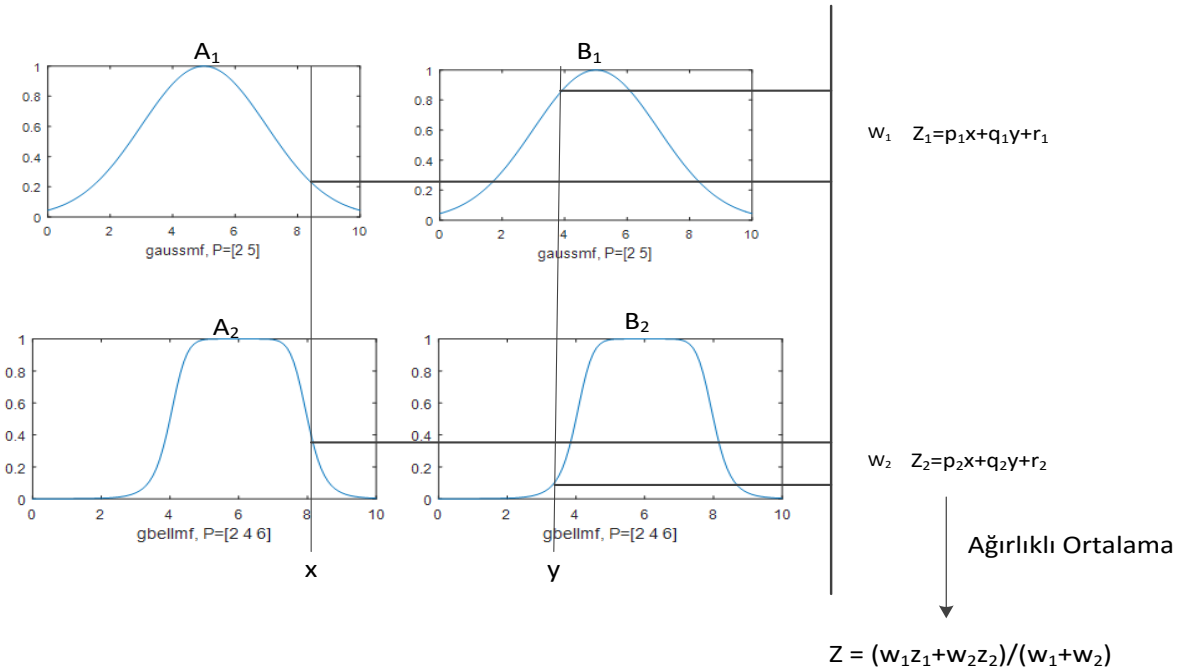
Bulanık çıkarım sistemi 5 ana bloktan oluşmaktadır. Bulanık kuralların sayısını içeren kural tabanı, bulanık kurallarda kullanılan bulanık kümelerin tanımlanan üyelik fonksiyonlarının veri tabanı, kuralların üzerinde çıkarım operasyonlarını gerçekleyen karar verme birimi, dış ortamdan alınan kesin değerleri (crisp values), bulanık değerlere

dönüştüren bulanıklaştırma ara yüzü ve son olarak bulanık değerleri kesin değerlere dönüştüren netleştirme biriminden oluşmaktadır. Sistemin yapısı Şekil 2.8'de şematize edilmiştir [1].



Şekil 2.8. *Bulanık Çıkarım Sistemi*

Birden fazla bulanık çıkarım yöntemi bulunmaktadır. En sık kullanılan bulanık çıkarım yöntemleri, Mamdani ve Takagi-Sugeno-Kang yöntemleridir. Bulanık sinir ağları anahatları ile Takagi-Sugeno-Kang çıkarım sistemine benzer yapıdadır. Bu yapıda, her bir kuralın çıktısı sabit bir terim ile birlikte girdi değişkenlerinin lineer birleşimleri şeklinde oluşturulur. Nihai çıktı ise her bir kuralın çıktısının ağırlıklandırılmış ortalamasıdır.



Şekil 2.9. Sugeno tip bulanık çıkarım sistemi

Şekil 2.9'da girdileri x ve y olup, çıktısı z fonksiyonu olan birinci derece Sugeno tip bulanık çıkarım sistemi gösterilmektedir. Kuralları ile birlikte sistem ifade edilecek olursa;

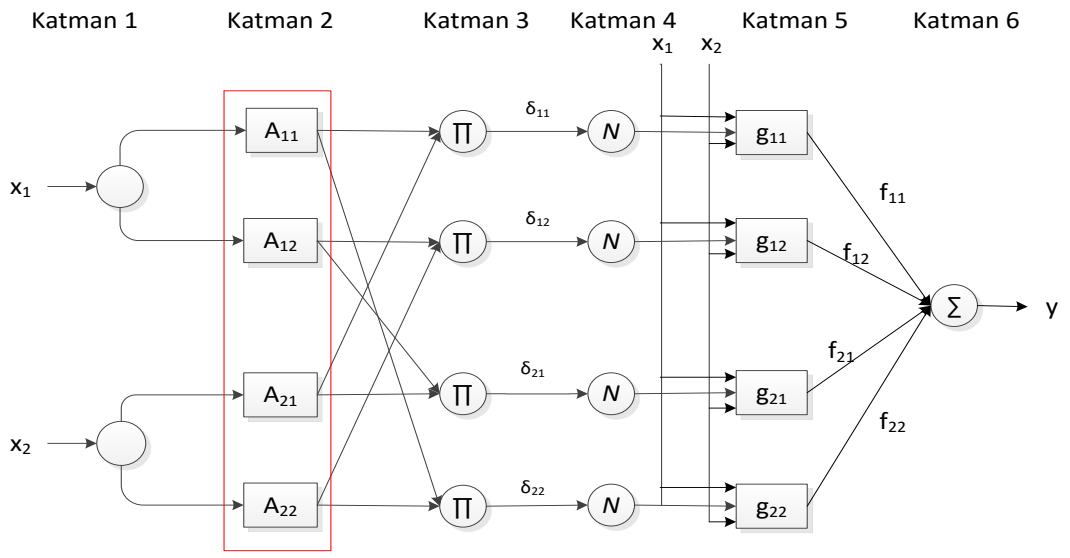
Kural 1: **Eğer** $x = A_1$ ve y = B_1 ise **O Halde** $z_1 = p_1x + q_1y + r_1$

Kural 2: **Eğer** $x = A_2$ ve y = B_2 ise **O Halde** $z_2 = p_2x + q_2y + r_2$

$i = 1,2$ için, x ve y girdi değişkenlerini, z_i çıkış değişkenini, A_i ve B_i üyelik fonksiyonlarını ve p_i ,q_i ,$r_i \in \mathbb{R}$ ise sonuç değişkenlerini ifade etmektedir.

2.2.2 ANFIS-AWN mimarisi

ANFIS'in Sugeno tipi bir bulanık çıkarım sistemi olduğundan daha önce bahsedilmişti. Sugeno çıkarım sisteminde de olduğu gibi bulanık kuralların IF (EĞER, antecenet part) kısmı üyelik fonksiyonlarından oluşmaktadır. Üçgensel (triangular), bell-shape, trapezoidal, Gaussian gibi birçok üyelik fonksiyonu bulunmaktadır Kuralların sonuç kısmı ise girdilerin lineer fonksiyonu ya da sabit terim olarak alınmaktadır. Üyelik fonksiyonlarında bulunan bilinmeyen parametreler ile çıktı kısmında bulunan bilinmeyen parametreler, hibrit öğrenme algoritması ile güncellenmektedir.



Şekil 2.10. *ANFIS-AWN yapısı*

Şekil 2.10'da örnek olarak iki girdili ve tek çıktıya sahip bir ANFIS-AWN yapısı yer almaktadır. Yapıda, her girdi için ikişer tane üyelik fonksiyonu bulunup, toplam dört kuraldan oluşmaktadır.

Kural 1: **Eğer** $x_1 = A_{11}$ **VE** $x_2 = A_{21}$ **O Halde** $g_{11} = p_{11}x_1 + q_{11}x_2 + r_{11}$

Kural 2: **Eğer** $x_1 = A_{12}$ **VE** $x_2 = A_{22}$ **O Halde** $g_{12} = p_{12}x_1 + q_{12}x_2 + r_{12}$

Kural 3: **Eğer** $x_1 = A_{12}$ **VE** $x_2 = A_{21}$ **O Halde** $g_{21} = p_{21}x_1 + q_{21}x_2 + r_{21}$

Kural 4: **Eğer** $x_1 = A_{11}$ **VE** $x_2 = A_{22}$ **O Halde** $g_{22} = p_{22}x_1 + q_{22}x_2 + r_{22}$

ANFIS ve AWN yapılarının sadece ikinci katmanı farklı olduğu için iki yapının da 6 katmanı aşağıda özetlenmiştir.

- Katman 1: Bu katman, girdi katmanıdır. Bu katmandaki her bir nöron dış ortamdan alınan sinyaller direkt bir sonraki katmana aktarılmaktadır.
- Katman 2: Bu katman, bulanıklaştırma katmanıdır. Kuralların **Eğer** kısmındaki üyelik fonksiyonları bu katmanda bulunmaktadır. Bu katmanın çıktısı, bulunduğu nörondaki üyelik fonksiyonunun değerinden oluşmaktadır. Üyelik fonksiyonu olarak Gaussian üyelik fonksiyonu kullanıldığını düşünürsek;

$$A_{ij}(x_i) = e^{-\frac{1}{2}\left(\frac{x_i-\mu_{ij}}{\sigma_{ij}}\right)^2} \qquad i = 1,2 \; ve \; j = 1,2 \qquad \textbf{(2.4)}$$

AWN yapısında ise kırmız çizgili kısımdaki üyelik fonksiyonları, dalgacık fonksiyonu ile yer değiştirmektedir.

Dalgacık fonksiyonu, Gaussian fonksiyonunun girdi bileşenine göreve türevidir. Dalgacık fonksiyonu i. nöron için 2.5'de ifade edilmektedir.

$$\phi_{ij}\left(\frac{x_i-\mu_{ij}}{\sigma_{ij}}\right) = (1-\left(\frac{x_i-\mu_{ij}}{\sigma_{ij}}\right)^2)e^{-\frac{1}{2}\left(\frac{x_i-\mu_{ij}}{\sigma_{ij}}\right)^2} \qquad i = 1,2\ ve\ \ j = 1,2 \qquad \textbf{(2.5)}$$

- Katman 3: Kural katmanıdır. Bu katmandaki her bir nöron bulanık kuralların işleyiş derecesini (firing strength) temsil etmektedir. Bu katmanda faklı birçok T-norm yapısı kullanılabilir. Ancak genellikle kullanılan T-norm yapıları minimum veya çarpım T-normudur. Bu kısımda geleneksel ANFIS yapısında kullanılan çarpım T-normu kullanılmıştır.

$$\delta_{ij} = \ A_{1i}(x_1).\ A_{2j}(x_2) \qquad i = 1,2\ ve\ j = 1,2 \qquad \textbf{(2.6)}$$

- Katman 4: Normalizasyon katmanıdır. İşleyiş derecelerinin normalize edilmesi işlemi bu katmanda gerçekleşir.

$$\overline{\delta_{ij}} = \frac{\delta_{ij}}{\sum_{i=1}^{2}\sum_{j=1}^{2}\ \delta_{ij}} \qquad i = 1,2\ ve\ j = 1,2 \qquad \textbf{(2.7)}$$

- Katman 5: Netleştirme katmanıdır. Bu katmanda girdi değerlerinin lineer birleşimleri (sabit terim ile birlikte) ile önceki katmanda elde edilen normalize işleyiş dereceleri çarpılır.

$$f_{ij} = \overline{\delta_{ij}}g_{ij}$$
$$g_{ij} = p_{ij}x_1 + q_{ij}x_2 + r_{ij} \qquad i = 1,2\ ve\ j = 1,2 \qquad \textbf{(2.8)}$$

- Katman 6: Çıktı katmanı olup, sistemin nihai çıktısı oluşturulur.

$$y = \sum_{i=1}^{2}\sum_{j=1}^{2} f_{ij} \qquad i = 1,2\ ve\ j = 1,2 \qquad \textbf{(2.9)}$$

ANFIS yapısındaki bilinmeyen parametrelerin öğrenime katılması için hibrit öğrenme algoritması bulunmaktadır. Hibrit öğrenme, ileri geçiş (forward pass) ve geri

geçiş (backward pass) olmak üzere ikiye ayrılmaktadır. Girdi parametreleri için türev tabanlı azalma algoritması (gradient descent) uygulanırken, çıktı parametreleri için ise en küçük kareler tahmin edicisi (Least square estimation) kullanılır. İleri geçiş aşamasında girdi parametreleri sabit iken çıktı parametreleri en küçük kareler tahmin edicisi ile güncellenmektedir. Geri geçiş kısmında ise girdi parametreleri türev tabanlı azalma algoritması ile güncellenirken çıktı parametreleri sabit tutulmaktadır [1].

AWN'de, 2. katmanda genel olarak kullanılan Gaussian fonksiyonu yerine, zaman-frekans uyumuna sahip olan [58] dalgacık dönüşümü kullanılmaktadır [59]. Diğer 5 katman da ANFIS ile benzer özellikleri taşımaktadır. ANFIS gibi çıktı kısmında yer alan fonksiyonun ismine göre adlandırılmaktadır. Çıktı kısmında sabit bir sayı bulunuyorsa sıfırıncı dereceden AWN, eğer girdilerin lineer birleşimlerinden oluşan bir fonksiyon bulunuyorsa, birinci dereceden AWN denilmektedir. Şekli 2.11'de örnek olarak, iki girdi ve tek çıktıya sahip, AWN modeli yer almaktadır. Parametre açısından da ele alınırsa; modelde her bir girdi için ikişer tane dalgacık fonksiyonu bulunmaktadır. Her fonksiyonda da μ, σ parametrelerinden ikişer tane olduğu göz önüne alınırsa, 8 tane dalgacık fonksiyonu parametresi bulunmaktadır. Çıktı kısmında model, sıfırıncı dereceden ise her kural için bir sabit terim olacağından toplam 4 parametre mevcuttur. Ancak model birinci dereceden ise sistem iki girdili olduğundan, her bir kuralda 3 tane parametre bulunmaktadır. Kural sayısının ise 4 olduğu göz önünde bulundurulursa, iki girdi ve tek çıktıya sahip bir AWN modelindeki çıktı parametresi sayısı toplam 12'dir.

Öğrenme algoritması olarak, türev tabanlı öğrenme algoritması kullanılmaktadır. Öncelikle çıktı parametrelerinin başlangıç değerlerinin belirlemek için bir defa en küçük kareler tahmin edicisi kullanılır. Daha sonra ise bütün parametreler, türev tabanlı bir optimizasyon yöntemi ile güncellenir.

2.3 Tezde Sunulan Modeller

Tezde iki farkı model sunulmaktadır. Tezde sunulan ilk model ANFIS ve AWN modellerinin 3. katmanında bulunan çarpım T-normu yerine, Einstein ve PH T-normu kullanılmıştır. Bu yapı kaotik zaman serileri ve sistem tanıma problemi üzerinde uygulanmıştır. PH T-normun λ parametresi, öğrenme algoritmasına katılmıştır. Öğrenme algoritması sonucunda λ'nın uygun değerleri bulunmuştur. Einstein T-norm, PH T-norm un özel bir halidir. PH T-norm, farklı λ değerleri için farklı T-norm yapıları üretmektedir.

Bu yüzden sunulan modellerde PH T-norm kullanılmıştır. Sistemin yapısına 3. bölümde ayrıntılı şekilde yer verilmiştir.

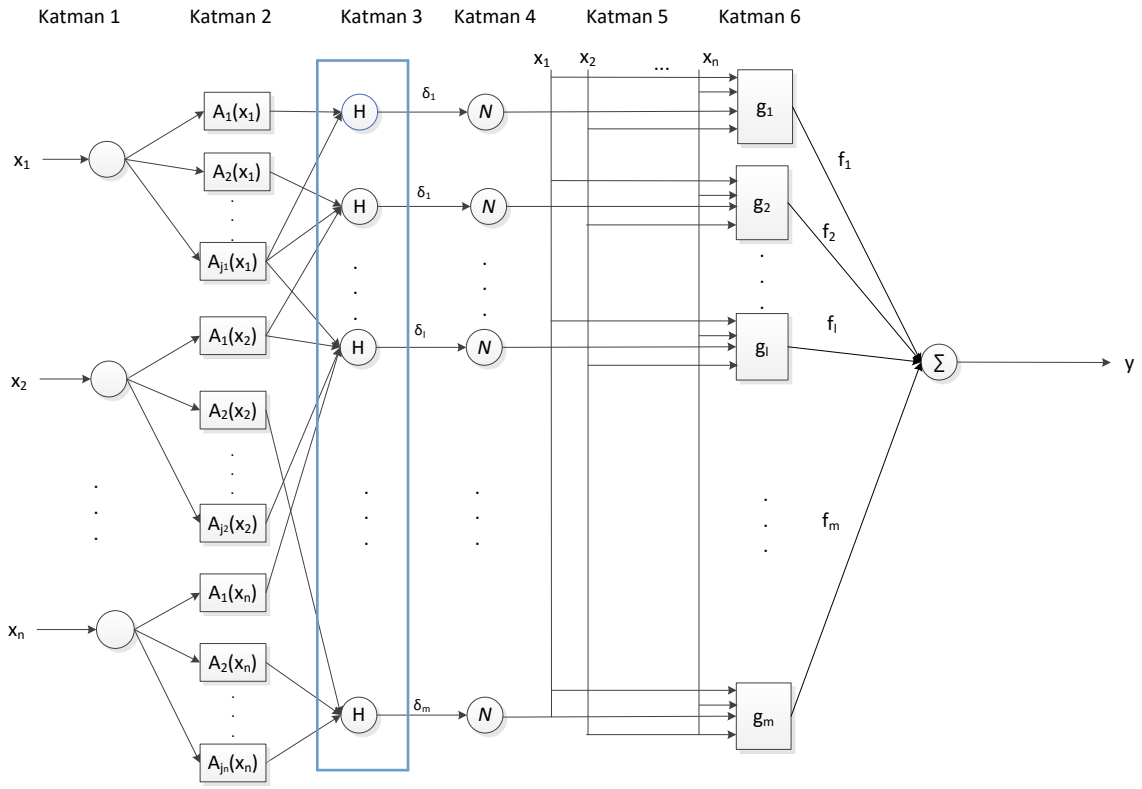
Tezde sunulan diğer model ise topluluk yapısının 3. bölümde sunulan modellerin üzerinde uygulanıp elde edilen çıktıların, ağırlıklandırılmış ortalamalarına, öğrenme algoritmasının uygulanmasıdır. Bu yapıda ağırlıklandırılmış ortalamaların her bir ağırlığı öğrenme algoritması ile belirlenmektedir. ANFIS için topluluk kavramını özetlemek gerekirse; farklı yapılarda n tane ANFIS alınır ve daha sonra her bir ANFIS'in çıktısı, belirlenen bir integrator için girdi teşkil eder. Burada "integrator", çıktı değerleri için entegrasyonu sağlayan, ortalama veya ağırlıklı ortalamanın hesaplanmasıdır. "Farklı yapılarda n tane ANFIS" kısmında ise vurgulanmak istenen, n tane ANFIS'in her biri için 2. katmanda kullanılan üyelik fonksiyonları farklı olabildiği gibi her bir ANFIS yapısı farklı girdi değerlerine de sahip olabilir. Her bir ANFIS'in 3. katmanında kullanılan çarpım T-normu yerine, her bir ANFIS'de farklı T-normlar da kullanılabilir. Tezde sunulan modelde, girdi için 3 ANFIS yapısı, finansal zaman serisi kestirimi üzerinde uygulanmıştır. İlki ANFIS 3, 2 ve 1. gün önceki verilerden yararlanarak bir sonraki gündeki veriyi tahmin için kullanılmıştır. İkinci ve sonuncu ANFIS de ise sırası ile 2, 4 ve 6 gün önceki, 1, 3 ve 5 gün önceki veriler kullanılmıştır. Ayrıca T-norm olarak da sistem üzerinde farklı T-normlar denenmiştir. Sistemin yapısı 4. bölümde ayrıntılı şekilde incelenmektedir.

3. ANFIS ve AWN YAPILARINDA FARKLI T-NORMLARIN KULLANIMI

Bu bölümde ANFIS ve AWN yapılarında, çarpım ve min. T-norm yapıları yerine, PH ve Einstein T-normlarının kullanılması amaçlanmıştır. Yapıların eğitimi için, türev tabanlı BFGS algoritması kullanılmıştır.

3.1 ANFIS-Einstein ve ANFIS-PH Yapıları

ANFIS-Einstein ve ANFIS-PH yapıları, ANFIS yapısına benzer olarak, katmanlar şeklinde ifade edilmiştir. Sunulan yapıları, çok girdili ve tek çıktılı bir sisteme genellemek mümkündür.



Şekil 3.1. *ANFIS-Parametrik Hamacher yapısı (mavi çizgili kısım)*

Şekil 3.1'de çok girdili ve tek çıktılı ASBÇ – PH yapısı verilmiştir. Görüldüğü gibi 3. katmanda Hamacher T-normu ifade eden "H" harfi kullanılmaktadır. ANFIS-

Einstein ve ANFIS-PH yapıları arasında sadece 3. katmanda farklılıklar olduğu için, katmanlar açıklanırken, Hamacher ve Einstein T-norm yapıları beraber verilmiştir. Bu yapıyı 6 katman şeklinde açıklayacak olursak;

- **Katman 1**: Bu katmanda n tane giriş olan, $x_1, x_2 , \dots , x_n$ doğrudan 2. katmana iletilir.
- **Katman 2**: Bulanıklaştırma katmanıdır. ANFIS yapısında da olduğu gibi, üyelik fonksiyonları olarak, Gaussian üyelik fonksiyonu kullanılmaktadır. Burada 1. girdi için j_1 tane, ikinci girdi için j_2 tane ve nihayetinde n. girdi için ise j_n tane üyelik fonksiyonu mevcuttur. Aşağıdaki Gaussian denklemi k. girdi için i_k. üyelik fonksiyonunu göstermektedir.

$$A_{i_k}(x_k) = \exp(-\frac{1}{2}\left(\frac{x_k-\mu_{i_k}}{\sigma_{i_k}}\right)^2) \ \ (k = 1,2, \dots, n \text{ ve } i_k = 1,2, \dots, j_k) \tag{3.1}$$

- **Katman 3**: Bu katmanda diğer ANFIS yapılarından farklı olarak PH ve Einstein T-norm yapıları kullanılmıştır. Bu katmanda, her bir kuralda hesaplanan değer, o kuralın işleyiş derecesi olarak adlandırılır.

<u>Parametrik-Hamacher</u>

$$\frac{\chi^{\,n}(A_n)}{\lambda^{n-1}+\sum_{j=1}^{n-1}\lambda^{n-j-1}(1-\lambda)^j\chi^{\,j}(A_n)-\sum_{i=1}^{n-1}(1-\lambda)^i\chi^{\,n}(A_n)} \tag{3.2}$$

Buradaki n, sistemdeki girdi sayısını vermektedir. Denklem 3.2 PH T-normun toplam sembolü ile genelleştirilmesini göstermektedir. Toplam sembolü ile genelleştirme yapılmasının sebebi ise, toplam sembolü üzerinde gradyan hesaplamalarının daha kolay olmasıdır. Çarpım sembolü ile genelleştirilmesi ise Dombi tarafından yapılmıştır [64]. Bu katmandaki PH çarpımının genelleştirilmesi, 3.3'te verilmektedir.

$$\delta_l = \frac{1}{1+\frac{1}{\lambda}(\prod_{k=1}^{n}\left(1+\lambda\frac{1-A_{i_k}(x_k)}{A_{i_k}(x_k)}\right)-1)} \tag{3.3}$$

$\lambda = 2$ için Einstein T-norm elde edilmektedir.

$$\delta_l = \frac{1}{1+\frac{1}{2}(\prod_{k=1}^{n}\left(1+2\frac{1-A_{i_k}(x_k)}{A_{i_k}(x_k)}\right)-1)} \tag{3.4}$$

Burada $l = i_1, i_2, \dots, i_n$, $i_1 = 1, \dots, j_1$, $i_2 = 1, \dots, j_2$, $\dots$, $i_n = 1, \dots, j_n$,

- **Katman 4**: Normalizasyon katmanıdır. Bu katmanın her bir çıktısı, 3. katmandan elde edilen her bir işleyiş derecesinin, bütün işleyiş derecelerine oranı olarak alınmaktadır.

$$\overline{\delta_l} = \frac{\delta_l}{\sum_{i=1}^{m} \delta_i} \qquad (l = 1, \dots, m) \qquad \textbf{(3.5)}$$

- **Katman 5**: Netleştirme katmanıdır. Bu katmanda birinci dereceden lineer fonksiyon kullanılacaktır. Çıktı parametreleri ile girdilerin lineer birleşimlerinden oluşmaktadır.

$$f_l = \overline{\delta_l} g_l \qquad (l = 1, \dots, m) \qquad \textbf{(3.6)}$$
$$g_l = q_l + \sum_{i=1}^{n} p_{il} x_i$$

- **Katman 6**: Bu katman, gelen bütün sinyalleri toplayıp, sonuç kısmına iletilmesini sağlamaktadır.

$$y = \sum_{l=1}^{m} f_l \qquad \textbf{(3.7)}$$

Birçok T-norm yapısı bulunmaktadır. Fakat ANFIS için hangi T-norm yapısının daha önemli olduğu henüz bilinmemektedir. Ancak sistemde yer alan PH yapısının λ'nın farklı değerleri için, bazı T-normları kapsadığı görüldüğünden, ANFIS yapısında kullanılmasının uygun olduğu düşünülmüştür. Sunulan parametrik modelin, ANFIS yapısından tek farkı 3. katmanda farklı T-norm yapılarının kullanılmasıdır. PH T-normun yapısındaki λ değeri değiştikçe, farklı T-norm yapıları oluşmaktadır.

Bu yapıya en uygun sistem, 2014 yılında Zhang ve Liao tarafından sunulmuştur [11]. Çalışmalarında çok girdili Hamacher yapısını kullanmışlardır. Çalışmada, λ parametresi öğrenme algoritması ile güncellenmiştir. Ancak her bir kural değeri için, ayrı λ değerleri belirlenmiştir. Bu durum, aynı iterasyonda bütün bir ANFIS yapısı içerisinde farklı λ değerleri kullanılmasına sebep olmuştur. Bu durum ise parametre sayısının artmasına ve dolayısıyla, öğrenmenin yavaşlamasına sebep olmaktadır. Ayrıca sonuçları tek bir veri seti üzerinde karşılaştırmışlardır. Tezde sunulan yapıda ise her bir iterasyonda

λ değerleri, ayrı ayrı güncellenmektedir. Yani bütün bir sistem için en son, tek bir uygun λ değeri bulunmaktadır. Sonuçlar ise üç farklı veri seti üzerinde denenmiştir. Veri setleri hakkındaki detaylı bilgiye 5. Bölümde yer verilecektir.

3.1.1 ANFIS-Einstein ve ANFIS-PH için eğitim algoritması

Tezde sunulan bütün yapılarda aynı eğitim algoritması kullanılmıştır. Genel anlamda ANFIS-PH yapısı üzerinden bütün parametreler incelenirse; girdi parametreleri, μ (merkez), σ (genişlik), T-norma ait parametre λ ve çıktı parametrelerinin p ve q değişkenleri olduğu görülmektedir. Şekil 3.1'de yer verilen, genelleştirilmiş yapıdaki, bütün parametreler hesaba katılacak olursa; μ ve σ parametreleri, üyelik fonksiyonlarında yer alan parametrelerdir. Her girdi için faklı sayıda üyelik fonksiyonu olduğunu göz önüne alalım. O halde, Şekil 3.1'in 2. katmanında da bahsedildiği gibi, n. girdi için j_n tane üyelik fonksiyonu bulunmaktadır. μ ve σ parametreleri: $\sum_{i=1}^{n} j_i$ şeklinde gösterilebilir. q parametresi ise çıktı kısmındaki birinci dereceden fonksiyonun sabit terimidir. Kural sayısı kadar q olacağından, Şekil 3.1'de görüldüğü gibi q parametrelerinin sayısı ise m tanedir. Burada $m = \prod_{i=1}^{n} l_i$ toplam kural sayısı olarak bulunmaktadır. p parametresi, girdilerin lineer birleşimi şeklinde yazıldığından, toplam $m * n$ tane bulunmaktadır. λ parametresi de bütün bir sistem için 1 tane bulunmaktadır. O halde bir ANFIS-PH yapısında bulunan, parametrelerin toplamı: $2* \sum_{i=1}^{n} j_i + \mathrm{m} + \mathrm{m} * \mathrm{n} + 1$ olarak bulunmaktadır. Sistemde bütün parametreler, gradyan tabanlı BFGS metodu ile güncellenmektedir. Yapılar eğitilirken, hedef fonksiyon olarak, ortalama hata kareler yöntemi seçilmiştir. Amaç hata fonksiyonu olan, ortalama kareler hatayı minimize etmektir.

$$E = \frac{1}{N} \sum_{k=1}^{N} (y - y_g)^2 \tag{3.8}$$

Burada N girdi sayısı iken, y, ANFIS-PH'ın çıktısı ve y_g ise gerçek değerleridir. $\lambda = 2$ için Einstein T-norm sağlandığından, bütün parametreler ortaktır. Ancak λ değeri belli olduğu için toplam $2* \sum_{i=1}^{n} j_i + \mathrm{m} + \mathrm{m} * \mathrm{n}$ parametreye sahiptir.

3.1.2 Broyden-Fletcher-Goldfarb-Shanno gradient yöntemi

BFGS, en hızlı ve en ideal yakınsamaya sahip, türev tabanlı, öğrenme algoritmalarından biridir [60]. BFGS, Quasi-Newton yapısından türeyen bir optimizasyon tekniğidir. Sınırlandırılmamış bir optimizasyon problemini düşünecek olursak;

$$\min f(x),\ x \in \mathbb{R}^n \tag{3.9}$$

Burada f, ikinci türeve sahip, sürekli bir fonksiyondur. BFGS'in genel algoritma yapısı aşağıdaki gibidir.

$$x_{k+1} = x_k - \alpha_k H_k g_k \tag{3.10}$$

Burada x_1, başlangıç vektörü, x vektörü ise güncellenecek parametre vektörüdür. H_k, Hessian matrisinin tersidir ve $[\nabla f(x_k)]^{-1}$ şeklinde ifade edilir.

$$g_k = \nabla f(x_k) \tag{3.11}$$

x, bilinmeyen parametrelerden oluşan bir parametre vektörüdür, her iterasyonda güncellemektedir. (3.10)'da verilen denklem performans ölçütünün, bütün bilinmeyen parametrelere göre türevidir. α_k ise arama yönü ile belirlenen optimal adım büyüklüğüdür.

$$d_k = -H_k g_k \tag{3.12}$$

(3.11) da verilen denklem algoritmanın arama yönüdür. Eğer algoritma optimal değerlerden uzaklaşırsa, arama yönü değişmekte, yeni bir yön aramaktadır. BFGS'de H_{k+1} matrisi, (3.12) yer alan denklem ile güncellenmektedir [60].
$\gamma_k = g_{k+1} - g_k$, $\delta_k = x_{k+1} - x_k$ olmak üzere;

$$H_{k+1} = \left[I - \frac{\delta_k(\gamma_k)^T}{(\delta_k)^T\gamma_k}\right] H_k \left[I - \frac{\delta_k(\gamma_k)^T}{(\delta_k)^T\gamma_k}\right]^T + \frac{\delta_k(\delta_k)^T}{(\delta_k)^T\gamma_k} \tag{3.13}$$

Burada γ_k, gradient vektörünün bir önceki adımla olan farkı iken, δ_k ise parametre vektörünün bir öneki adım ile olan farkıdır. Hessian matrisinin başlangıç değeri, birim matris olarak alınmaktadır. Yani $H_0 = I$ şeklinde ifade edilmektedir.

3.1.3 ANFIS-PH için parametre gradyanları

Tez çalışmasında, gradyan tabanlı BFGS algoritması kullanılacağı için, hedef fonksiyonun, bütün parametrelere göre birinci dereceden türevi gerekmektedir. Aşağıda gerekli bütün gradyanlar hesaplanmıştır.

Öncelikle, diğer gradyanlarda da kullanılacak olan, hatanın çıkış fonksiyonuna göre türevi bulunmalıdır. Bu türev (3.14)'da verilmektedir.

$$\frac{\partial E}{\partial y} = \frac{2}{N}\sum_{k=1}^{N}(y - y_g) \tag{3.14}$$

(3.15) ve (3.16)'de çıktı parametrelerine göre türevler bulunmaktadır. Türevler, zincir kuralı gereği, son katmandan ilk katmana doğru hesaplandığı için, çıktı parametrelerine göre türevin, klasik ANFIS yapısından farkı yoktur.

$$\frac{\partial E}{\partial q_l} = \frac{\partial E}{\partial y}\overline{\delta_l} \tag{3.15}$$

$$\frac{\partial E}{\partial p_{il}} = \frac{\partial E}{\partial y}\overline{\delta_l}x_i \tag{3.16}$$

(3.17) ve (3.18)'de girdi parametrelerine göre gradyanlar bulunmaktadır.

$$\frac{\partial E}{\partial \mu_{i_k}} = \frac{\partial E}{\partial y}\frac{\partial y}{\partial A_{i_k}(x_k)}\left(\frac{x_k - \mu_{i_k}}{\sigma_{i_k}}\right)\exp\left(-\frac{1}{2}\left(\frac{x_k - \mu_{i_k}}{\sigma_{i_k}}\right)^2\right) \tag{3.17}$$

$$\frac{\partial E}{\partial \sigma_{i_k}} = \frac{\partial E}{\partial y}\frac{\partial y}{\partial A_{i_k}(x_k)}\frac{(x_k - \mu_{i_k})^2}{{\sigma_k}^3}\exp\left(-\frac{1}{2}\left(\frac{x_k - \mu_{i_k}}{\sigma_{i_k}}\right)^2\right) \tag{3.18}$$

Parametre indisleri $i = 1, \dots, n$; $l = 1, \dots, m$; $k = 1, \dots, n$; $i_k = 1,2, \dots, j_k$ şeklindedir.

$$\frac{\partial E}{\partial \lambda} = \frac{\partial E}{\partial y}\frac{\partial y}{\partial f_l}\frac{\partial f_l}{\partial \overline{\delta_l}}\frac{\partial \overline{\delta_l}}{\partial \delta_l}\frac{\partial \delta_l}{\partial \lambda} \tag{3.19}$$

$$\frac{\partial y}{\partial f_l}\frac{\partial f_l}{\partial \overline{\delta_l}}\frac{\partial \overline{\delta_l}}{\partial \delta_l} = g_l \frac{(\sum_{i=1}^{m} \delta_i) - \delta_l}{(\sum_{i=1}^{m} \delta_i)^2} \tag{3.20}$$

$$\frac{\partial \delta_l}{\partial \lambda} = \frac{-\chi^{\,n}(A_n)R_n}{Q_n^2} \tag{3.21}$$

$$\frac{\partial E}{\partial \lambda} = \frac{\partial E}{\partial y} g_l \frac{(\sum_{i=1}^{m} \delta_i) - \delta_l}{(\sum_{i=1}^{m} \delta_i)^2} \left(\frac{-\chi^{\,n}(A_n)R_n}{Q_n^2}\right) \tag{3.22}$$

(3.21)'de hatanın λ parametresine göre türevi yer almaktadır ve n girdi sayısını belirtmektedir. Ayrıca χ fonksiyonu ile R_n ve Q_n değerlerine, Bölüm 2.1.1'de değinilmiştir. Genelliliği bozmamak adına, verilen yapıda her bir girdi için farklı sayıda üyelik fonksiyonu olabileceğinden, (3.21) aşağıdaki şekilde de ifade edilebilir. Diğer yapılarda benzer şeklide genelleştirilebilir.

$$\frac{\partial \delta_l}{\partial \lambda} = \frac{-\chi^{\,i_k}\left(A_{i_k}(x_k)\right)R_{i_k}}{Q_{i_k}^2} \tag{3.23}$$

Burada x_k, k. girdiyi temsil etmektedir. Üyelik fonksiyonlarına göre türev hesaplanırken, (3.22)'de yer alan gradyan hesabından yararlanılmaktadır.

$$\frac{\partial y}{\partial A_{i_k}(x_k)} = \frac{\partial y}{\partial f_l}\frac{\partial f_l}{\partial \overline{\delta_l}}\frac{\partial \overline{\delta_l}}{\partial \delta_l}\frac{\partial \delta_l}{\partial A_{i_k}(x_k)} \tag{3.24}$$

$$\frac{\partial y}{\partial A_{i_k}(x_k)} = g_l \frac{(\sum_{i=1}^{m} \delta_i) - \delta_l}{(\sum_{i=1}^{m} \delta_i)^2} S_{n,k} \tag{3.25}$$

$$S_{n,k} = \frac{(\chi^{n}(A_n)/a_k)Q_n - (\sum_{j=1}^{n-1} \lambda^{n-j-1}(1-\lambda)^j \chi^{j-1}(A_n \backslash a_k) - \sum_{i=1}^{n-1}(1-\lambda)^i(\chi^{n}(A_n)/a_k))\chi^{n}(A_n)}{Q_n^2} \quad \textbf{(3.26)}$$

$$S_n = \frac{1}{1+\frac{1}{\lambda}(\prod_{k=1}^{n}\left(1+\lambda\frac{1-A_{i_k}(x_k)}{A_{i_k}(x_k)}\right)-1)} \quad \textbf{(3.27)}$$

(3.26)'da $\frac{\partial \delta_l}{\partial A_{i_k}(x_k)}$ değeri, sadelik açısından $S_{n,k}$'ya eşitlenmiştir. (3.27)'de PH'ın çarpım sembolü ile ifadesi yer almaktadır. Daha önceden de bahsedildiği gibi, gradyan işlemlerinde toplam sembolü kullanılırken, genel gösterimlerde çarpım sembolü yer alan gösterim kullanılmaktadır. Eğer, bütün katmanlardaki değerler, iç içe yazılıp, 1. girdi için, temsili bir gradyan hesaplanmak istenirse,

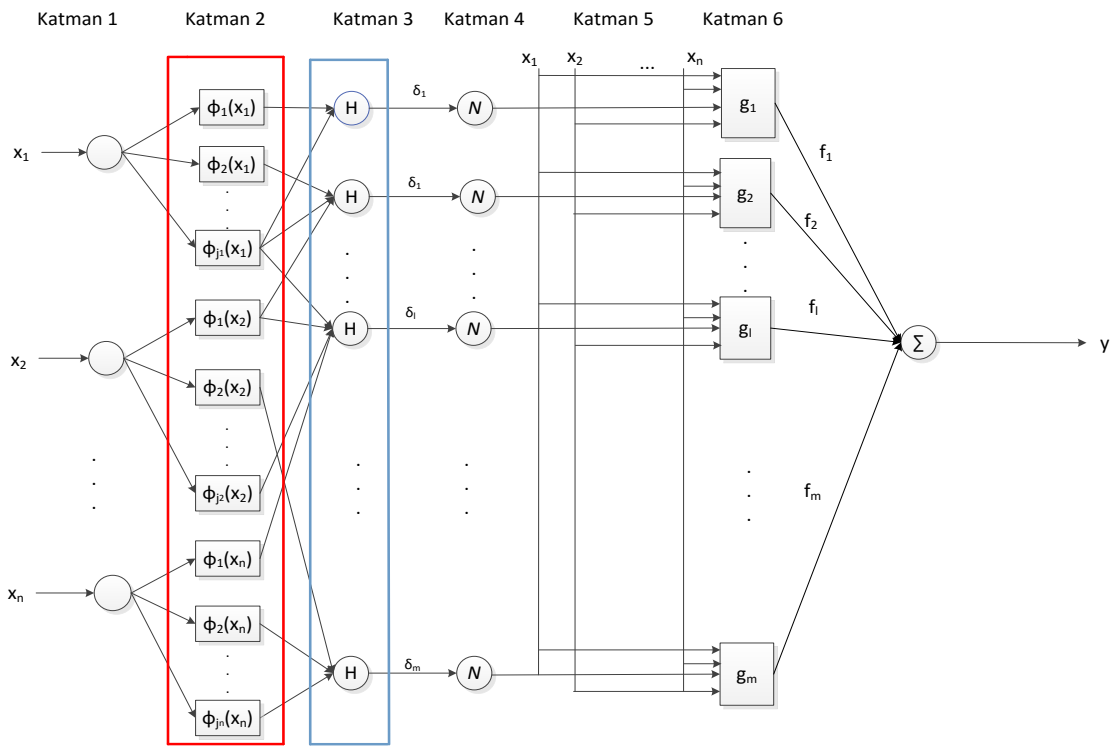
$$\frac{\partial y}{\partial A_{i_1}(x_1)} = \frac{\sum_{i_2=1}^{j_2}\sum_{i_3=1}^{j_3}\cdots\sum_{i_3=n}^{j_n} g_{i_1,i_2,\ldots,i_n} S_{n,k=i_1} - y\sum_{i_2=1}^{j_2}\sum_{i_3=1}^{j_3}\cdots\sum_{i_3=n}^{j_n} S_{n,k=i_1}}{\sum_{i_1=1}^{j_1}\sum_{i_2=1}^{j_2}\sum_{i_3=1}^{j_3}\cdots\sum_{i_n=1}^{j_n} S_n} \quad \textbf{(3.28)}$$

$i_1 = 1,2,\ldots,j_1$

$S_{n,k}$'da a_{i_1}. üyelik fonksiyonuna göre türev alınacağı için $k = i_1$ durumu göz önüne alınmıştır. Çıkışın üyelik fonksiyonlarına göre türevi (3.27) ve (3.28)'de olmak üzere, iki farklı şekilde de hesaplanmıştır. (3.28)'de yer alan gradyan, diğer girdi değerleri için de benzer şeklide hesaplanmaktadır. Ayrıca Einstein T-norm, PH'da $\lambda = 2$ için sağlandığından, Einstein T-norm'un bütün gradyanları, $\lambda = 2$ için benzer şekilde verilebilir.

3.2 AWN-Einstein ve AWN-PH Yapıları

Bölüm 2.3'de Adaptif Dalgacık Ağlarından bahsedilmişti. Adaptif Dalgacık Ağlarının, ANFIS yapısından farkı, ikinci katmanda Gaussian üyelik fonksiyonu yerine, dalgacık fonksiyonunun kullanılmasıdır. Ayrıca AWN-Einstein ve AWN-PH yapılarının, ANFIS-PH gibi 3. katmanında, PH T-normu kullanılmaktadır.



Şekil 3.2 *AWN-Parametrik Hamacher yapısı (mavi çizgili kısım Parametrik-Hamacher, kırmızı çizgili kısım dalgacık fonksiyonu)*

Şekil 3.2'de çok girdili ve tek çıktılı bir AWN-PH yapısı verilmiştir. Sistemin klasik ANFIS'den farkı, 2.ve 3. katmanlarında farklı fonksiyonların kullanılmasıdır. Yapılar, Bölüm 3.1'de yapıldığı gibi katmanlar şeklinde açıklanmıştır. AWN yapısında μ ve σ parametrelerinden farklı olarak, ancak aynı anlama gelen b ve c parametreleri kullanılmaktadır. Bu parametreler sırsıyla öteleme (translation) ve dönme (dilation) parametreleridir.

- **Katman 1**: Dış dünyadan alınan tam değerler (crisp values) direk olarak 2. katmana aktarılmaktadır.

- **Katman 2**: Bulanıklaştırma katmanıdır. Bu katmanda Gauss fonksiyonunun türevi olan, dalgacık fonksiyonu kullanılmaktadır.

$$A_{i_k}(x_k) = (1 - \left(\frac{x_k - b_{i_k}}{c_{i_k}}\right)^2)\exp(-\frac{1}{2}\left(\frac{x_k - b_{i_k}}{c_{i_k}}\right)^2) \quad \textbf{(3.29)}$$

Burada $k = 1,2,\dots,n$ ve $i_k = 1,2,\dots,j_k$

- **Katman 3**: 2. katmandan aktarılan veriler, PH T-normu için girdi değerleri olarak alınmaktadır. Bu katmanın çıktısı, kuralın işleyiş derecesini vermektedir.

Parametrik-Hamacher-toplam ifadesi

$$\frac{\chi^{n}(A_n)}{\lambda^{n-1}+\sum_{j=1}^{n-1}\lambda^{n-j-1}(1-\lambda)^{j}\chi^{j}(A_n)-\sum_{i=1}^{n-1}(1-\lambda)^{i}\chi^{n}(A_n)} \quad \textbf{(3.30)}$$

Parametrik-Hamacher-çarpım ifadesi

$$\delta_l = \frac{1}{1+\frac{1}{\lambda}\left(\prod_{k=1}^{n}\left(1+\lambda\frac{1-A_{i_k}(x_k)}{A_{i_k}(x_k)}\right)-1\right)} \quad \textbf{(3.31)}$$

$\lambda = 2$ için Einstein T-norm elde edilmektedir.

$$\delta_l = \frac{1}{1+\frac{1}{2}\left(\prod_{k=1}^{n}\left(1+2\frac{1-A_{i_k}(x_k)}{A_{i_k}(x_k)}\right)-1\right)} \quad \textbf{(3.32)}$$

Burada $l = i_1, i_2, \dots, i_n$, $i_1 = 1, \dots, j_1$, $i_2 = 1, \dots, j_2$, $\dots$, $i_n = 1, \dots, j_n$,
n ise girdi sayısıdır.

- **Katman 4**: Normalizasyon katmanıdır. 3. katmandan elde edilen işleyiş dereceleri, normalize işlemine tabii tutulmaktadır.

$$\overline{\delta_l} = \frac{\delta_l}{\sum_{i=1}^{m}\delta_i} \qquad (l = 1, \dots, m) \qquad \textbf{(3.33)}$$

- **Katman 5**: Netleştirme katmanı olan bu katmanda, kullanılan 1. dereceden lineer denklem, girdi değerleri ile oluşturulmaktadır.

$$f_l = \overline{\delta_l} g_l \qquad (l = 1, \dots, m) \qquad \textbf{(3.34)}$$

$$g_l = q_l + \sum_{i=1}^{n} p_{il} x_i \qquad \textbf{(3.35)}$$

- **Katman 6**: Bu katmanda, gelen bütün sinyaller toplanıp, AWN-PH yapısının çıktısı elde edilmektedir.

$$y = \sum_{l=1}^{m} f_l \tag{3.36}$$

AWN-PH'da, ANFIS-PH yapısında da kullanılan öğrenme algoritması, BFGS kullanılmıştır. AWN ve ANFIS yapılarından PH kullanılması bakımından bir fark bulunmayıp, sadece 2. katmandaki üyelik fonksiyonları farklılık gösterdiği için, iki yapının da parametre sayıları birbirleri ile aynıdır. Girdi değişkenleri, 2. katmanda bulunan fonksiyonların içinde yer almaktadır. Bu yüzden, eğitim algoritması için girdi değerlerine göre gradyanlar değişirken, çıktı değerlerine ve λ'ya göre gradyanlar değişmemektedir. AWN-PH yapısı için gradyan hesaplamaları, Bölüm 3.2.1'de genel olarak hesaplanmıştır.

3.2.1 AWN-PH için parametre gradyanları

Öncelikle hata fonksiyonunun sistemin çıktısına olan gradyan hesabı düşünülürse, bu kısımda bir değişiklik olmadığından, aynı ifade kullanılacaktır.

$$\frac{\partial E}{\partial y} = \frac{2}{N} \sum_{k=1}^{N} (y - y_g) \tag{3.37}$$

Zincir kuralı gereği, gradyanlar, son katmandan, ilk katmana doğru uygulandığı için çıktı parametrelerine göre gradyanlar da aşağıdaki gibidir.

Parametre indisleri $i = 1, \dots, n$; $l = 1, \dots, m$; $k = 1, \dots, n$; $i_k = 1,2, \dots, j_k$ şeklindedir.

$$\frac{\partial E}{\partial q_l} = \frac{\partial E}{\partial y} \overline{\delta_l} \tag{3.38}$$

$$\frac{\partial E}{\partial p_{il}} = \frac{\partial E}{\partial y} \overline{\delta_l} x_i \tag{3.39}$$

b ve c parametrelerinin gradyanları aşağıdaki şekilde hesaplanır:

$$\frac{\partial E}{\partial b_{i_k}} = \frac{\partial E}{\partial y} \frac{\partial y}{\partial \varphi_{i_k}(x_k)} \left(\frac{x_k - b_{i_k}}{{c_{i_k}}^2}\right)\left(3 - \left(\frac{x_k - b_{i_k}}{c_{i_k}}\right)^2\right) \exp\left(-\frac{1}{2}\left(\frac{x_k - b_{i_k}}{c_{i_k}}\right)^2\right) \tag{3.40}$$

$$\frac{\partial E}{\partial c_{i_k}} = \frac{\partial E}{\partial y}\frac{\partial y}{\partial \varphi_{i_k}(x_k)}\frac{(x_k - b_{i_k})^2}{{c_{i_k}}^3}\left(3 - \left(\frac{x_k - b_{i_k}}{c_{i_k}}\right)^2\right)\exp\left(-\frac{1}{2}\left(\frac{x_k - b_{i_k}}{c_{i_k}}\right)^2\right) \tag{3.41}$$

$$\frac{\partial E}{\partial \lambda} = \frac{\partial E}{\partial y}\frac{\partial y}{\partial f_l}\frac{\partial f_l}{\partial \overline{\delta_l}}\frac{\partial \overline{\delta_l}}{\partial \delta_l}\frac{\partial \delta_l}{\partial \lambda} \tag{3.42}$$

λ'ya göre türev hesabı yapılırken, (3.42)'de hiçbir çarpan değişmediği için λ'ya göre türevde de değişiklik olmamaktadır. O halde

$$\frac{\partial E}{\partial \lambda} = \frac{\partial E}{\partial y} g_l \frac{(\sum_{i=1}^{m}\delta_i) - \delta_l}{(\sum_{i=1}^{m}\delta_i)^2}\left(\frac{-\chi^n(A_n)R_n}{Q_n^2}\right) \tag{3.43}$$

(3.43) eşitliği aynen korunmaktadır. İlk girdi için yapıyı genellemek gerekirse, (3.26) ve (3.27) eşitliklerinden yararlanılarak, (3.44) elde edilmektedir.

$$\frac{\partial y}{\partial \varphi_{i_1}(x_1)} = \frac{\sum_{i_2=1}^{j_2}\sum_{i_3=1}^{j_3}\cdots\sum_{i_3=n}^{j_n}\varphi_{i_1,i_2,\ldots,i_n}S_{n,k=i_1} - y\sum_{i_2=1}^{j_2}\sum_{i_3=1}^{j_3}\cdots\sum_{i_3=n}^{j_n}S_{n,k=i_1}}{\sum_{i_1=1}^{j_1}\sum_{i_2=1}^{j_2}\sum_{i_3=1}^{j_3}\cdots\sum_{i_n=1}^{j_n}S_n} \tag{3.44}$$

Tezde, bütün parametreler Bölüm 3.1.3 ve Bölüm 3.2.1'de yer alan gradyan hesaplamalarına göre güncellenmektedir. ANFIS-PH yapısında μ ve σ parametrelerinin başlangıç değerleri 1 seçilirken, diğer parametrelere ise (λ parametresi hariç) 0 ile 1 arasında rastgele değerler verilmektedir. AWN-PH yapısında λ parametresi hariç, bütün parametrelerin başlangıç değerleri 0 ile 1 arasında seçilmiştir. λ parametresi iki yapı içinde, başlangıç değerleri olarak, birbirinden farklı birçok reel sayı verilip, denenmiştir.

4. TOPLULUK ANFIS-PH ve AWN-PH YAPILARI

Topluluk yapıları oluşturulurken, her bir yapının çıktısı, integrator ile işleme girerek son tahmin değerleri elde edilir. Sistem katmanlar halinde verilmektedir.

- **Katman 1:** Veri seti gerekli önişlemlerden geçerek, her bir yapıya girdi olarak verilmektedir. Sistemin girdileri $u_{t-1}, u_{t-2}, \dots, u_{t-n}$ olacak şekilde sisteme dahil olmaktadır. Burada t girdi sayısına göre değişiklik gösteren bir sayıdır.
- **Katman 2:** 1. katmandan iletilen veriler, ANFIS-PH ya da AWN-PH yapıları için girdi teşkil etmektedir. Bu katmanda yapılar çalıştırılır ve sonuçlar bir sonraki katmana iletilir.
- **Katman 3:** Bütün sistemlerin sonuçlarının yer aldığı katmandır. AWN-PH yapısını örnek almak gerekirse, bu katmanda AWN-PH 1'in çıktısı x_1, AWN-PH 2'nin çıktısı x_2 ve AWN-Parametrik Hamacher n'in çıktısı ise x_n olarak adlandırılmaktadır. O halde 3. katman, $[x_1, x_2, \dots, x_n]$ vektöründen oluşmaktadır.
- **Katman 4:** İntegrator işlemlerinin yapıldığı katmandır. Bu katmanda iki farklı integrator işlemi kullanılmıştır. Bu işlemler ağırlıklandırılmış ortalama (eşitlik 4.1) ve ortalama işlemleridir (eşitlik 4.2). Aslında bu ifadeler, aynı zamanda Sugeno tip bulanık çıkarım sisteminde de kullanılan ifadelerdir.

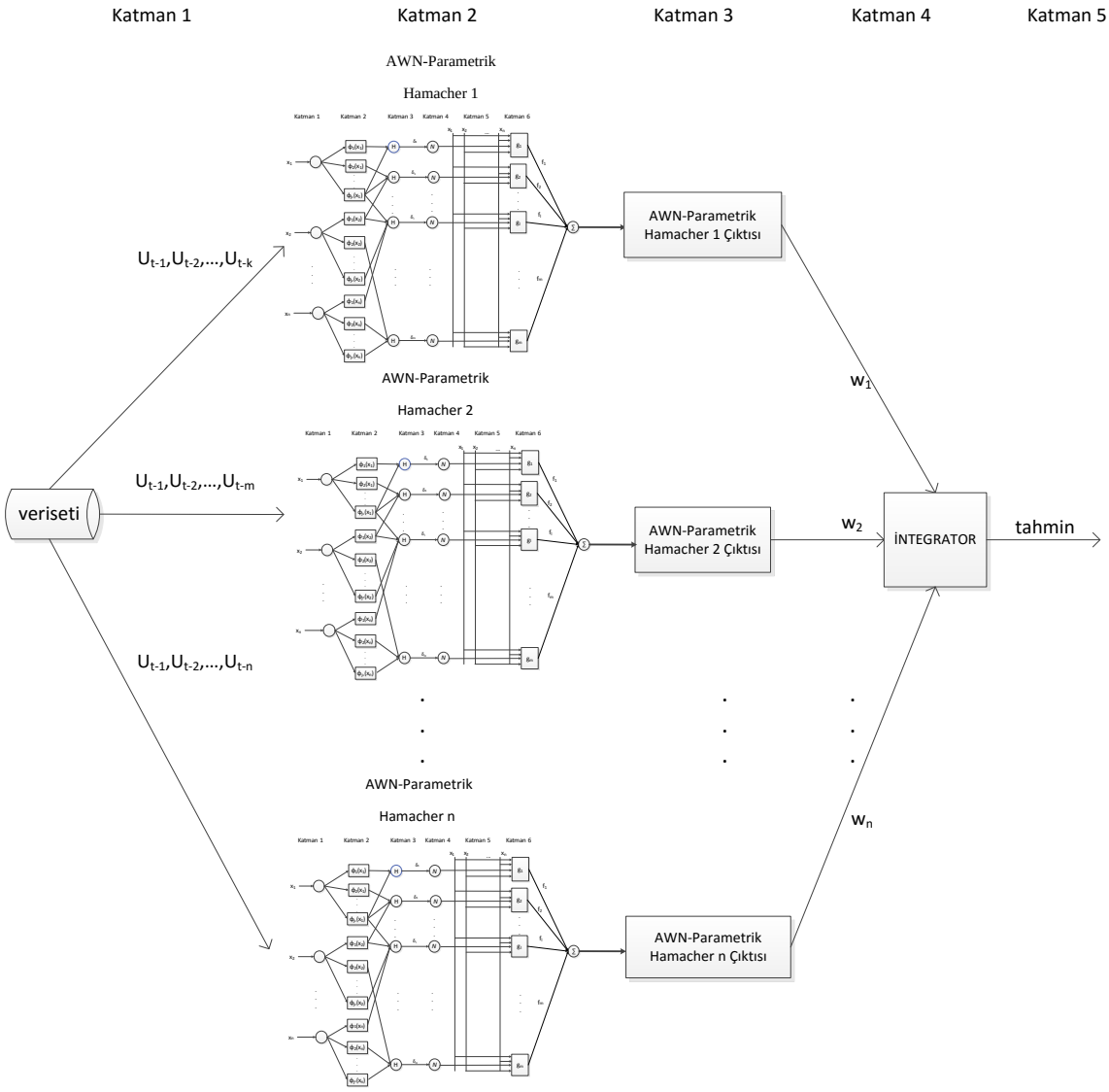
$$y = \frac{w_1x_1+w_2x_2+\cdots+w_nx_n}{w_1+w_2+\cdots+w_n} = \frac{\sum_{i=1}^{n} w_ix_i}{\sum_{i=1}^{n} w_i} \tag{4.1}$$

$$y = \frac{x_1+x_2+\cdots+x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{4.2}$$

- **Katman 5:** Çıktı katmanıdır. Bu katmanda, sistem tarafından bir çıktı üretilir. Gerçek değerler ile Topluluk yapısının ürettiği değer arasında, performans ölçütü hesaplanır.

Örnek olarak AWN-PH yapısının Topluluk sistemi Şekil 4.1'de verilmektedir. 2. katmanda ANFIS-PH yapısı da kullanılmaktadır. 3. katmandan elde edilen $[x_1, x_2, \dots, x_n]$ vektörünün uygun çarpanlar ile çarpılması sistemin performans ölçütünün hesaplanması

açısından önemli yer tutmaktadır. Bu yüzden sistemde w_i'ler öğrenme algoritması ile uygun değerler elde edilerek, integrator işleminde geçmektedir.



Şekil 4.1 *Topluluk AWN-PH yapısı*

Şekil 4.1'de yer alan sistemin tahmini çıktısına y_t ve gerçek çıktıya ise y_g ifadeleri kullanılacak olursa, kullanılan eğitim algoritması ortalama kareler hata temel alınarak eğitilmiştir.

$$\text{E=OKH} = \frac{1}{N}\sum_{k=1}^{N}(y_t - y_g)^2 \qquad \textbf{(4.3)}$$

Ayrıca $[x_1, x_2, \dots, x_n]$ vektörünün yapısında, $1 \leq i \leq n$ olacak şekilde her bir x_i değeri farklı bir sistemin çıktısı olduğundan, kendi aralarında birer vektörü temsil etmektedirler. Şekil 4.1'de yer alan sistemde türev tabanlı öğrenme algoritması kullanıldığı için, uygun w_i değerlerinin bulunmasında w_i değişkenlerine göre gerekli gradyanların bulunması gerekmektedir. Bu gradyanlar (4.4) ve (4.5)'de verilmektedir.

$$\frac{\partial E}{\partial y}\frac{\partial y}{w_i} = \frac{\partial E}{\partial y}\frac{x_i \sum_{i=1}^{n} w_i - \sum_{i=1}^{n} w_i x_i}{(\sum_{i=1}^{n} w_i)^2} \tag{4.4}$$

$$\frac{\partial E}{\partial y} = \frac{2}{N}\sum_{k=1}^{N}(y_t - y_g) \tag{4.5}$$

Hesaplanan gradyanlar ile hatayı minimize eden en uygun w_i katsayıları bulunmaktadır. Daha sonra ağırlıklandırılmış ortalama yapılarak, sistemin çıktısı elde edilmektedir.

5. DENEYSEL ÇALIŞMALAR

Tezin bu bölümünde sunulan yapıların farklı veri setlerinde değerlendirmeleri yapılmaktadır. ANFIS-PH ve AWN-PH yapıları kaotik zaman serileri ve sistem tanıma problemleri üzerinde uygulanmıştır. Elde edilen sonuçlar ise farklı çalışmalar ile karşılaştırılmaktadır. Ens-ANFIS-PH ve Ens-AWN-PH yapıları ise finansal zaman serileri üzerinde denenmiş ve sonuçlar sunulmuştur. Tezde, kaotik zaman serileri olarak Box-Jenkins ve Mackey-Glass zaman serileri kullanılmıştır. Bu iki zaman serisi Benchmark problemleri olarak bilinmektedir. Ayrıca diğer bir veri seti ise sistem tanıma problemlerinden seçilmiştir. Tezde kullanılan finansal zaman serileri ise Dow-Jones, S&P 500 ve EUR/USD zaman serileridir.

Bu bölümün özeti şu şekildedir; Bölüm 5.1'de, kullanılan veri setleri tanıtılmıştır. Bölüm 5.2'de, verisetlerinin sunulan yapılar üzerinde uygulanma biçimi verilmektedir. Bu bölüm deney metodu olarak geçmektedir. Bölüm 5.3'de ANFIS-PH ve AWN-PH yapılarının deney sonuçları verilmektedir. Bölüm 5.4'de Ens-ANFIS-PH ve Ens-AWN-PH yapılarının deney sonuçları verilmektedir.

5.1 ANFIS/AWN-PH Deneyleri

Bu bölümde ANFIS-PH ve AWN-PH yapıları ayrı ayrı Box-Jenkins, Mackey-Glass ve Sistem tanılama verileri üzerinde denenmiş ve her bir yapı için farklı λ değerleri ile HKOK sonuçları sunulmuştur. Ayrıca sunulan yapılar literatürde yer alan diğer çalışmalar ile karşılaştırılmıştır.

5.1.1 Box-Jenkins zaman serisi tahmini

Box-Jenkins zaman serisi, 1994 yılında Box ve ark. [61] tarafından iki değişkenli bir zaman serisinin analizini göstermek için kullanılmıştır. Gaz ocağı içerisindeki, hava ve metan gazları, CO_2 içeren karışımı elde etmek için birleştirilmiştir. Bu süreçte u_t girdileri, gazın akış miktarını belirler iken, y_t çıktısı ise CO_2 yoğunluğunu belirlemektedir. y(t), verisinin tahmini için, $u(t-4)$ ve $y(t-1)$ verileri kullanılmıştır.

Veri seti 292 veriden oluşmaktadır. İlk 200 veri, yapıyı eğitmek için kullanılırken, kalan 92 veri ise test verisi olarak kullanılmaktadır. Modelde her bir girdi için 2'şer üyelik

fonksiyonu kullanılıp, 4 kuraldan oluşmaktadır. ANFIS-PH ve AWN-PH yapılarının her ikiside 400 iterasyonda eğitilmişlerdir. Ayrıca her iki yapıda da λ da dahil olmak üzere 21 parametre kullanılmıştır. Sistemin performansını ölçmek için Hata Kareler Ortalamasının Karekökü (HKOK) kullanılmıştır.

$$\text{HKOK}=\sqrt{\frac{1}{N}\sum_{k=1}^{N}(y-y_g)^2} \tag{5.1}$$

Tablo 5.1 ve 5.2'de sırasıyla ANFIS-PH ve AWN-PH yapılarının farklı λ değerleri için oluşan farklı T-normlar ile sonuçları verilmiştir.

Tablo 5.1. *Box-Jenkins zaman serisinin λ değerleri için ANFIS-PH sonuçları*

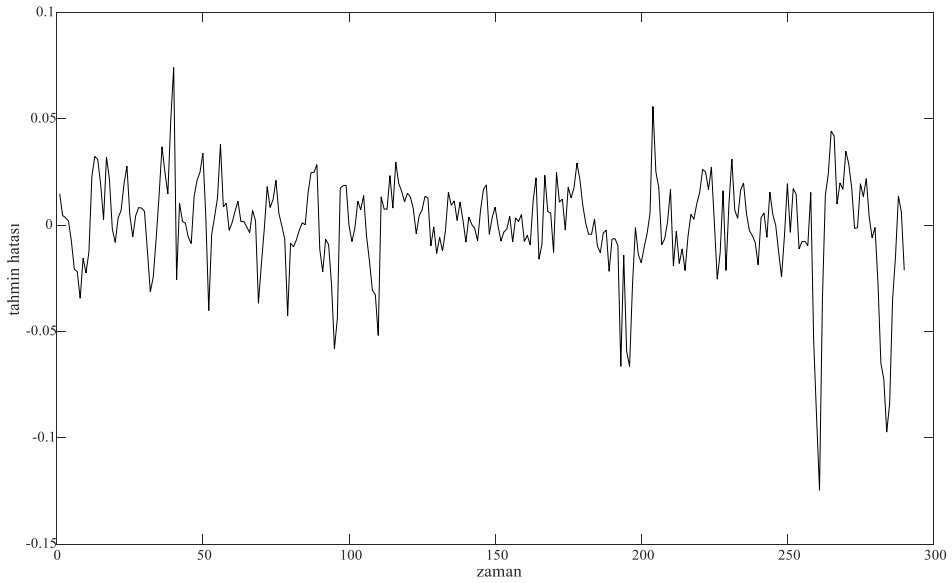
λ **değerleri**	**HKOK-Eğitim**	**HKOK-Test**
$\lambda = 1$ (çarpım)	0,0238	0,0315
$\lambda = 0$ (Klasik Hamacher)	0,0234	0,0309
$\lambda = 2$ (Einstein)	0,0238	0,0313
$\lambda = 0{,}035$	0,0231	0,0299

Tablo 5.1'de görüldüğü gibi en iyi sonucu veren λ=0,035 iken, Tablo 5.2'de ise en iyi λ değerinin sıfıra yakınsadığı görülmüş ve en iyi test sonucu klasik Hamacher' in verdiği görülmektedir. Tablolarda yer alan λ değerleri başlangıç parametrelerinin birbirinden farklı değerleri için denenmiştir. Diğer λ değerleri ise örnek olarak verilmiştir.

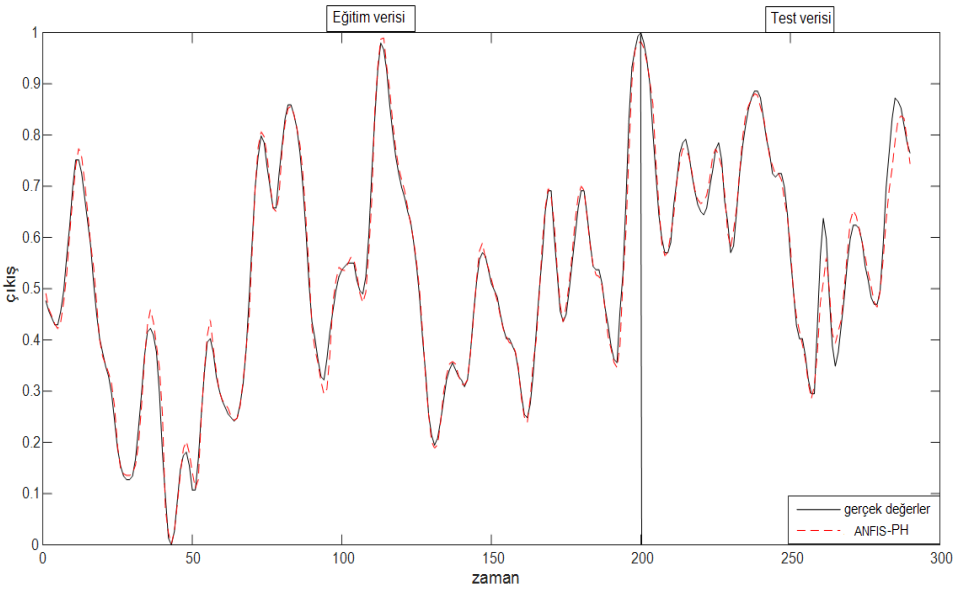
Tablo 5.2. *Box-Jenkins zaman serisinin λ değerleri için AWN-PH sonuçları*

λ **değerleri**	**HKOK-Eğitim**	**HKOK-Test**
$\lambda = 1$ (çarpım)	0,0193	0,0317
$\lambda = 0$ (Klasik Hamacher)	0,0237	0,0301
$\lambda = 2$ (Einstein)	0,0227	0,0349
$\lambda = 1{,}2671$	0,0237	0,0361

İki tabloda dikkate alındığında, en iyi sonucu ANFIS-PH ($\lambda = 0{,}035$) verdiği görülmektedir. Şekil 5.1'de tahmin hatası verilirken, Şekil 5.2'de tahmini ve gerçek değerleri arasındaki ilişki verilmiştir.



Şekil 5.1. *Box-Jenkins zaman serisi için ANFIS-PH (λ=0,035) tahmin hatası*



Şekil 5.2 *Box-Jenkins zaman serisi için ANFIS-PH (λ=0,035) tahmini ve gerçek değerleri*

Tablo 5.3. *Box-Jenkins zaman serisi için modellerin test sonuçlarının karşılaştırılması*

Model	HKOK
Tong's model[62]	0,685
Pedrycz's model[63]	0,566
Xu's model[64]	0,573
Sugeno's model[65]	0,596
Surmann's model[66]	0,400
FuNN model[67]	0,071
HyFIS model[36]	0,042
WNN[68]+gradient	0,084
WNN[68]+hybrid	0,081
LLWNN[68]+gradient	0,017
LLWNN[68]+hybrid	0,013
AWN-Z [59] (2 MFs)	0,033
AWN-F [59] (2 MFs)	0,031
Sunulan Model (AWN-PH) (2 MFs)	0,0301
Sunulan Model (ANFIS-PH) (2 MFs)	0,0299

Tezde sunulan ANFIS-PH ve AWN-PH yapılarının, çıktı kısmının, birinci dereceden denklemlerden oluşturulduğuna yer verilmişti. Tablo 5.3'e bakıldığında ise Box-Jenkins zaman serisi için, PH T-normun AWN yapısına katkıda bulunduğu söylenilebilir. Ancak görüldüğü gibi Box-Jenkins zaman serisi için PH T-normun ANFIS yapısına olan etkisinin AWN yapısına olan etkisinden daha fazla olduğu görülebilir. Diğer çalışmalar ile kıyaslandığında ise en iyi sonuca ulaşılmasa dahi "LLWNN" modelinden sonra en iyi sonuca ulaşıldığı görülmektedir. LLWNN modelde 56 parametre kullanılırken, önerilen modellerde bu sayı daha azdır.

5.1.2 Mackey-Glass Zaman Serisi Tahmini

Bu bölümde seçkin tahmin problemlerinden biri olan ve Bulanık-Sinir ağları sistemlerinde sıkça kullanılan Mackey-Glass zaman serisi kullanılmıştır. Zaman serisi verileri 5.2'de verilen diferansiyel denklem ile üretilmektedir.

$$\frac{dx}{dt} = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} \tag{5.2}$$

Kaotik zaman serisi oluşması için $\tau \geq 17$ seçilmesi gerekmektedir. Bu kısımda başlangıç değer koşulları $\tau = 17$ ve $x(0) = 1{,}2$ olarak seçilmiştir. Deneyde, 5.2 diferansiyel denklemi ile oluşturulan, $t = 118$'den 1117'ye kadar olan, 1000 veri kullanılmıştır. Girdi değerleri olarak $x(t), x(t-6), x(t-12)$ ve $x(t-18)$ kullanılmış ve $x(t+6)$ tahmini yapılmıştır. 1000 veri içerisinden ilk 500 veri sunulan modelin eğitimi için ayrılırken, kalan 500 veri, test verisi olarak kullanılmıştır. Sunulan Model 4000 iterasyonda eğitilmiştir. 4 girdiye sahip olan Mackey-Glass zaman serisinde, her bir girdi için ikişer üyelik fonksiyonu kullanılmıştır. Sistemde toplam 16 kural ve 97 parametre bulunmaktadır. Tablo 5.4 ve 5.5'de belirli λ değerleri için HKOK sonuçları verilmiştir.

Tablo 5.4. *Mackey-Glass zaman serisinin λ değerleri için ANFIS-PH sonuçları*

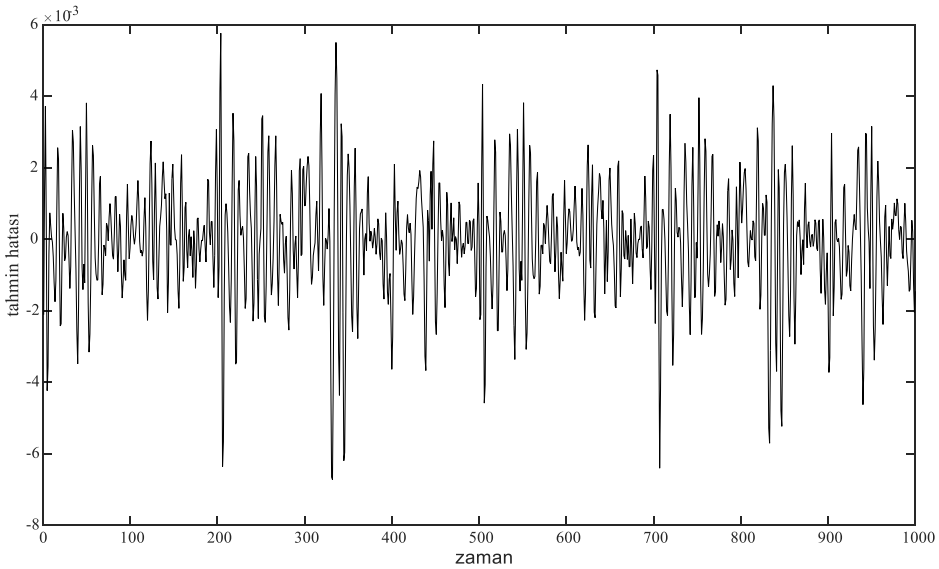
λ değerleri	HKOK-Eğitim	HKOK-Test
$\lambda = 1$ (çarpım)	0,00224	0,00225
$\lambda = 0$ (Klasik Hamacher)	0,00264	0,00256
$\lambda = 2$ (Einstein)	0,00221	0,00213
$\lambda = 0{,}0153$	0,00174	0,00168

ANFIS-PH yapısı için en iyi sonucu $\lambda = 0{,}0153$ için verirken, AWN-PH yapısı için ise en iyi sonucu Einstein T-norm'un verdiği görülmüştür. Öğrenme algoritması ile en iyi λ değerinin 2'ye yakınsadığı görülmüştür. İki tablo birbirleri ile karşılaştırıldığından Mackey-Glass zaman serisi için en iyi sonucu $\lambda = 2$ için verdiği görülmüştür.

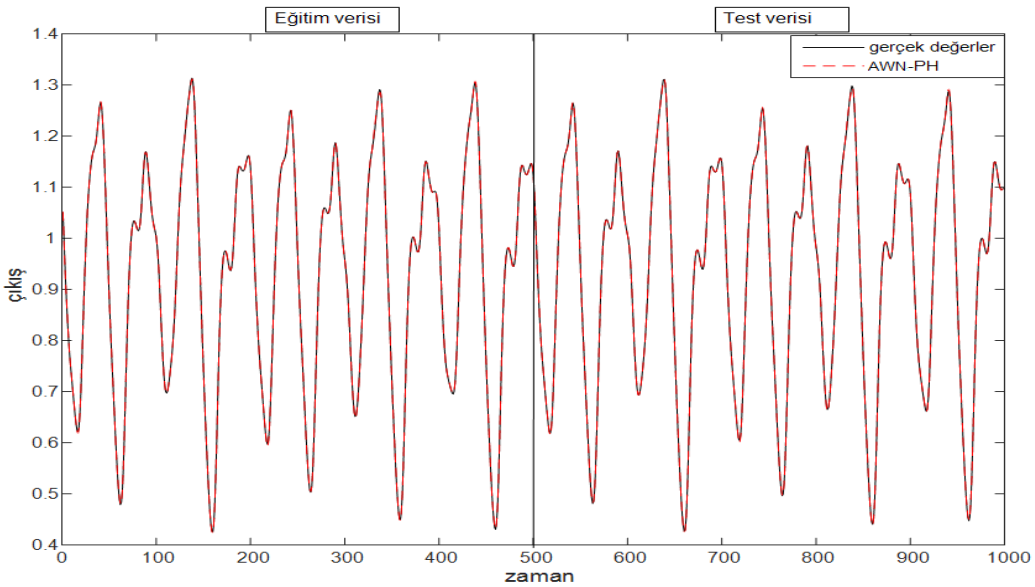
Tablo 5.5. *Mackey-Glass zaman serisinin λ değerleri için AWN-PH sonuçları*

λ değerleri	HKOK-Eğitim	HKOK-Test
$\lambda = 1$ (çarpım)	0,00183	0,00178
$\lambda = 0$ (Klasik Hamacher)	0,00332	0,00327
$\lambda = 2$ (Einstein)	0,00162	0,00156
$\lambda = 1{,}7576$	0,00171	0,00161

Mackey-Glass zaman serisi için Şekil 5.4'de tahmin hatası ve Şekil 5.5'de ise tahmini ve gerçek değerleri arasındaki ilişki verilmiştir.



Şekil 5.3. *Mackey-Glass zaman serisi için AWN-PH (λ=2) tahmin hatası*



Şekil 5.4. *Mackey-Glass serisi için AWN-PH (λ=2) tahmini ve gerçek değerleri*

Tablo 5.6. *Mackey Glass zaman serisi için modellerin test sonuçlarının karşılaştırılması*

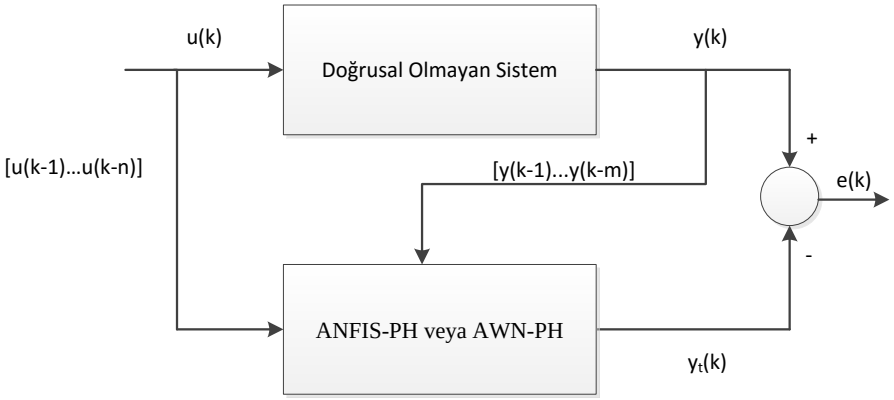
Model	HKOK
Auto-regressive model	0,19
Cascade correlation NN	0,06
Backpropagation NN	0,02
Sixth-order polynomial	0,04
Linear prediction method	0,55
Classical RBF (23 nöron) [69]	0,0114
PG-RBF network [70]	0,0028
Genetic algorithm and fuzzy system [71]	0,049
WNN [68] + gradient	0,0071
WNN [68]+ hybrid	0,0059
LLWNN [68] + gradient	0,0041
LLWNN [68] + hybrid	0,0036
AWN-Z [59]	0,00982
AWN-F [59]	0,00178
Sunulan Model (ANFIS-PH) (2MFs)	0,00168
Sunulan Model (AWN-PH) (2MFs)	0,00156

Tablo 5.6'da, sunulan modellerin, diğer modeller arasındaki yeri gösterilmektedir. AWN-PH ($\lambda = 2$) yapısının en iyi yapı olduğu görülmektedir. Bunu nedeni ise dalgacık üyelik fonksiyonlarının hızlı geçiş durumlarını daha iyi yakalamasıdır.

5.1.3 Sistem tanımlama örneği

Şekil 5.5'de ANFIS-PH veya AWN-PH yapıları ile seri-paralel sistem tanılamanın yapısı verilmiştir. $u(k)$ kontrol sinyali, modelin girdi sinyalleri, $y(k)$ doğrusal olmayan sistemin çıktısıdır. Sistemde hedef çıktı $y(k)$ iken, tahmini çıktı ise $y_t(k)$ çıktısıdır. Sistemde kullanılan yapıların bütün parametreleri $e(k)$ hata tahminine göre güncellenmektedir. Örnekte yer alan doğrusal olmayan sistem (5.3)'de verilmiştir.

$$y(k) = 0.72y(k-1) + 0.025y(k-2)u(k-2) + 0.01u^2(k-3) + 0.02u(k-4) \quad \textbf{(5.3)}$$



Şekil 5.5. *ANFIS-PH veya AWN-PH ile Sistem Tanılama Modeli*

$y(k)$ çıktısının hesaplanması için çıktının 1, 2 ve 3 adım önceki verisinden yararlanılırken, girdinin ise 2, 3 ve 4 adım önceki verisinden yararlanılmıştır. Sistemde $y(k+1)$'i tahmin etmek için $u(k-1)$ ve $y(k)$ verileri kullanılmaktadır. Modelde her girdi için 2'şer üyelik fonksiyonu seçilmiştir. Modelin eğitimi için 900 veri kullanılmıştır. Kullanılan veriler [72] ve [73] yapılarında kullanılan veriler ile benzerdir. Eğitim girdilerinin yarısı için [-2,2] arasındaki bağımsız ve benzerlik dağılımdan elde edilen rastgele veriler kullanılırken, kalan yarısı için ise $1.05\sin(\frac{\pi k}{45})$, sinüs dalgasından elde edilen veriler kullanılmıştır. Bütün veriler eğitildikten sonra, test sonuçlarını, diğer çalışmalar ile karşılaştırmak amaçlı, diğerleri ile aynı test sinyali kullanılmıştır. Test için 1000 veri kullanılmıştır. Test veri sinyalleri parçalı fonksiyon şeklinde (5.4)'de ifade edilmektedir.

$$u(k)=\begin{cases}\sin(\pi k/25) & k<250\\ 1.0 & 250\le k<500\\ -1.0 & 500\le k<750\\ 0.3\sin(\pi k/25)+0.1\sin(\pi k/32)+0.6\sin(\pi k/10) & 750\le k<1000\end{cases} \quad \textbf{(5.4)}$$

Tablo 5.7. *Sistem tanılama probleminin λ değerleri için ANFIS-PH sonuçları*

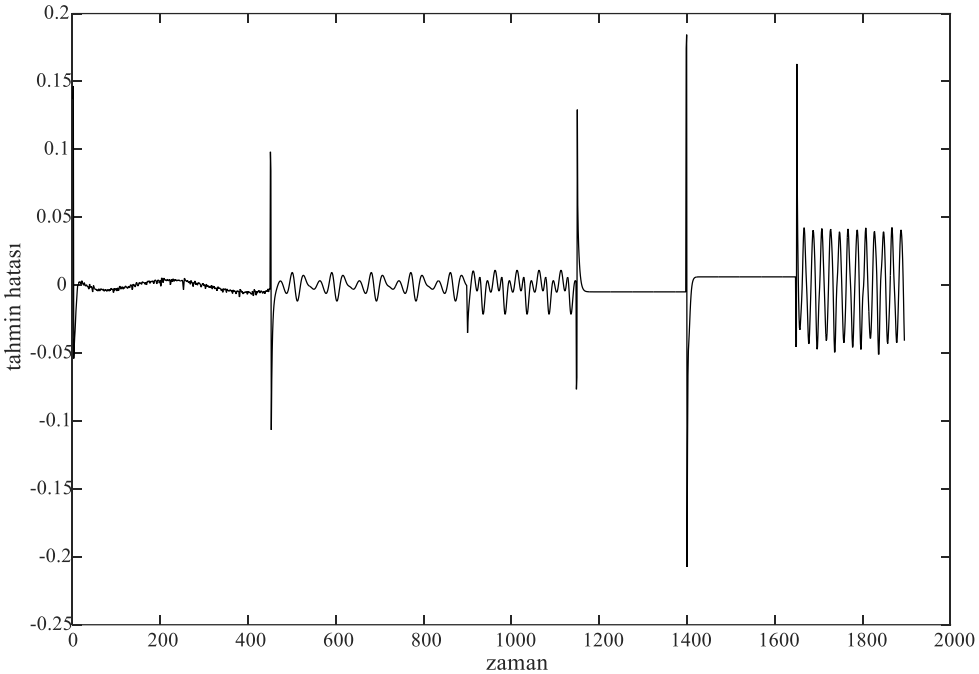
λ **değerleri**	**HKOK-Eğitim**	**HKOK-Test**
$\lambda = 1$ (çarpım)	0,0097	0,0228
$\lambda = 0$ (Klasik Hamacher)	0,0092	0,0251
$\lambda = 2$ (Einstein)	0,0092	0,0244
$\lambda = 0.0153$	0,0088	0,0226

Tablo 5.7 ve 5.8'e bakıldığında AWN-PH yapısı için $\lambda = 1{,}0806$'da en iyi sonuca ulaştığı görülmektedir. ANFIS-PH yapısı için ise en iyi sonucu $\lambda = 0{,}0153$ değerinde verdiği görülmektedir. İki tablo dikkate alındığında en iyi sonucu AWN-PH ($\lambda = 1{,}0806$) verdiği görülmektedir.

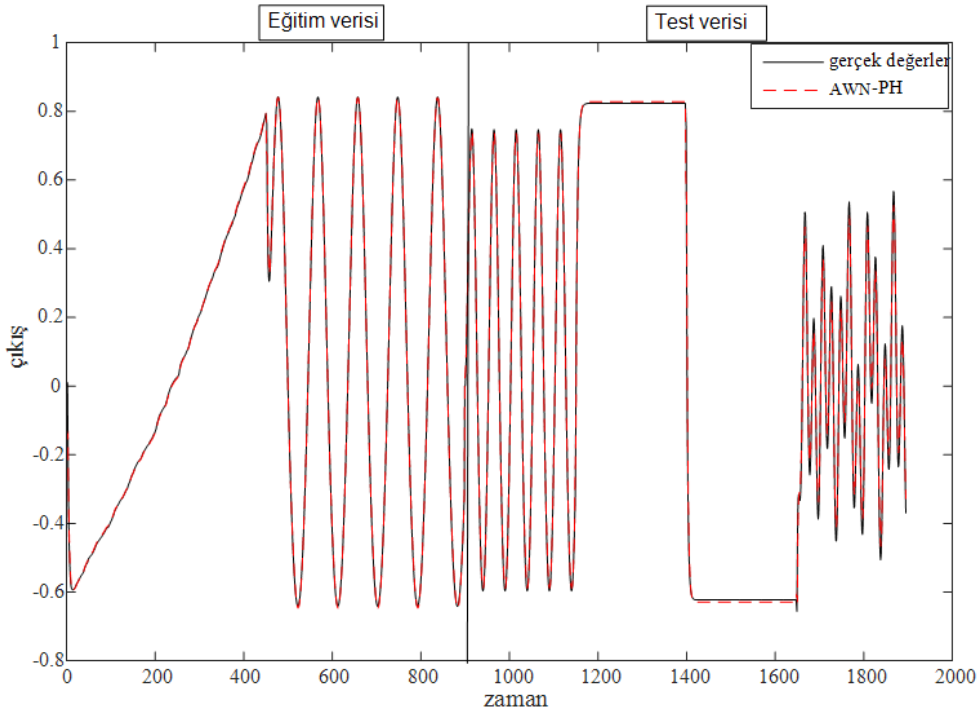
Tablo 5.8. *Sistem tanılama probleminin λ değerleri için AWN-PH sonuçları*

λ **değerleri**	**HKOK-Eğitim**	**HKOK-Test**
$\lambda = 1$ (çarpım)	0,0093	0,0232
$\lambda = 0$ (Klasik Hamacher)	0,0096	0,0264
$\lambda = 2$ (Einstein)	0,0083	0,0301
$\lambda = 1{,}0806$	0,0107	0,0209

Şekil 5.6'da en iyi modelin tahmin hatası verilmiştir. Verilen şekilde eğitim ve test hataları birlikte verilmektedir. Şekil 5.7'de ise sistem tanılama örneği için tahmini ve gerçek değerler bir arada verilmektedir. Her iki şekilde de ilk 900 veri eğitim verisi iken, kalan 1000 veri ise test verisi olarak alınmıştır. Sunulan modellerin her biri 100 iterasyonda eğitilmiştir.



Şekil 5.6. *Sistem tanılama için AWN-PH (λ=1,0806) tahmin hatası*



Şekil 5.7. *Sistem tanılama için AWN-PH (λ=1,0806) tahmini ve gerçek değerler*

Tablo 5.9. *Sistem tanıma problemi için modellerin test sonuçlarının karşılaştırılması*

Model	Parametre sayısı	HKOK Eğitim	HKOK Test
ERNN[74]	54	0,036	0,078
RSONFIN[72]	49	0,03	0,06
TRFN-S[73]	33	0,0067	0,0313
FWNN [75]	27	0,019736	0,022609
FWNN [75]	43	0,018713	0,020169
AWN-Z [59]	12	0,009368	0,022933
AWN-F [59]	20	0,009391	0,023259
Sunulan Model (ANFIS-PH) (2MFs)	21	0,008823	0,022612
Sunulan Model (AWN-PH) (2MFs)	21	0,010733	0,020883

Tablo 5.9'da, karşılaştırılan bütün yapıların eğitim ve test HKOK sonuçları dikkate alındığında, eğitim hata sonucuna göre test sonucu en iyi olan değerin, tezde sunulan model (AWN-PH) olduğu görülmektedir. Ayrıca parametre sayılarına bakıldığında, AWN-PH'ın diğer çalışmalardan daha az parametre ile daha iyi sonuç verdiği görülebilir.

5.2 Ens-ANFIS-PH ve Ens-AWN-PH Deneyleri

Tezin bu kısmında Ens-ANFIS-PH ve Ens-AWN-PH yapıları, finansal zaman serileri üzerinde denenmiştir. Finansal zaman serilerinde günlük değişmeler fazla olduğu için tahmin hatasının gerçek verilere oranı finansal veri tiplerine daha uygundur. Böylece kötü tahminde bulunulan verinin hatası gerçek veri üzerinden hesaplanabilmektedir. Hata sonuçları ise Ortalama Mutlak Yüzde Hata (OMYH) ile verilmiştir. Ortalama Mutlak Yüzde Hata (5.5)'de verilmektedir.

$$\frac{100}{N}\sum_{i=1}^{N}\left|\frac{x_i^g - x_i^t}{x_i^g}\right| \quad \text{g=gerçek veriler, t=tahmini veriler} \qquad \textbf{(5.5)}$$

Euro-Dolar veri seti hariç, diğer veri setleri (Dow-Jones, S&P 500) [76-78]'e benzer şekilde, normalizasyon işlemlerine tabii tutularak değerler [0,1] aralığına indirgenmiştir. Euro-Dolar veri seti ise belli aralıklarda değiştiği için normalizasyon

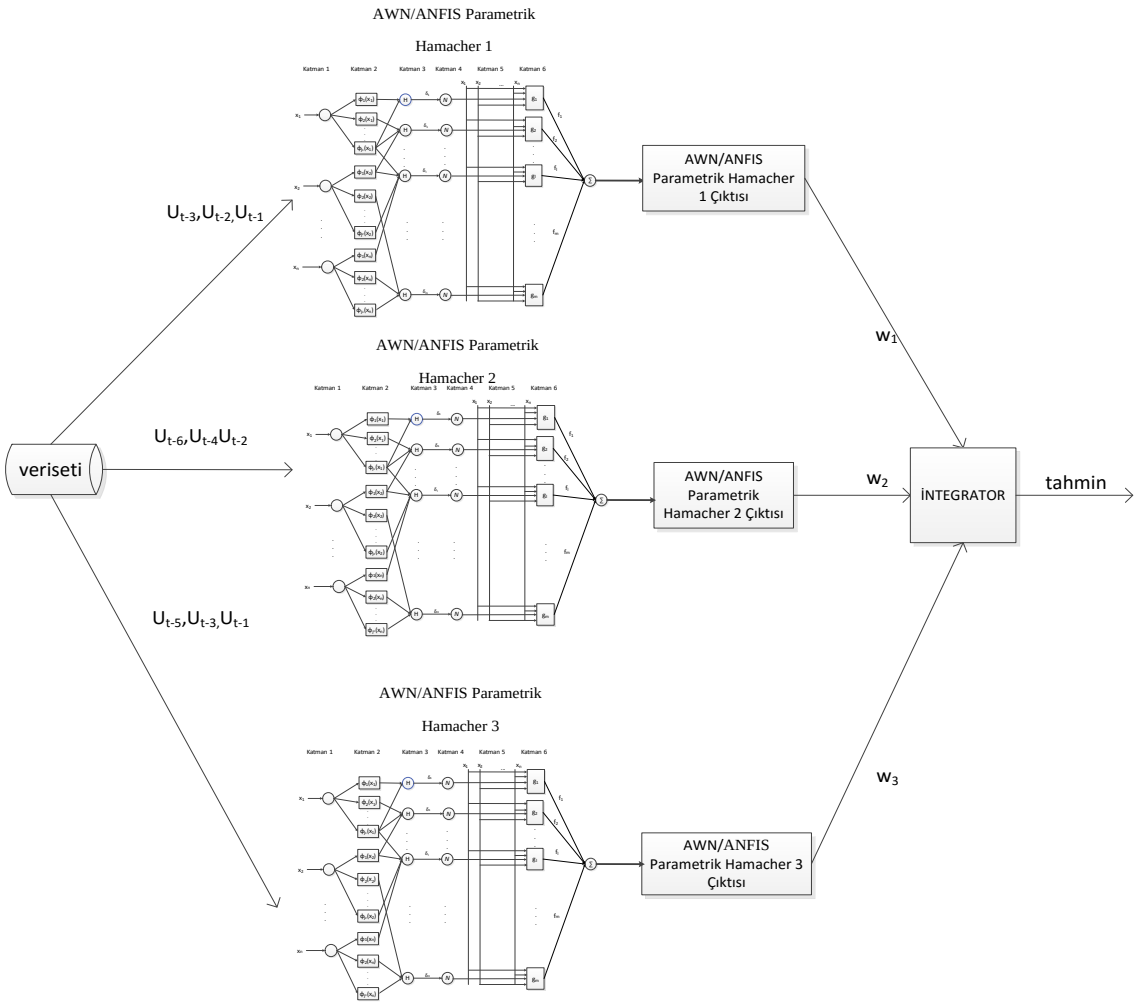
işlemine gerek duyulmamıştır. Normalizasyon yöntemi, genel olarak kullanılan max-min. normalizasyonu, (5.6)'da verilmektedir.

$$x_{norm} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{5.6}$$

5.2.1 Dow-Jones zaman serisi tahmini

Dow-Jones zaman serisi Amerika Birleşik Devletleri'nin önemli hisse senetlerinden biridir. En büyük 30 şirketi içermektedir. İlk yayını 26 Mayıs 1896 yılında yapılmıştır [80]. Topluluk yapılarında, Dow-Jones zaman serisi tahmini yapılırken, öncelikle ANFIS-PH ve AWN-PH yapıları dikkate alınmıştır. Tezde ANFIS-PH ve AWN-PH yapıları 3 girdi ve tahmin edilen 1 çıktıdan oluşmaktadır. Aynı zamanda topluluk yapısı da 3 girdi ve 1 çıktıdan oluşmaktadır. ANFIS-PH ve AWN-PH yapılarında 3, 2 ve 1 gün önceki, 5, 3 ve 1 gün önceki ve son olarak, 6, 4 ve 2 gün önceki veriler kullanılmıştır. Veri seti olarak, 09.05.2012 - 24.03.2016 tarihleri arasındaki günlük 1000 veri kullanılmıştır. Veriler yahoo.finance.com [81]'dan çekilmiştir. Her bir veri Amerikan borsasındaki ilk 30 şirketin günlük kapanış değerlerini göstermektedir. 1000 veri, 3 parçaya ayrılmaktadır. 1000 veri içerisinden ilk 500 veri ile sistem eğitilir, devamındaki 200 veri w_i'lerin eğitimi için ve son 300 veri ise sistemi test etmek için kullanılmıştır. Bu veri setinin seçilmesini sebebi, bu tarihler arasında, yakın zamandaki güncel veriler kullanılmak istenmiştir. Girdi ve çıktı değerleri, vektör formatı ile gösterilmek istenirse;

X(1)=[u_{t-3},u_{t-2},u_{t-1};u_t], X(2)=[u_{t-6},u_{t-4},u_{t-2};u_t] ,X(3)=[u_{t-5},u_{t-3},u_{t-1};u_t] şeklinde gösterilebilir. Burada yer alan X(1), X(2) ve X(3) girdileri ise topluluk yapısının girdilerini oluşturmaktadır.



Şekil 5.8 *Ens-ANFIS ve Ens-AWN-PH yapısı*

Dow-Jones finansal zaman serisinin, Ens-ANFIS-PH ve Ens-AWN-PH test sonuçları, ortalama ve ağırlıklı ortalamaları ile birlikte tablo 5.10 ve 5.11'de verilmektedir.

Farklı başlangıç parametreleri ile değerlendirildiğinde seçilen parametrelere göre farklı λ değerleri üretilmektedir. Bütün tablolarda son satırda yer alan parametre, en uygun λ değerleridir. İki tabloda birbirleri ile ilişkilendirildiğinde, iki yapı içinde en iyi sonucu $\lambda = 2$ için Einstein T-normun verdiği görülmektedir. Yapılar karşılaştırıldığı zaman ise Ens-AWN-PH yapısının daha iyi tahmin bulunduğu görülmektedir.

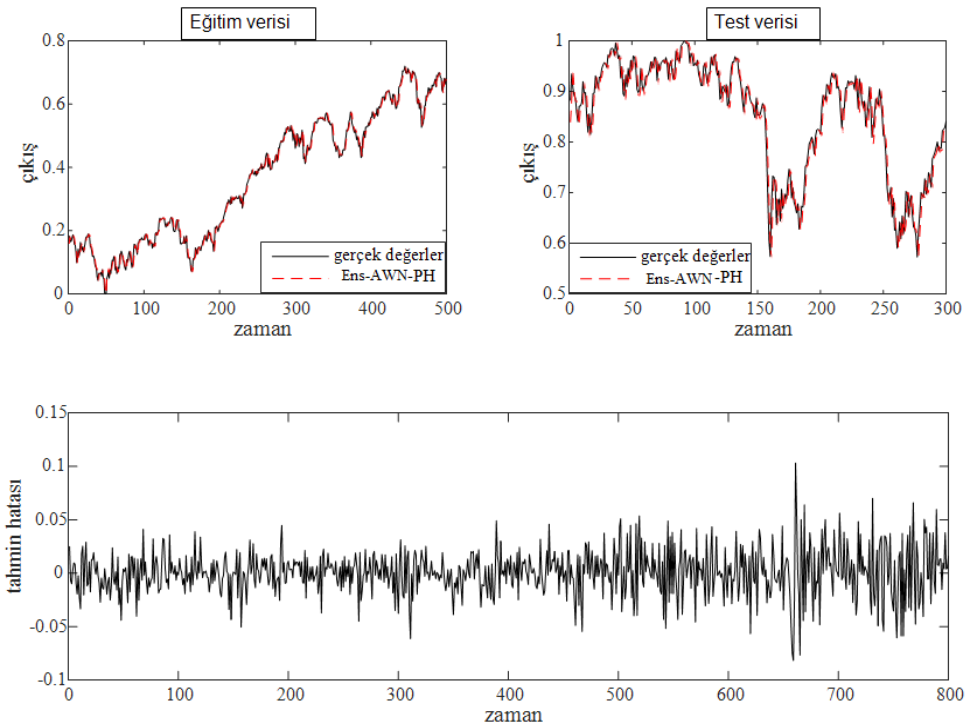
Tablo 5.10. *Dow-Jones zaman serisinin λ değerleri için Ens-ANFIS-PH sonuçları*

Ens-ANFIS-PH	OMYH-Ort.	OMYH-Ağır. Ort.	w_1	w_2	w_3
$\lambda = 0$	2,7560	2,7324	1,15	0,09	1,12
$\lambda = 1$	2,7458	2,7304	1,20	0,11	1,24
$\lambda = 2$	2,7277	2,7240	0,81	0,14	1,52
$\lambda = 0{,}8$	2,8120	2,7420	1,51	0,18	0,81

Tablo 5.11. *Dow-Jones zaman serisinin λ değerleri için Ens-AWN-PH sonuçları*

Ens-AWN-PH	OMYH-Ort.	OMYH-Ağır. Ort.	w_1	w_2	w_3
$\lambda = 0$	2,8526	2,8226	0,88	0,38	1,34
$\lambda = 1$	2,8567	2,8052	1,17	0,72	1,04
$\lambda = 2$	2,7392	2,7064	0,96	0,43	1,43
$\lambda = 0{,}3085$	2,8215	2,7317	1,10	0,41	1,26

Şekil 5.9'da Eğitim ve test verilerinin, sunulan yapı ile gerçek değerleri arasındaki ilişkisi verilmektedir. Ayrıca 800 veri için tahmin hatası da yer almaktadır.



Şekil 5.9. *Dow-Jones zaman serisinin Ens-AWN-PH yapısı için tahmin hatası ve tahmini*

Farklı λ değerleri ile sunulan modeller, klasik ANFIS ($\lambda = 1$) ve klasik AWN ($\lambda = 1$)yapılarından daha iyi tahminde bulunmuştur. Ancak, Şekil 5.9 ayrıntılı bir şekilde incelendiğinde, Dow-Jones veri seti için, test verisine uygulanan topluluk yapısının tahmini hata değerlerinin, özellikle 800 ile 1000. günler arasında arttığı görülmektedir. Bu aralık veri setinde 02.06.2015 – 24.03.2016 tarihleri arasına denk gelmektedir. Genel olarak bu tarihler arasına bakıldığında borsanın kapanış değerlerinde ani düşüşlerin ve ani çıkışların meydana geldiği görülmektedir. Bu durum ise Dow-Jones zaman serisinin tahminini zorlaştırmaktadır.

5.2.2 S&P 500 zaman serisi tahmini

S&P 500 borsa endeksi, Standart & Poor kredi derecelendirme kuruluşu tarafından yapılmaktadır. İlk 500 Amerikan şirketini içermektedir. Amerika hisse senedi piyasasının yaklaşık %75'ini kapsamaktadır [80]. S&P 500 için kullanılan veri seti, Dow-Jones ile aynı tarihler arasından seçilmiştir. Girdi değerleri ve veri setinin bölünme şekli de Dow-Jones finansal zaman serisi ile aynıdır. Tablo 5.12 ve 5.13'de λ değerleri ile birlikte sonuçlar verilmektedir.

Tablo 5.12. *S&P 500 zaman serisinin λ değerleri için Ens-ANFIS-PH sonuçları*

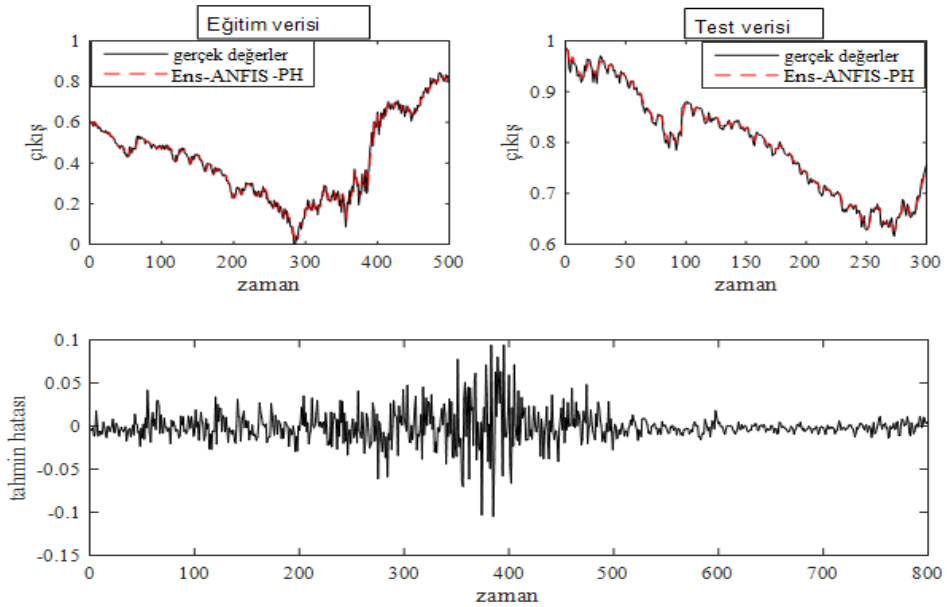
Ens-ANFIS-PH	OMYH-Ort.	OMYH-Ağır. Ort.	w_1	w_2	w_3
$\lambda = 0$	1,2572	1,2248	1,21	0,31	1,08
$\lambda = 1$	1,3089	1,2197	1,06	0,13	1,36
$\lambda = 2$	1,1979	1,1515	1,15	0,63	1,12
$\lambda = 0{,}7536$	1,2314	1,1313	1,14	0.52	1,19

Tablo 5.12 ve 5.13 dikkate alındığında Ens-ANFIS-PH için $\lambda = 0{,}7536$ ve Ens-AWN-PH için ise $\lambda = 0{,}5646$ değerinin uygun olduğu görülmüştür. Ancak iki yapı arasında en uygun λ değerinde, Ens-ANFIS-PH yapısının daha iyi tahminde bulunduğu görülmektedir.

Tablo 5.13. *S&P 500 zaman serisinin λ değerleri için Ens-AWN-PH sonuçları*

Ens-AWN-PH	OMYH-Ort.	OMYH-Ağır. Ort.	w_1	w_2	w_3
$\lambda = 0$	1,2864	1,2451	0,56	0,94	1,11
$\lambda = 1$	1,297	1,2503	0,21	0,58	1,61
$\lambda = 2$	1,2589	1,1617	0,64	0,72	1,43
$\lambda = 0{,}5646$	1,2122	1,1559	1,07	0,58	1,22

Şekil 5.10'da en iyi sonucu veren Ens-ANFIS-PH ($\lambda = 0{,}7536$) yapısı yer almaktadır. Bu yapıda tahmin hatası ile Sunulan yapının gerçek ve tahmini verileri arasındaki ilişki verilmektedir.



Şekil 5.10. *S&P 500 zaman serisinin Ens-ANFIS yapısı için tahmin hatası ve tahmini*

Sunulan yapıların klasik ANFIS ($\lambda = 1$) ve klasik AWN ($\lambda = 1$) yapılarından daha iyi tahminde bulunduğu görülmektedir. İlk 500 veri sadece eğitim için kullanılmıştır. Son 300 veri ise sunulan topluluk modelinin testi için kullanılmıştır. Şekil 5.10 incelendiğinde sunulan topluluk yapısının, test verisinin, tahmin hatasında etkili olduğu görülmektedir.

5.2.3 Euro/Dolar zaman serisi tahmini

Euro/Dolar Zaman Serisi, 1 Euro'nun Dolar cinsinden miktarını vermektedir. Veri seti olarak, usforex [82]'den alınan günlük veriler kullanılmıştır. 27.03.2011-05.01.2014 tarihleri arasındaki 1000 veri kullanılmıştır. Euro/Dolar Zaman Serisi, 1 ile 2 arasındaki sayılardan oluştuğu için, normalize işlemine ihtiyaç duyulmamıştır. Girdi verileri ile çıkış verisi diğer finansal zaman serileri gibi u_{t-3},u_{t-2},u_{t-1} girdileri mevcut iken u_t çıkış tahmini yapılmıştır. Benzer şekilde u_{t-6},u_{t-4},u_{t-2} ve u_{t-5},u_{t-3},u_{t-1} verileri mevcut iken u_t çıkış tahmini yapılmıştır. Elde edilen bu tahminlere sırasıyla; $x_1(t)$,$x_2(t)$ ve $x_3(t)$ denilirse, her bir $x_i(t)$ çıkışı, topluluk yapısı için girdi teşkil etmektedir. 1000 verinin dağılımı ilk 500 veri, eğitim verisi, sonraki 200 veri w_i'lerin eğitim için, son 300 veri ise test verisi olarak kullanılmıştır. Tablo 5.14 ve 5.15'de λ değerlerine karşılık gelen sonuçlar verilmiştir.

Tablo 5.14 'de Ens-ANFIS-PH yapısı en iyi değerinin $\lambda = 0{,}7536$'da almaktadır. Tablo 5.15'de ise Ens-AWN-PH yapısı en iyi değerinin Einstein ($\lambda = 2$) T-normda almaktadır. İki tablo karşılaştırıldığı zaman en iyi sonucu veren yapının, Einstein T-norm için Ens-AWN-PH yapısı olduğu görülmektedir.

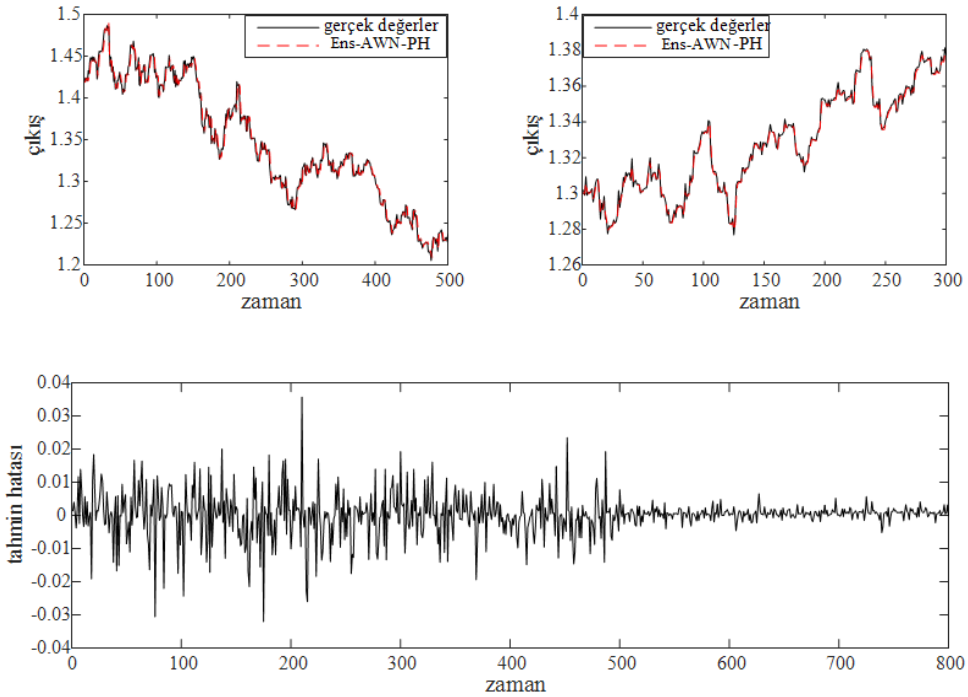
Tablo 5.14. *Euro/Dolar zaman serisinin λ değerleri için Ens-ANFIS-PH sonuçları*

Ens-ANFIS-PH	**OMYH-Ort.**	**OMYH-Ağır. Ort.**	w_1	w_2	w_3
$\lambda = 0$	0,2890	0,2834	1,21	0,31	1,08
$\lambda = 1$	0,2881	0,2821	1,0946	0,7763	1,0950
$\lambda = 2$	0,2885	0,2822	1.15	0,63	1,12
$\lambda = 0{,}7536$	0,2858	0,2803	1,14	0,52	1,19

Tablo 5.15. *Euro/Dolar zaman serisinin λ değerleri için Ens-AWN-PH sonuçları*

Ens-AWN-PH	OMYH-Ort.	OMYH-Ağır. Ort.	w_1	w_2	w_3
$\lambda = 0$	0,3027	0,2992	1,27	0,84	1,04
$\lambda = 1$	0,2920	0,2855	1,01	0,99	1,00
$\lambda = 2$	0,2847	0,2796	1,08	0,80	1,08
$\lambda = 0{,}303$	0,3013	0,2989	1,05	0,81	1,10

Şekil 5.10 ‘da Euro/Dolar zaman serisinin, Ens-AWN-PH ($\lambda = 2$) yapısı için, tahmini hatası ve gerçek değer ile olan ilişkisi verilmiştir.



Şekil 5.11. *Euro/Dolar zaman serisinin Ens-AWN-PH yapısı için tahmin hatası ve tahmini*

Şekil 5.11’de Euro/Dolar veri setinin test için kullanılan son 300 verisi dikkate alındığında, sunulan topluluk metodunun test verisi için iyi tahminde bulunduğu görülmektedir. Ayrıca bu veri seti üzerinde sunulan yapıların klasik yapılardan daha iyi sonuç verdiği görülmektedir.

6. SONUÇ

Tezde 2 farklı model sunulmaktadır. Modellerden ilki ANFIS ve AWN sistemlerinde çarpım T-normu yerine, parametresi ile birçok T-normu kapsayan, PH T-normu kullanılmıştır. Yapılan deneylerde, sistemin çıktılarının λ ile direk ilişkili olduğu görülmektedir. İlk Sunulan model olan, ANFIS–PH ve AWN–PH yapılarında, veri setlerine göre elde edilen en iyi değerlerin sonuçları Tablo 6.1'de verilmektedir.

Tablo 6.1. *ANFIS–PH ve AWN–PH modellerinin karşılaştırılması (HKOK Değerleri)*

Model	Box-Jenkins	Mackey-Glass	Sistem tanıma	Ortalama
ANFIS–PH	0,0299	0,00168	0,0226	**0,01806**
AWN–PH	0,0301	0,00156	0,0209	**0,01752**
Klasik ANFIS	0,0315	0,00225	0,0228	**0,01885**
Klasik AWN	0,0317	0,00178	0,0232	**0,01889**

Tablo 6.1'da, yapıların veri setlerine göre en iyi ağırlıklı ortalama sonuçları karşılaştırılmıştır. Veri setleri üzerinde hata ortalamalarına bakıldığında, sunulan ANFIS-PH yapısının, Çarpım T-normu ile oluşturulan klasik ANFIS yapısından daha iyi sonuç verdiği görülmektedir. Ayrıca aynı durumun AWN yapısı için de geçerli olduğu görülmektedir. İki yapı karşılaştırıldığı zaman ise uygulanan son ortalama işlemi sonucunda, AWN–PH yapısının en iyi sonucu verdiği görülmektedir. Sunulan modellerin bir diğer avantajı ise diğer yapılara göre daha az parametre ile daha iyi tahminde bulunmasıdır.

Tezde sunulan diğer model ise ilk sunulan modellerin üzerinde, topluluk yapısının uygulanmasıdır. Tezin bu kısmında, Ens-ANFIS–PH ve Ens-AWN–PH modelleri sunulmuştur. Sunulan iki yapının da, farklı veri setleri için en iyi ağırlıklı ortalama değerleri Tablo 6.2'de verilmektedir.

Tablo 6.2. *Ens-ANFIS–PH ve Ens-AWN–PH modellerinin karşılaştırılması (OMYH Değerleri)*

Model	Dow-Jones	S&P 500	Euro/Dolar	Ortalama
Ens-ANFIS–PH	2,7240	1,1313	0,2803	**1,3785**
Ens-AWN–PH	2,7064	1,1559	0,2796	**1,3806**

Tablo 6.2'den Ens-ANFIS–PH yapısının çok az bir fark ile daha iyi sonuç verdiği görülmektedir. Topluluk yapısının en büyük avantajı, paralel olarak her bir girdi için birden fazla girdi alınabilmesidir. Normalde (topluluk harici durumlarda), her bir yapıyı

eğitirken kullandığımız yöntemde de olduğu gibi, kullanılan modellerde, tek girdi çeşidi sisteme verilir ve çıktı elde edilir. Ancak tezde önerilen topluluk yapısına bakıldığında, önerilen topluluk modellerinin test aşamasında her bir girdi için u_{t-3},u_{t-2},u_{t-1} u_{t-5}, u_{t-3},u_{t-1} , u_{t-6},u_{t-4},u_{t-2} girdileri ayrı girdiler olarak alınmıştır. Bu yüzden her bir veri seti test edilirken, u_{t-6},u_{t-5},u_{t-4},u_{t-3},u_{t-2},u_{t-1} girdilerinin her biri kullanılmış olmaktadır.

Nihayetinde, her iki yapı içinde, λ değeri değiştikçe kullanılan T-norm yapısı değişeceği için, kullanılan farklı T-norm yapılarının (Einstein gibi), genel olarak kullanılan çarpım ($\lambda = 1$) T-normundan daha iyi sonuçlar elde edildiği görülmektedir. Ayrıca bazı durumlarda, optimizasyon algoritmasının sonucu olarak PH'ın özel durumu olan Einstein ($\lambda = 2$) T-normun daha iyi sonuçlar verdiği görülmektedir.

Sunulan çalışmalardan yola çıkılarak, ileri çalışmalarda, türev tabanlı olarak kullanılan öğrenme algoritması yerine, koloni optimizasyon yöntemleri kullanılabilir. Ayrıca 3. katmanda kullanılan PH T-normu yerine, trigonometrik T-norm yapıları da kullanılabilir. 2. katmanda ve 6. katmanda yer alan fonksiyonlar değiştiği taktirde farklı yapılar elde edilebilmektedir (Tezde sunulan AWN yapısı gibi). Elde edilen farklı yapılar üzerinde PH T-normu uygulanabilir.

KAYNAKÇA

[1] Jang, J. S. (1993). ANFIS: adaptive-network-based fuzzy inference system.*IEEE transactions on systems, man, and cybernetics*, *23*(3), 665-685.

[2] Mizutani, E., & Jang, J. S. (1995, November). Coactive neural fuzzy modeling. In *Neural Networks, 1995. Proceedings., IEEE International Conference on*(Vol. 2, pp. 760-765). IEEE.

[3] Yilmaz, S., & Oysal, Y. (2010). Fuzzy wavelet neural network models for prediction and identification of dynamical systems. *IEEE transactions on neural networks*, *21*(10), 1599-1609.

[4] Jang, J. S. R., & Mizutani, E. (1996, June). Levenberg-Marquardt method for ANFIS learning. In *Fuzzy Information Processing Society, 1996. NAFIPS., 1996 Biennial Conference of the North American* (pp. 87-91). IEEE.

[5] Nauck, D., & Kruse, R. (1999). Neuro-fuzzy systems for function approximation. *Fuzzy Sets and Systems*, *101*(2), 261-271.

[6] Ghomsheh, V. S., Shoorehdeli, M. A., & Teshnehlab, M. (2007, June). Training ANFIS structure with modified PSO algorithm. In *Control & Automation, 2007. MED'07. Mediterranean Conference on* (pp. 1-6). IEEE.

[7] Gupta, M. M., & Qi, J. (1991). Theory of T-norms and fuzzy inference methods. *Fuzzy sets and systems*, *40*(3), 431-450.

[8] Gál, L., & Kóczy, L. T. (2011). Fuzzy rule base extraction by bacterial type algorithms using selected t-norms. *Acta Technica Jaurinensis*, *4*(1), 157-175.

[9] Farahbod, F., & Eftekhari, M. (2012). Comparison of different t-norm operators in classification problems. *arXiv preprint arXiv:1208.1955*.

[10] Ciaramella, A., Tagliaferri, R., Pedrycz, W., & Di Nola, A. (2006). Fuzzy relational neural network. *International Journal of Approximate Reasoning*,*41*(2), 146-163.

[11] Zhang, F. Y., & Liao, Z. G. (2014). A New ANFIS Model based on Multi-Input Hamacher T-norm and Subtract Clustering. *The Open Mechanical Engineering Journal*, *8*, 833-838.

[12] Zhang, F., & Liao, Z. (2015). Stock Price Forecasting Based on Multi-Input Hamacher T-Norm and ANFIS. In *Proceedings of the Ninth International*

Conference on Management Science and Engineering Management (pp. 55-66). Springer Berlin Heidelberg.

[13] Gál, L., Lovassy, R., Rudas, I. J., & Kóczy, L. T. (2014). Learning the optimal parameter of the Hamacher t-norm applied for fuzzy-rule-based model extraction. *Neural Computing and Applications*, *24*(1), 133-142.

[14] Widyantoro, D. H., & Yen, J. (2000, May). A fuzzy similarity approach in text classification task. In *IEEE International conference on fuzzy systems* (Vol. 2, pp. 653-658).

[15] Fausett, L. (1994). *Fundamentals of neural networks: architectures, algorithms, and applications*. Prentice-Hall, Inc..

[16] Haykin, S., & Network, N. (2004). A comprehensive foundation. *Neural Networks*, *2*(2004).

[17] Hagan, M. T., Demuth, H. B., Beale, M. H., & De Jesús, O. (1996). *Neural network design* (Vol. 20). Boston: PWS publishing company.

[18] Jaeger, H. (2002). *Tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the" echo state network" approach* (pp. 1059-74). GMD-Forschungszentrum Informationstechnik.

[19] Chui, C. K. (1992). Wavelets: a tutorial in theory and applications. *Wavelet Analysis and its Applications, San Diego, CA: Academic Press,| c1992, edited by Chui, Charles K.*, *1*.

[20] Ruskai, M. B., Beylkin, G., & Coifman, R. (1992). Wavelets and their Applications. *Jones and Bartlett Books in Mathematics, Boston: Jones and Bartlett, 1992, edited by Ruskai, Mary B.; Beylkin, Gregory; Coifman, Ronald*.

[21] Zhang, Q. (1994). *Using wavelet network in nonparametric estimation* (Doctoral dissertation, INRIA).

[22] Zhang, Q., & Benveniste, A. (1992). Wavelet networks. *IEEE transactions on Neural Networks*, *3*(6), 889-898.

[23] Zhang, J., Walter, G. G., Miao, Y., & Lee, W. N. W. (1995). Wavelet neural networks for function learning. *IEEE Transactions on Signal Processing*, *43*(6), 1485-1497.

[24] Lin, C. T., & Lu, Y. C. (1995). A neural fuzzy system with linguistic teaching signals. *IEEE Transactions on Fuzzy Systems*, *3*(2), 169-189.

[25] Lin, Y., & Cunningham, G. A. (1995). A new approach to fuzzy-neural system modeling. *IEEE Transactions on Fuzzy systems*, *3*(2), 190-198.

[26] Nauck, D. (1995). Beyond neuro-fuzzy: perspectives and directions. *EUFIT*,*95*(995), 59-1164.

[27] Nauck, D., & Klawonn, F. (1996, September). Neuro-fuzzy classification initialized by fuzzy clustering. In *Proc. Fourth European Congress on Intelligent Techniques and Soft Computing (EUFIT96)* (pp. 1551-1555).

[28] Nauck, D., & Kruse, R. (1994). Choosing appropriate neuro-fuzzy models. In*Proc. Second European Congress on Fuzzy and Intelligent Technologies (EUFIT94)* (pp. 552-557).

[29] Nauck, D., Klawonn, F., & Kruse, R. (1993, June). Combining neural networks and fuzzy controllers. In *Austrian Conference on Fuzzy Logic in Artificial Intelligence* (pp. 35-46). Springer Berlin Heidelberg.

[30] Mastorocostas, P. A., & Theocharis, J. B. (2002). A recurrent fuzzy-neural model for dynamic system identification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *32*(2), 176-190.

[31] Subathra, B., Seshadhri, S., & Radhakrishnan, T. K. (2015). A comparative study of neuro fuzzy and recurrent neuro fuzzy model-based controllers for real-time industrial processes. *Systems Science & Control Engineering*, *3*(1), 412-426.

[32] Zhang, D., Bai, X. L., & Cai, K. Y. (2004). Extended neuro-fuzzy models of multilayer perceptrons. *Fuzzy Sets and Systems*, *142*(2), 221-242.

[33] Gorzalczany, M. B., & Gluszek, A. (2002). Neuro-fuzzy systems for rule-based modelling of dynamic processes. In *Advances in Computational Intelligence and Learning* (pp. 135-146). Springer Netherlands.

[34] Johansen, T. A., & Babuska, R. (2002). On multi-objective identification of Takagi-Sugeno fuzzy model parameters. *IFAC Proceedings Volumes*, *35*(1), 115-120.

[35] WU†, Z. Q., & Harris, C. J. (1997). A neurofuzzy network structure for modelling and state estimation of unknown nonlinear systems. *International journal of systems science*, *28*(4), 335-345.

[36] Kim, J., & Kasabov, N. (1999). HyFIS: adaptive neuro-fuzzy inference systems and their application to nonlinear dynamical systems. *Neural Networks*, *12*(9), 1301-1319.

[37] Lin, C. J., & Chen, C. H. (2005). Identification and prediction using recurrent compensatory neuro-fuzzy systems. *Fuzzy Sets and Systems*, *150*(2), 307-330.

[38] Lima, C. A. M., Coelho, A. L. V., & Von Zuben, F. J. (2002). Fuzzy systems design via ensembles of ANFIS. In *Fuzzy Systems, 2002. FUZZ-IEEE'02. Proceedings of the 2002 IEEE International Conference on* (Vol. 1, pp. 506-511). IEEE.

[39] Hashem, S., & Schmeiser, B. (1995). Improving model accuracy using optimal linear combinations of trained neural networks. *IEEE Transactions on Neural Networks*, *6*(3), 792-794.

[40] Chen, D. W., & Zhang, J. P. (2005, August). Time series prediction based on ensemble ANFIS. In *2005 International Conference on Machine Learning and Cybernetics* (Vol. 6, pp. 3552-3556). IEEE.

[41] Soto, J., Castillo, O., & Soria, J. (2010). Chaotic time series prediction using ensembles of ANFIS. In *Soft Computing for Intelligent Control and Mobile Robotics* (pp. 287-301). Springer Berlin Heidelberg.

[42] Lei, K. S., & Wan, F. (2012, July). Applying ensemble learning techniques to ANFIS for air pollution index prediction in Macau. In *International Symposium on Neural Networks* (pp. 509-516). Springer Berlin Heidelberg.

[43] Faulina, R., Lusia, D. A., Otok, B. W., & Kuswanto, H. (2012, September). Ensemble method based on anfis-arima for rainfall prediction. In *Statistics in Science, Business, and Engineering (ICSSBE), 2012 International Conference on* (pp. 1-4). IEEE.

[44] Breiman, L. (1996). Stacked regressions. *Machine learning*, *24*(1), 49-64.

[45] Gabralla, L. A., Wahby, T. M., Ojha, V. K., & Abraham, A. (2014, December). Ensemble of adaptive neuro-fuzzy inference system using particle swarm optimization for prediction of crude oil prices. In *Hybrid Intelligent Systems (HIS), 2014 14th International Conference on* (pp. 141-146). IEEE.

[46] Menger, K. (1942). Statistical metrics. *Proceedings of the National Academy of Sciences*, *28*(12), 535-537.

[47] Schweizer, B., & Sklar, A. (2011). *Probabilistic metric spaces*. Courier Corporation.

[48] Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, *8*(3), 338-353.

[49] Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on systems, Man, and Cybernetics*, (1), 28-44.

[50] Weber, S. (1983). A general concept of fuzzy connectives, negations and implications based on t-norms and t-conorms. *Fuzzy sets and systems*, *11*(1-3), 115-134.

[51] Bandler, W., & Kohout, L. (1980). Fuzzy power sets and fuzzy implication operators. *Fuzzy Sets and Systems*, *4*(1), 13-30.

[52] Giles, R. (1976). Łukasiewicz logic and fuzzy set theory. *International Journal of Man-Machine Studies*, *8*(3), 313-327.

[53] Yager, R. R. (1980). On a general class of fuzzy connectives. *Fuzzy sets and Systems*, *4*(3), 235-242.

[54] Dombi, J. (1982). A general class of fuzzy operators, the DeMorgan class of fuzzy operators and fuzziness measures induced by fuzzy operators. *Fuzzy sets and systems*, *8*(2), 149-163.

[55] Aczél, J., & Alsina, C. (1982). Characterizations of some classes of quasilinear functions with applications to triangular norms and to synthesizing judgements.*Aequationes Mathematicae*, *25*(1), 313-315.

[56] Gál, L., Lovassy, R., & Kóczy, L. T. (2010, July). Function approximation performance of fuzzy neural networks based on frequently used fuzzy operations and a pair of new trigonometric norms. In *Fuzzy Systems (FUZZ), 2010 IEEE International Conference on* (pp. 1-8). IEEE.

[57] Yandong, Y. (1985). Triangular norms and TNF-sigma-algebras. *Fuzzy Sets and Systems*, *16*(3), 251-264.

[58] Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. *IEEE transactions on information theory*, *36*(5), 961-1005.

[59] Oysal, Y., & Yilmaz, S. (2010). An adaptive wavelet network for function learning. *Neural Computing and Applications*, *19*(3), 383-392.

[60] Gill, P. E., Murray, W., & Wright, M. H. (1981). Practical optimization.

[61] Box, G. E. (1998). P, Jenkins, GM and Reinsel, GC (1994). Time series analysis. forecasting and control. *Prentice Hall, Englewood Cliffs, NJ. CH. Hammer/J. of Economic Behavior & Org*, *33*, 333-362.

[62] Tong, R. M. (1980). The evaluation of fuzzy models derived from experimental data. *Fuzzy sets and systems*, *4*(1), 1-12.

[63] Pedrycz, W. (1984). An identification algorithm in fuzzy relational systems.*Fuzzy sets and systems*, *13*(2), 153-167.

[64] Xu, C. W., & Lu, Y. Z. (1987). Fuzzy model identification and self-learning for dynamic systems. *IEEE Transactions on Systems, Man, and Cybernetics*,*4*(17), 683-689.

[65] Sugeno, M., & Yasukawa, T. (1991). Linguistic modeling based on numerical data. In *IFSA* (Vol. 91, pp. 264-267).

[66] Surmann, H., Kanstein, A., & Goser, K. (1993). Self-organizing and genetic algorithms for an automatic design of fuzzy control and decision systems. In *In Proc. EUFIT'93*.

[67] Kasabov, N. K., Kim, J., Watts, M. J., & Gray, A. R. (1997). FuNN/2—a fuzzy neural network architecture for adaptive learning and knowledge acquisition.*Information sciences*, *101*(3), 155-175.

[68] Chen, Y., Yang, B., & Dong, J. (2006). Time-series prediction using a local linear wavelet neural network. *Neurocomputing*, *69*(4), 449-465.

[69] Cho, K. B., & Wang, B. H. (1996). Radial basis function based adaptive fuzzy systems and their applications to system identification and prediction. *Fuzzy sets and systems*, *83*(3), 325-339.

[70] Rojas, I., Pomares, H., Bernier, J. L., Ortega, J., Pino, B., Pelayo, F. J., & Prieto, A. (2002). Time series analysis using normalized PG-RBF network with regression weights. *Neurocomputing*, *42*(1), 267-285.

[71] Kim, D., & Kim, C. (1997). Forecasting time series with genetic fuzzy predictor ensemble. *IEEE Transactions on Fuzzy systems*, *5*(4), 523-535.

[72] Juang, C. F., & Lin, C. T. (1999). A recurrent self-organizing neural fuzzy inference network. *IEEE Transactions on Neural Networks*, *10*(4), 828-845.

[73] 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.

[74] Elman, J. L. (1990). Finding structure in time. *Cognitive science*, *14*(2), 179-211.

[75] Abiyev, R. H., & Kaynak, O. (2008). Fuzzy wavelet neural networks for identification and control of dynamic plants—a novel structure and a comparative study. *IEEE transactions on industrial electronics*, *55*(8), 3133-3140.

[76] Araújo, R. D. A., Vasconcelos, G. C., & Ferreira, T. A. (2007, September). Hybrid differential evolutionary system for financial time series forecasting. In*2007 IEEE Congress on Evolutionary Computation* (pp. 4329-4336). IEEE.

[77] Ferreira, T. A., Vasconcelos, G. C., & Adeodato, P. J. (2008). A new intelligent system methodology for time series forecasting with artificial neural networks.*Neural Processing Letters*, *28*(2), 113-129.

[78] Araújo, R. D. A., Madeiro, F., de Sousa, R. P., Pessoa, L. F. C., & Ferreira, T. A. E. (2006, July). An evolutionary morphological approach for financial time series forecasting. In *2006 IEEE International Conference on Evolutionary Computation* (pp. 2467-2474). IEEE.