T.C.
İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YAPAY SİNİR AĞI MODELLERİ İLE HİSSE SENETLERİNİN DEĞİŞİM TAHMİNİNDE FARKLI DÜĞÜM SAYILARI, ZAMAN ÇERÇEVELERİ VE AKTİVASYON FONKSİYONLARININ ETKİSİNİN ARAŞTIRILMASI

YÜKSEK LİSANS TEZİ

Mustafa Ozan DUMAN

Elektrik Elektronik Mühendisliği Anabilim Dalı

Tez Danışmanı: Prof. Dr. Mehmet Emin TAĞLUK

ŞUBAT 2025

T.C
İNÖNÜ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YAPAY SİNİR AĞI MODELLERİ İLE HİSSE SENETLERİNİN DEĞİŞİM TAHMİNİNDE FARKLI DÜĞÜM SAYILARI, ZAMAN ÇERÇEVELERİ VE AKTİVASYON FONKSİYONLARININ ETKİSİNİN ARAŞTIRILMASI

YÜKSEK LİSANS TEZİ

Mustafa Ozan DUMAN
36223615935

Elektrik Elektronik Mühendisliği Ana Bilim Dalı

Tez Danışmanı: Prof. Dr. Mehmet Emin TAĞLUK

ŞUBAT 2025

TEŞEKKÜR VE ÖNSÖZ

Bu tez çalışmasının her aşamasında yardım, öneri, bilgi, deneyim ve desteklerini esirgemeden beni her konuda yönlendiren danışman hocam Sayın Prof. Dr. Mehmet Emin TAĞLUK'a,

Ayrıca tüm hayatım boyunca olduğu gibi bu çalışmalarım süresince de bana destek olan aileme,

teşekkür ederim.

ONUR SÖZÜ

Yüksek lisans tezi olarak sunduğum "YAPAY SİNİR AĞI MODELLERİ İLE HİSSE SENETLERİNİN DEĞİŞİM TAHMİNİNDE FARKLI DÜĞÜM SAYILARI, ZAMAN ÇERÇEVELERİ VE AKTİVASYON FONKSİYONLARININ ETKİSİNİN ARAŞTIRILMASI" başlıklı bu çalışmanın bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurmaksızın tarafımdan yazıldığını ve yararlandığım bütün kaynakların hem metin içinde hem de kaynakçada yöntemine uygun biçimde gösterilenlerden oluştuğunu belirtir, bunu onurumla doğrularım.

Mustafa Ozan DUMAN

İÇİNDEKİLER

TEŞEKKÜR VE ÖNSÖZ ..... **i**
ONUR SÖZÜ ..... **ii**
İÇİNDEKİLER ..... **iii**
ÇİZELGELER DİZİNİ ..... **iv**
ŞEKİLLER DİZİNİ ..... **v**
SEMBOLLER VE KISALTMALAR ..... **vi**
ÖZET ..... **vii**
ABSTRACT ..... **viii**
1. GİRİŞ **1**
2. LİTERATÜR TARAMASI **7**
3. HİSSE SENEDİNİN TAHMİN EDİLEBİLİRLİĞİ **11**
3.1 Borsa 11
3.2 Hisse Senedini Etkileyen Faktörler 12
3.2.1 Çalışmada kullanılan makroekonomik veriler 13
3.2.2 Çalışmada kullanılan mikroekonomik faktörler 14
3.3 Rastgele Yürüyüş Hipotezi 16
3.4 Rastgele Olmayan Yürüyüş Hipotezi 16
3.5 Etkin Piyasalar Hipotezi 17
3.6 Dow Teorisi 18
3.7 Hisse Senetlerini Tahmin Etmede Kullanılan Bazı Geleneksel Yöntemler 19
3.7.1 Temel analiz 19
3.7.2 Teknik analiz 20
3.7.3 YSA ile geleneksel yöntemlerin karşılaştırması 20
4. YAPAY SİNİR AĞLARI **21**
4.1 Yapay Sinir Ağlarının Tanımı 21
4.2 Yapay Sinir Hücresinin Temel Elemanları 21
4.3 MLP'nin Geri Yayılım Algorithması ile Eğitilmesi 28
4.4 Zaman Serisi Tahmininde Yaygın Kullanılan Bazı YSA Modelleri 31
4.4.1 RNN 31
4.4.1.1 LSTM 33
5. UYGULAMA **37**
5.1 MLP ile Farklı Düğüm Sayıları ve Aktivasyon Fonksiyonlarının Denenmesi 37
5.2 MLP ve LSTM'in Değişen Zaman Çerçevelerinde Farklı Frekanstaki Değişkenlerden Oluşan Veriler Kullanılarak Denenmesi 40
6. BULGULAR **44**
6.1 MLP'nin Farklı Düğüm Sayıları ve Aktivasyon Fonksiyonları ile Denenmesinin Sonuçları 44
6.2 MLP ve LSTM'in Değişen Zaman Çerçevelerinde Farklı Frekanstaki Değişkenlerden Oluşan Veriler Kullanılarak Denenmesin Sonuçları 48
7. TARTIŞMA **53**
8. SONUÇ VE ÖNERİLER **59**
KAYNAKLAR **62**
ÖZGEÇMİŞ **66**

ÇİZELGELER DİZİNİ

Çizelge 5.1 : Farklı aktivasyon fonksiyonları ve düğüm sayıları içeren MLP ile yapılan denemelerde kullanılan verilerin değişkenleri ve tanımları.....38
Çizelge 5.2 : LSTM ve MLP'nin farklı zaman çerçevelerinde denenmelerinde kullanılan verilerin makroekonomik değişkenleri ve frekansları ..42
Çizelge 6.1 : Farklı aktivasyon fonksiyonları ve düğüm sayıları içeren MLP ile yapılan denemelerden alınan en iyi sonuçlar ve algılayıcı sayıları.....45
Çizelge 6.2 : Farklı aktivasyon fonksiyonları ve düğüm sayıları içeren MLP ile yapılan denemelerin en iyisinin parametreleri ile farklı hisselerin test ve eğitim verilerinin denenmesinin MAPE sonuçları.....46
Çizelge 6.3 : Farklı aktivasyon fonksiyonları ve düğüm sayıları içeren MLP ile yapılan denemelerin en iyisinin parametreleri ile farklı hisselerin test ve eğitim verilerinin denenmesinin EDOC istatistikleri 47
Çizelge 6.4 : Farklı aktivasyon fonksiyonları ve düğüm sayıları içeren MLP ile yapılan denemelerin en iyisinin parametreleri ile farklı hisselerin test ve eğitim verilerinin denenmesinin karmaşıklık matrisleri.....47
Çizelge 6.5 : LSTM ve MLP'nin bütün hisselerin test ve eğitim verileri ile farklı zaman çerçevelerindeki denenmelerinin ortalama MAPE sonuçları 51
Çizelge 6.6 : LSTM ve MLP'nin bütün hisselerin test ve eğitim verileri ile farklı zaman çerçevelerindeki denenmelerinin ortalama ortalama %EDOC doğruluğu 52
Çizelge 7.1 : YSA ile hisse senedi tahmini konusunda yapılmış bazı önceki çalışmaların ve bu çalışmanın sonuçları 57

ŞEKİLLER DİZİNİ

Şekil 4.1 : Yapay sinir hücresinin sistematik gösterimi 22
Şekil 4.2 : Sigmoid fonksiyonu ve türevi 24
Şekil 4.3 : *Tanh* fonksiyonu ve türevi..... 25
Şekil 4.4 : ReLU ve türevi 26
Şekil 4.5 : Çok katmanlı YSA'nın genel gösterimi 27
Şekil 4.6 : Yinelemeli sinir hücresinin katlanmış ve açık halde gösterimi 32
Şekil 4.7 : LSTM hücresi sembolik gösterimi 34
Şekil 6.1 : Farklı aktivasyon fonksiyonları ve düğüm sayıları içeren MLP ile yapılan denemelerin en iyisinin öğrenme süreci 45
Şekil 6.2 : Farklı aktivasyon fonksiyonları ve düğüm sayıları içeren MLP ile yapılan denemelerin en iyisinin tahminleri ve gerçek değerleri..... 46
Şekil 6.3 : LSTM ve MLP'nin farklı hisseler için test ve eğitim verileri ile farklı zaman çerçevelerindeki denenmelerinin MAPE sonuçları..... 50
Şekil 6.4 : LSTM ve MLP'nin farklı hisseler için test ve eğitim verileri ile farklı zaman çerçevelerindeki denenmelerinin %EDOC doğruluğu sonuçları ... 51

SEMBOLLER VE KISALTMALAR

ABD : Amerika Birleşik Devletleri

ADAM : Adaptive Moment Estimation (uyarlanabilir moment tahmini)

ARIMA : Autoregressive Integrated Moving Average (bütünleşik otoregresif hareketli ortalama)

BIST : Borsa İstanbul

CNN : Convolutional Neural Networks (evreşimli sinir ağları)

EDOC : Estimated Direction Of Change (tahmini değişim yönü)

EMH : Efficient Market Hypothesis (etkin piyasalar hipotezi)

FAVÖK : Faiz, Amortisman ve Vergi Öncesi Kar

FVÖK : Faiz ve Vergi Öncesi Kar

GRU : Gated Recurrent Unit (geçitli yinelemeli birim)

GSYİH : Gayrisafi Yurtiçi Hasıla

LSTM : Long Short Term Memory (uzun kısa süreli bellek)

MAE : Mean Absolute Error (ortalama mutlak hata)

MAPE : Mean Absolute Percentage Error (ortalama mutlak yüzde hatası)

MLP : Multi Layer Perceptron (çok katmanlı algılayıcı)

MSE : Mean Squared Error (ortalama hata karesi)

NASDAQ : National Association of Securities Dealers Automated Quotations

NN : Neural Network (sinir ağı)

NYSE : New York Stock Exchange (New York borsası)

OOS : Out Of Sample (örneklem dışı)

ReLU : Rectifier Linear Unit (doğrultulmuş doğrusal birimler)

RF : Random Forest (rastgele orman)

RMSE : Root Mean Squared Error (ortalama karekök sapması)

RNN : Recurrent Neural Network (yinelemeli sinir ağı)

Tanh : Hyperbolic Tangent Function (hiperbolik tanjant fonksiyonu)

TL : Türk Lirası

USD : Amerika Birleşik Devletleri Doları

YSA : Yapay Sinir Ağları

ÖZET

Yüksek Lisans Tezi

YAPAY SİNİR AĞI MODELLERİ İLE HİSSE SENETLERİNİN DEĞİŞİM TAHMİNİNDE FARKLI DÜĞÜM SAYILARI, ZAMAN ÇERÇEVELERİ VE AKTİVASYON FONKSİYONLARININ ETKİSİNİN ARAŞTIRILMASI

Mustafa Ozan DUMAN

İnönü Üniversitesi
Fen Bilimleri Enstitüsü
Elektrik–Elektronik Mühendisliği Anabilim Dalı

66 + VIII sayfa

2025

Danışman: Prof. Dr. Mehmet Emin TAĞLUK

Çok sayıda değişkenin etkisi altında olan sistematik değişkenlerin rastgele yürüyüşünü, modern tekniklerle dahi tahmin etmek oldukça zordur. Bu tür karmaşık problemlerin çözümünde, yapay zekâ tabanlı modellerin etkinliği ve hangi modelin hangi parametrelerle daha iyi sonuç vereceği araştırmacılar tarafından merak edilmektedir. Bu tezde, bu tür bir problem olan hisse senetlerinin gelecekteki değerlerini tahmin etmek için farklı YSA modelleri çeşitli parametrelerle test edilerek, bu parametrelerin tahmin performansı üzerindeki etkileri incelenmiştir. Çalışmada, BIST ve NASDAQ borsalarından seçilen bazı hisselere ait, farklı frekansta değişkenler içeren verilerle, farklı düğüm sayıları ve aktivasyon fonksiyonlarına sahip LSTM ve MLP modelleri çeşitli zaman çerçevelerinde test edilmiştir. Denemeler iki ana örnek üzerinden gerçekleştirilmiştir. İlk örnekte farklı düğüm sayıları ve aktivasyon fonksiyonlarının etkisi incelenmiştir. Bu kapsamda, ReLU veya sigmoid aktivasyon fonksiyonu kullanılan ve farklı sayılarda düğüm içeren MLP ile tahminler yapılmıştır. İkinci örnekte, çeşitli frekansta değişkenler içeren veriler kullanılarak, çeşitli modellerin ve farklı zaman çerçevelerinin tahmin sonuçları üzerindeki etkisi incelenmiştir. Burada 1, 2, 3, 4 ve 5 günlük zaman çerçevelerinde LSTM ve MLP ile tahminler yapılmıştır. Üretilen tahminler, gerçek değerlerle MAPE ve EDOC ile karşılaştırılmıştır. Sonuçlar, düğüm sayılarının yaklaşık hesap kuralıyla belirlenebileceğini ve aktivasyon fonksiyonu seçiminde ReLU'nun Sigmoid'e göre daha iyi bir seçenek olabileceğini göstermektedir. Daha fazla sayıda ve farklı frekansta değişken içeren verilerin kullanımının tahmin doğruluğunu artırdığı görülmüştür. Kısa zaman çerçevelerinde yapılan tahminlerin rastgele yürüyüş hipoteziyle uyumlu olarak başarız olduğu görülmüştür. Zaman çerçevelerinin uzunluğu arttıkça, MAPE'de çok büyük olmayan bir artışla birlikte yön doğruluğunun belirgin şekilde iyileştiği ve LSTM modelinin MLP'ye göre daha iyi sonuçlar verdiği görülmüştür. Bu durum, elde edilen sonuçların Dow teorisiyle uyumlu olduğunu ve hisse senetlerinin geçmiş değerlerinin, gelecekteki hareketleriyle bir ilişki taşıdığını göstermektedir.

Anahtar Kelimeler: Yapay Sinir Ağları, Hisse Senedi Tahmini, Zaman Serileri, LSTM, MLP, Zaman Çerçeveleri, Düğüm Sayıları, Aktivasyon Fonksiyonları

ABSTRACT

Master Thesis

INVESTIGATION OF THE EFFECT OF DIFFERENT NODE NUMBERS, TIME FRAMES AND ACTIVATION FUNCTIONS ON STOCK CHANGE PREDICTION WITH ARTIFICIAL NEURAL NETWORK MODELS

Mustafa Ozan DUMAN

Inonu University
Graduate School of Nature and Applied Sciences
Department of Electrical and Electronics Engineering

66 + VIII pages

2025

Supervisor: Prof. Dr. Mehmet Emin TAĞLUK

The random walk of systematic variables, which are influenced by numerous factors, is very difficult to predict even with modern techniques. In solving such complex problems, the effectiveness of artificial intelligence-based models and which models yield better results with specific parameters is a subject of interest for researchers. This thesis investigates the prediction of future stock prices, a problem of this nature, using different artificial neural network models, testing various parameters and examining their impact on prediction performance. In the study, data containing variables with different frequencies from selected stocks of the BIST and NASDAQ stock exchanges were used, and LSTM and MLP models with varying numbers of nodes and activation functions were tested across different time frames. The experiments were carried out in two main examples. The first example examined the effect of different numbers of nodes and activation functions. In this context, predictions were made using MLP models with ReLU or sigmoid activation functions and varying numbers of nodes. In the second example, the effects of using data with variables at different frequencies, various models, and different time frames on prediction results were tested. In this section, predictions were made using LSTM and MLP models over 1, 2, 3, 4, and 5-day time frames. The generated predictions were compared with actual values using MAPE and EDOC metrics. The results suggest that the number of nodes can be determined by an approximate calculation rule, and that ReLU might be a better choice than Sigmoid for the activation function. Moreover, it was observed that using a greater number of variables with different frequencies improved prediction accuracy. Predictions made with short time frames were found to fail, consistent with the random walk hypothesis. As the length of the time frames increased, direction accuracy improved significantly with only a small increase in MAPE, and the LSTM model performed better than the MLP. These findings are consistent with Dow Theory, indicating that the past values of stocks have a relationship with their future movements.

Keywords: Artificial Neural Networks, Stock Price Prediction, Time Series, LSTM, MLP, Time Frames, Node Numbers, Activation Functions

1. GİRİŞ

Gelişmiş yöntemler kullanılsa bile, çok sayıda değişkenin etkisi altındaki sistemlerin gelecekteki hareketlerini tahmin etmek oldukça zordur. Bu tür zorlu sorunları yapay zeka tabanlı modellerin çözebileceği ve hangi modelin hangi parametrelerle daha iyi sonuç vereceği konusu, birçok araştırmaya konu olmaktadır. Borsada işlem gören hisselerin gelecekteki fiyatlarını tahmin etmek de bu tür zorluklara bir örnektir.

Hisse senedi, bir şirketin sahip olduğu sermayenin eşdeğer parçalardan biridir. Bir kişinin bir şirkete ait hisse senedi alması, o kişiyi şirkete ortak yapar. Bu, aynı zamanda şirketin kâr ve zararına da ortak oldukları anlamına gelir. Borsa, yatırımcıların şirket hisselerini alıp satabildiği, halka açık bir pazardır. Borsanın ortaya çıkışıyla birlikte hisse alım satımı daha önemli hale gelmiştir. Yatırımcılar, kazanç elde edebilmek için aldıkları hisselerin değerinin zamanla artmasını beklerler. Değer kazanması beklenen hisselerin alınması, değer kaybedeceği düşünülenlerin ise satılması, kâr elde etme açısından önemlidir. Bu durum, hisse tahmininin önemli bir konu haline gelmesine yol açmaktadır [1].

Bu çalışmada, hisse senedi tahmini amacıyla BIST ve NASDAQ borsalarından seçilen hisseler analiz edilmiştir. Türkiye'de borsa işlemlerini gerçekleştiren tek kurum olan BIST, ülkenin finans piyasasının ana merkezi konumundadır [2]. NASDAQ ise uluslararası düzeyde en önemli borsalardan biri olarak kabul edilir ve küresel piyasalarda büyük bir öneme sahiptir [3].

Hisse verileri, bir tür zaman serisidir. Bu nedenle, hisse senedi tahmini de bir zaman serisi tahmini yapmaktır. Gelecekteki değeri sayısal olarak tahmin etmeye çalışmak, aynı zamanda regresyon analizi yapmaktır. Hisse senetleri, diğer zaman serilerine göre daha karmaşık ve doğrusal olmayan bir yapıya sahiptir. Bunun nedeni, ekonomik, politik ve sosyal birçok faktörden aynı anda etkilenmeleridir.

Hisse değerini etkileyen en önemli faktörler, makroekonomik ve mikroekonomik faktörler olarak iki grupta incelenebilir. Makroekonomik faktörler, işletme dışındaki genel ekonomik etkenlerdir. Mikroekonomik faktörler ise işletmenin kendisine özgü olan etkenlerdir [4]. Ne kadar çok faktör hakkında bilgi edinilir ve bu bilgiler kullanılırsa, o hisse

hakkında elde edilen bilgi de o kadar artar. İdeal olarak, tüm olası faktörlerin kullanılması gerekir. Ancak hisse senetleri için bu durum mümkün görünmemektedir. Çünkü hisse senetleri çok fazla faktörden etkilenir. Örneğin, bir ünlünün bir hisse hakkında yaptığı yorum bile o hissenin değerinde değişikliklere neden olabilir. Ünlülerin yorumları gibi pek çok farklı etken de mevcuttur ve bunların hepsine ulaşmak neredeyse imkansızdır. Yine de, mümkün olduğunca makroekonomik, mikroekonomik ve diğer faktörlerden elde edilen veriler kullanılarak hisse senetleri hakkında bilgi edinmek ve işlem yapmak mümkündür.

Hisse senetlerinin doğrusal olmayan yapısı nedeniyle, tahmin edilebilirliği konusunda uzmanlar arasında farklı görüşler bulunmaktadır. Bazı uzmanlar hisselerin tahmin edilebileceğini savunurken, diğerleri bunun mümkün olmadığını öne sürmektedir. Tahmin edilebileceğini düşünenler, bu amaçla çeşitli teoriler, hipotezler ve yöntemler geliştirmiştir. Örneğin, Dow teorisi ve rastgele olmayan yürüyüş hipotezi bu teoriler arasında yer alırken [5, 6], yöntemler arasında makine öğrenmesi ve yapay zeka gibi güncel teknolojiler ile teknik analiz ve temel analiz gibi geleneksel yaklaşımlar bulunmaktadır. Buna karşılık, hisselerin tahmin edilemeyeceğini savunan uzmanlar da kendi görüşlerini destekleyen teoriler ortaya koymuştur. Bu görüşlerin başında, EMH ve rastgele yürüyüş hipotezi gelmektedir. EMH'ye göre, hisse fiyatları mevcut tüm bilgileri zaten yansıttığı için, geçmiş verilere dayanarak gelecekteki değerleri tahmin etmek mümkün değildir [7]. Rastgele yürüyüş hipotezi ise, hisse senetlerinin kısa vadede hangi yönde hareket edeceğinin tamamen rastgele olduğunu ve bu nedenle tahmin edilemeyeceğini iddia etmektedir [8].

Günümüzde yapay zekanın bir alt kümesi olan YSA, zaman serisi tahmininde sıkça kullanılmaktadır. YSA, insan beynindeki biyolojik sinir hücrelerinin çalışma şeklinden esinlenerek tasarlanmıştır. Bu ağlardaki temel sinir hücreleri, dışarıdan gelen bilgileri farklı girdiler olarak alır. Her bir girdi, kendine özgü bir ağırlıkla çarpılarak işlenir ve işlem elemanına iletilir. Gelen çarpımların tümü toplanır ve bu toplam, hücrenin doğrusal olmayan bir aktivasyon fonksiyonu ile işlenir. Sonuç, o sinir hücresinin çıktısıdır ve diğer hücrelere veri olarak iletilir. Birçok sinir hücresinin (algılayıcıların) bir araya gelmesiyle YSA oluşturulur. Bu ağların öğrenebilmesi için veri setleriyle eğitilmeleri gerekmektedir. Öğrenme tamamlandığında, tahmin, sınıflandırma veya veri analizi gibi özel görevlerde kullanılabilirler. YSA'lar, görüntü işleme, öneri sistemleri, konuşma tanıma, dil çevirisi ve zaman serisi tahmini gibi birçok alanda yaygın olarak kullanılır [9].

YSA, bir tür makine öğrenmesi yöntemi olmasına rağmen, geleneksel makine öğrenmesi yöntemlerinden, örneğin doğrusal regresyon, karar ağaçları ve rastgele orman

gibi yöntemlerden bazı önemli farklara sahiptir. Bu farklardan biri, geleneksel ML modellerinin istatistiksel tekniklere dayanarak, önceden belirlenmiş kurallar ve kalıplara göre verilerdeki desenleri tanımaya çalışmasıdır. YSA, verilerdeki desenleri öğrenerek tanımaya yönelik tasarlanmıştır. YSA, önceden tanımlı kurallar yerine, büyük veri setlerini işleyerek ve nöronlar arasındaki bağlantıların ağırlıklarını ayarlayarak, kalıpları öğrenir ve doğru tahminlerde bulunur.

Sinir ağlarındaki öğrenme süreci, nöronlar arasındaki bağlantıların ağırlıklarını işledikleri verilere göre ayarlamayı içerir. Bu ayarlama, hataların ağırlıkları geri doğru yayılarak güncellendiği geri yayılım adı verilen bir yöntemle yapılır. Buna karşılık, karar ağaçları veya destek vektör makineleri gibi geleneksel makine öğrenmesi algoritmaları, farklı optimizasyon teknikleri kullanır, örneğin açgözlü algoritmalar veya dışbükey optimizasyon. Sinir ağlarının uyarlanabilirliği, daha fazla veri sağlandıkça performanslarının iyileşmesini sağlar.

Sinir ağları, özellikle derin öğrenme modelleri, çoğu geleneksel makine öğrenmesi algoritmasından daha karmaşıktır. Bu karmaşıklık, verilerdeki karmaşık ilişkileri modellemelerine olanak tanırken, aynı zamanda daha fazla hesaplama gücü gerektirir ve aşırı uyum (overfitting) riskini artırır. Geleneksel algoritmalar ise genellikle daha basittir, bu da onları daha hızlı eğitilebilir ve daha kolay anlaşılır hale getirir. Ancak, karmaşık desenleri sinir ağları kadar etkili bir şekilde tespit edemeyebilirler.

Pratikte, sinir ağları bilgisayarla görme ve doğal dil işleme gibi alanlarda devrim yaratmış ve birçok görevde geleneksel algoritmalardan çok daha başarılı olmuştur. Ancak, geleneksel makine öğrenmesi algoritmaları, yapılandırılmış veri analizi ve açıklanabilirliğin daha önemli olduğu durumlarda hala oldukça etkilidir.

Hisse tahmininde gelecekteki sonuçların doğru bir şekilde tahmin edilmesi önemli olduğu için, YSA genellikle geleneksel makine öğrenmesi yöntemlerine tercih edilmektedir. Ayrıca, hisse senedi verileri arasındaki karmaşık ilişkileri analiz etme konusunda, YSA'nın doğrusal olmayan aktivasyon fonksiyonları, katmanları ve bu katmanlardaki algılayıcıları daha etkili bir şekilde çalışmasını sağlar. Ancak, YSA'yı oluştururken kullanılan bu parametrelerin dikkatlice seçilmesi gerekmektedir.

Hisse senedi tahminlerinde yaygın olarak kullanılan YSA modellerinden biri, MLP modelidir. MLP, bir giriş katmanı, bir veya daha fazla gizli katman ve bir çıktı katmanından oluşur. Bu katmanlar, her katmandaki nöronların, bir sonraki katmandaki nöronlarla

tamamen bağlantılı olduğu nöronlardan ya da düğümlerden oluşur. Giriş katmanındaki düğümler, bağımsız değişkenleri temsil eder. Her gizli katman, girdilerin ağırlıklı toplamını hesaplayarak doğrusal olmayan bir dönüşüm uygular ve bu verileri çıktı katmanına ileterek sonucu hesaplar. Çıktı düğümleri ise sistemin verdiği sonuca göre istenen çıkış değerlerini temsil eder [10, 11]. Bu çalışmada MLP ile denemeler yapılmıştır.

Sinir ağındaki gizli katman sayısı, çözülmek istenen probleme göre belirlenir. Bazı problemler sadece bir gizli katman gerektirirken, bazıları birden fazla gizli katman ihtiyacı duyabilir. Araştırmalar, çok fazla gizli katmana sahip ağların, sadece bir veya iki gizli katmana sahip ağlardan daha iyi sonuçlar vermediğini ortaya koymuştur [12]. Bu nedenle, bu çalışmada kullanılan MLP modeli, iki gizli katmana sahip olacak şekilde tasarlanmıştır. Ancak, bu katmanlardaki düğüm sayısını seçmek oldukça zordur çünkü yanlış seçimler, aşırı uyum veya yetersiz uyum (underfitting) gibi problemlere yol açabilir. Bu da ağın verimliliğini ve işlem süresini olumsuz etkiler [13].

Gizli katmanlardaki düğüm sayısı, ağın verimli çalışabilmesi için çok önemlidir. Eğer gizli katmanda çok az düğüm varsa, YSA verileri doğru şekilde işleyip kullanmak için yeterli kapasiteye sahip olamayacaktır ve yetersiz uyuma sebep verecektir. Diğer taraftan, çok fazla düğüm olması durumunda ise ağ, giriş verilerindeki desenleri ezberleyebilir ve bu da ağın yeni veriler üzerinde doğru tahminler yapma yeteneğini zayıflatır. Bu durum ise aşırı uyuma sebep olabilir. Genellikle, gizli katmandaki en uygun düğüm sayısı, deneme-yanılma yöntemiyle bulunur.

Aşırı uyum, gizli katmanlardaki düğüm sayısının çözülmek istenen problemin karmaşıklığına göre çok fazla olduğunda meydana gelir. Bu durumda ağ, aşırı derecede eğitilir ve bu da ağın işlem süresini olumsuz etkiler. Aşırı uyum, ağın eğitim verilerini çok iyi ezberlemesi sonucu, test verilerine genelleme yapabilme yeteneğini kaybetmesine yol açar [13].

Yetersiz uyum, ağdaki gizli katmanlardaki düğüm sayısının, çözülmesi gereken problemin karmaşıklığına göre az olduğu durumlarda meydana gelir. Bu durumda ağ, problemi yeterince iyi çözemez. Bu, yetersiz eğitim olarak da adlandırılır ve ağın verimliliğini olumsuz şekilde etkiler. Böyle bir durumda ağın işlem süresi çok düşük olur ve verimsiz sonuçlar ortaya çıkar [13].

Bahsedilen bu iki durumundan dolayı bu çalışmada farklı sayıda düğüm sayıları denenmiş ve karşılaştırılmıştır.

MLP içindeki bir diğer önemli parametre, aktivasyon fonksiyonunun seçilmesidir. Doğru aktivasyon fonksiyonunun seçilmesi, ağın genel performansı üzerinde büyük bir etkiye sahiptir. Aktivasyon fonksiyonu, bir sinir hücresinin çıktısını hesaplar ve türevi, geri yayılım algoritmasında kullanılır. YSA'da, parametrelerin doğrusal olmayan karmaşık fonksiyonlarını etkili bir şekilde ele alabilmek için aktivasyon fonksiyonu çok önemlidir. Aktivasyon fonksiyonunun temel görevi, YSA'daki her düğümün durumunu belirlemektir. Aktivasyon fonksiyonu seçmeden önce, her fonksiyonun özelliklerini, ağın ihtiyaçlarına göre dikkatlice incelemek gerekir. Aktivasyon fonksiyonunun monotonluğu, türevleri ve aralığı, bu seçimde dikkate alınması gereken önemli özelliklerdir [14]. Bu çalışmada farklı aktivasyon fonksiyonları denenmiş ve performansları karşılaştırılmıştır.

Hisse senedi tahmini için yaygın olarak kullanılan bir diğer YSA modeli de LSTM'dir. LSTM, bir tür RNN'dir ve gradyan kaybolması gibi yaygın bir soruna karşı geliştirilmiştir. Veriler sıralı olduğunda, genellikle RNN kullanılır. RNN'nin temel yapı taşı, gizli bir durumu saklayan ve her zaman adımında güncellenen bir yinelemeli birimdir. Bu geri bildirim döngüsü, ağın geçmiş girdilerden öğrenmesini ve bu bilgiyi mevcut işleme dahil etmesini sağlar. Ancak RNN'leri eğitmek daha karmaşıktır, çünkü tüm zaman adımlarında gradyanlar toplanır ve bu da gradyanın kaybolmasına yol açabilir. Bu nedenle, uzun dizilerle çalışırken standart RNN'ler verimli olamaz. LSTM, bazı gradyanların bozulmadan geçmesine izin vererek bu sorunu çözer. LSTM'lerin eğitimi, MLP'lere göre daha fazla hesaplama gücü gerektirir, çünkü yalnızca mevcut zaman damgasındaki verileri değil, geçmişteki verileri de hafızasında tutar. Bu özellik, geçmiş verilere bağlı olarak daha doğru tahminler yapabilmesini sağlar, bu yüzden hisse senedi tahminlerinde sıkça tercih edilir [15]. Bu sebeple bu çalışmada LSTM ile denemeler de yapılmıştır.

YSA uygulamalarının yazımında birçok yardımcı program bulunmaktadır. Bunların arasından yıllardır dünya çapından birçok araştırmacının kullandığı MATLAB birçok özelliği ile öne çıkmaktadır [16]. Bu çalışmada da uygulama platformu olarak kullanılabilirliği ve olanakları sebebiyle bu uygulama tercih edilmiştir.

Bu tezde, YSA kullanarak borsada işlem gören hisselerin gelecekteki değer ve yönünü tahmin etmeye yönelik bir zaman serisi tahmini yapılacaktır. Farklı algılayıcı sayıları, aktivasyon fonksiyonları, zaman çerçeveleri, modeller ve hisse verilerini oluşturan değişkenlerin frekanslarındaki farklılıkların etkileri incelenip değerlendirilerek analiz edilecektir. Çalışmanın amacı, hisse senedi tahmini gibi karmaşık durumlar için YSA parametrelerinin seçiminde karar destek sistemlerinin oluşturulmasına katkı sağlamaktır.

Hisse tahmini için YSA kullanılarak yapılan önceki çalışmaların detaylı bir incelemesi bir sonraki bölümde sunulmuştur. Bu incelemeden elde edilen bulgulara göre, daha önceki çalışmalarda önemli özelliklerin hepsi bir arada test edilmemiştir. Çoğunlukla yalnızca farklı modeller, düğüm sayıları, aktivasyon fonksiyonları, farklı frekansta değişkenler, farklı zaman çerçeveleri, sonuçların MAPE ile EDOC kullanılarak karşılaştırılması gibi bir veya birkaç özellik denenmiştir. Bu çalışmada ise tüm bu özellikler ayrıntılı bir şekilde test edilip karşılaştırılacak ve böylece literatüre katkı sağlanacaktır.

İlk bölümde, bu çalışmanın yapılmasının gerekçesi, amacı, kapsamı ve kullanılacak yöntemler açıklanmıştır. İkinci bölümde, daha önce yapılan benzer çalışmalar ele alınmıştır. Üçüncü bölümde hisse senetleri ve bunların tahmin edilebilirliği hakkında bilgi verilmiştir. Dördüncü bölümde ise yapay sinir ağları teorisi ve kullanımı detaylandırılmıştır. Beşinci bölümde yapılan uygulama ile ilgili ayrıntılar sunulmuştur. Altıncı bölümde elde edilen sonuçlar paylaşılmış, yedinci bölümde ise bu sonuçlar yorumlanmıştır. Sekizinci bölümde ise genel sonuçlar ve öneriler yer almaktadır. Kaynakça ve özgeçmiş tezin sonunda, Türkçe ve İngilizce özetler ise başında yer almaktadır.

2. LİTERATÜR TARAMASI

YSA kullanarak yapılan birçok hisse senedi tahmini çalışması bulunmaktadır. Ancak, bu çalışmaların sayısının çok fazla olması nedeniyle hepsinden bahsetmek mümkün değildir. Ayrıca, alandaki sürekli gelişmeler nedeniyle, en güncel çalışmaların incelenmesi daha önemlidir. Bu yüzden bu bölümde, son beş yıl içinde önde gelen dergi ve yayınlarda yer alan bazı önemli çalışmalar incelenmiş ve özetlenmiştir.

[17]'de, CNN, RNN ve LSTM modelleri, bir sonraki günün kapanış değerini tahmin etmek amacıyla karşılaştırılmıştır. Kullanılan veriler, 3 Temmuz 1997 ile 24 Ocak 2022 arasındaki Shanghai Composite Index'inin 6 parametreden oluşan verilerini içermektedir. Eğitim için ilk 4751 gün, doğrulama ve test için ise son 1190 gün kullanılmıştır. CNN modeli, 1D zaman serisi verilerini işlemek için tasarlanmıştır. Hem LSTM hem de RNN modelleri, tek bir gizli katmandan oluşmakta ve bu katmanda 64 üniteden oluşmaktadır. Tüm modellerde aktivasyon fonksiyonu olarak *tanh* fonksiyonu kullanılmıştır. Eğitimde ADAM iyileştirici, MAE kayıp fonksiyonu ve yığın büyüklüğü olarak 64 seçilmiştir. Tahminler ile gerçek değerler MAE ve RMSE metrikleri ile karşılaştırılmıştır. Çalışma sonunda, tek katmanlı RNN modelinin, diğer hibrit ve sade modellere göre daha iyi sonuç verdiği gözlemlenmiştir.

Bir diğer çalışma olan [18]'de, dokuz NYSE hissesinin aylık kapanış değerlerinin hareket yönü, farklı düğüm ve katman sayılarıyla yapılandırılmış MLP, RNN ve LSTM sinir ağları kullanılarak tahmin edilmiştir. Çalışma, 1971 ile 2015 yılları arasındaki günlük kapanış değerlerinden oluşan bir veri setini kullanmıştır. Veri seti, eğitim ve test için ikiye ayrılmıştır. Sinir ağı, her yıl, önceki beş yılın verileriyle eğitilmiştir. Yapılan analizlerde, hisse senetlerinin fiyat artış veya azalış yönünü tahmin etme doğruluğu ortalama %50 olarak hesaplanmıştır.

[19]'da, MLP, RNN, LSTM ve diğer modeller, Çin'in A hisse senedi borsasındaki kömür, fotovoltaik ve enerji sektörlerindeki hemen hemen tüm şirketlerin ertesi günkü kapanış değerlerini tahmin etmek için karşılaştırılmıştır. Çalışmada, 20 Mayıs 2014 ile 31 Aralık 2021 tarihleri arasındaki veriler kullanılmıştır. Bu veriler, farklı frekanslara sahip ve ticaret verilerinden hesaplanan 20 teknik istatistikten oluşmaktadır. Son 500 güne ait veriler

test için ayrılırken, geri kalan veriler eğitim için kullanılmıştır. Sonuçlar MAE ve MAPE kullanılarak karşılaştırılmıştır. Elde edilen sonuçlara göre, tekli modellerin MAPE değerleri 2 ile 17 arasında değişmiş olup, LSTM modelinin diğerlerinden daha iyi performans gösterdiği belirlenmiştir.

[20]'de MLP ve LSTM gibi çeşitli teknikler kullanılarak Fas Hisse Senedi Endeksi (MSI 20) günlük kapanış değerleri tahmin edilmiştir. Çalışma, 18 Aralık 2020 ile 9 Şubat 2023 arasındaki verilere dayanmakta olup, veriler günlük açılış, en yüksek, en düşük ve kapanış fiyatlarını içermektedir. Verilerin ilk %90'ı eğitim için, geri kalan %10'u ise test için ayrılmıştır. Modellerin en uygun parametrelerini belirlemek için ızgara arama algoritması kullanılmıştır. Eğitim ve test veri kümeleriyle eğitilen modellerden tahminler üretilmiş ve elde edilen tahminler, gerçek değerlerle MAPE kullanılarak karşılaştırılmıştır. Sonuçlara göre, MLP modeli hem eğitim hem de test verilerinde LSTM modelinden daha iyi performans göstermiştir.

[21]'de, ARIMA ve LSTM modelleri, dünyanın en değerli hisselerinden biri olan NVIDIA'nın ertesi gün kapanış değerini tahmin etmek için karşılaştırılmıştır. Çalışmada kullanılan veriler, 12 Ağustos 2013 ile 11 Ağustos 2023 arasındaki günlük kapanış değerlerini içermektedir. LSTM modeli, bir gizli katmandan oluşmakta ve bu katmanda 32 LSTM hücresi bulunmaktadır. ARIMA modeli ise 4, 1, 2 parametrelerine sahipti. Veriler 5 günlük dilimler halinde kullanılarak, sonraki günün kapanış değeri tahmin edilmiştir. Bu işlem, tüm olası günlerde tekrarlanmış ve tahmin edilen değerlerle gerçek sonuçlar RMSE kullanılarak karşılaştırılmıştır. Sonuçlara göre, LSTM modeli ARIMA modelinden daha iyi performans sergilemiştir.

[22]'de, CNN ve LSTM modelleri birleştirilerek, on bir uluslararası borsa indeksinin ertesi gün kapanış değerinin hareket yönü tahmin edilmeye çalışılmıştır. Kullanılan veriler, 4 Ocak 2010 ile 29 Aralık 2017 arasındaki günleri kapsamaktadır. CNN modeline gelen veriler 3D formatında işlenip, bir sonraki katman olan LSTM modeline iletilmiştir. LSTM katmanında ise farklı sayılarda hücrelerle denemeler yapılmıştır. Sonuçlar kesinlik (precision), duyarlılık (recall), doğruluk (accuracy) ve F1 skoru (F1-score) gibi ölçütlerle değerlendirilmiştir. Elde edilen sonuçlara göre, önerilen model diğer modellere göre daha iyi performans göstermiştir.

[23]'te, hisselerin günlük verileri ile haberler birleştirilerek, ertesi gün sonundaki yön tahmini yapılmaya çalışılmıştır. Bu amaçla MLP, LSTM, GRU ve bunların kombinasyonları

kullanılmıştır. Çalışmada, 85 hissenin 1 Ocak 2014 ile 1 Ocak 2016 arasındaki verileri kullanılmıştır. Verilerin ilk %80'i eğitim için, kalan %20'si ise test için ayrılmıştır. Sonuçlar kesinlik, duyarlılık, doğruluk ve F1 skoru gibi ölçütlerle karşılaştırılmıştır. Elde edilen sonuçlara göre, önerilen hibrit modelin, LSTM ve GRU'nun tek başına kullanılmasına kıyasla daha iyi performans gösterdiği görülmüştür.

[24]'te, iki hissenin günlük, aylık ve yıllık zaman periyotları sonunda alacağı değerler tahmin edilmeye çalışılmıştır. Bu amaçla LSTM, NN+RF, NN+MLP gibi klasik ve kombinasyon modelleri kullanılmış ve karşılaştırılmıştır. Verilerin bazı değişkenleri, diğer verilere dayalı olarak türetilmiş ve ayrıca haberlerden de veri elde edilerek tüm verilere dahil edilmiştir. Günlük tahminler için kullanılan veriler 1 Ocak 2019 ile 1 Aralık 2020 arasındaki dönemi kapsarken, aylık tahminler için veriler Ocak 2020'den Aralık 2020'ye kadar olan dönemi içermektedir. Yıllık tahminler için ise 2000 sonrası veriler kullanılmıştır. Tahminlerin doğruluğu, gerçek sonuçlarla karşılaştırılarak MSE ölçütü ile değerlendirilmiştir. Sonuçlar, önerilen modelin diğer modellere göre daha başarılı olduğunu göstermektedir.

[25]'te, Nike, Goldman Sachs, Johnson & Johnson, Pfizer ve JP Morgan Chase & Co.'nun bir sonraki gün kapanış değerleri, YSA ve RF modelleri kullanılarak tahmin edilmiştir. Bu tahminler daha sonra gerçek kapanış değerleriyle karşılaştırılmıştır. Veri seti Yahoo Finance'ten alınmış olup, 4 Mayıs 2009 ile 5 Mayıs 2019 arasındaki dönemi kapsamaktadır. Verilerin ilk 8 yılı eğitim amacıyla, son 2 yılı ise test amacıyla kullanılmıştır. Kullanılan YSA modeli, sırasıyla 6-3-1 düğüm yapısına sahip katmanlardan oluşmaktadır. Veriler, her bir şirketin hisse senedinin borsada işlem gördüğü günlerdeki açılış, en yüksek, en düşük ve kapanış değerlerini içeren finansal verilerden oluşturulmuştur. Tahminler, gerçek verilerle RMSE ve MAPE ölçütleri kullanılarak karşılaştırılmıştır. Sonuçlar, YSA modelinin RF modelinden daha başarılı olduğunu göstermektedir. Yapılan beş şirket için YSA ile yapılan tahminlerin MAPE değerleri, 0,7 ile 1,09 arasında değişmiştir.

Yapılan önceki çalışmalarda, hisse senedi tahminiyle ilgili farklı yöntemler ve denemeler gerçekleştirilmiştir. Çoğu çalışmada tek bir YSA modeli kullanılmışken, bazı araştırmalar farklı modelleri karşılaştırarak değerlendirme yapmıştır. Ayrıca, bazı çalışmalar farklı düğüm sayıları, aktivasyon fonksiyonları gibi değişkenleri deneyerek sonuçları incelemiştir. Çoğunlukla, hisse senetlerinin gelecekteki değerleri regresyon yöntemleriyle tahmin edilmiştir. Bu çalışmalardan bir kısmı sadece gelecekteki değere yakınsama üzerine odaklanırken, bir kısmı ise hisse senedinin gelecekteki yönünü de tahmin etmeye çalışmıştır.

Veriler genellikle günlük hisse senedi değerlerinden oluşmuş olsa da, bazı çalışmalarda haftalık, aylık veya çeyreklik veriler de kullanılmıştır. Tahminlerin doğruluğu genellikle regresyon analiziyle değerlendirilmiş ve MAPE, RMSE gibi yöntemler kullanılmıştır. Yön tahminlerinde ise kesinlik, duyarlılık, doğruluk ve F1 skoru gibi parametreler tercih edilmiştir [26]. Bazı çalışmalar, gerçek değerlerle oldukça yakın sonuçlar elde ederken, bazıları ise beklenen doğruluğa ulaşamamıştır. Farklı çalışmaların farklı sonuçlar alması, hisselerin farklı ülkelerde ve borsalarda tahmin edilebilirliğinin değişken olduğunu göstermektedir. Ayrıca, çalışmalar farklı zaman vadeleri sonrasındaki değerleri tahmin etmeye çalışmıştır. Çoğunlukla, hisse senetlerinin bir sonraki işlem günü sonundaki değeri tahmin edilmeye çalışılmıştır, ancak bazı çalışmalar haftalık, aylık veya yıllık tahminler yapmayı hedeflemiştir. Önceki çalışmalarda dikkati çeken bir diğer nokta ise, aynı hisse senedinin farklı zaman vadeleri sonundaki alacağı değerlerin tahminlerinin sonuçlarının farklı olmasıdır. Bu durum, hisse senetlerinin zamanla dalgalanmasının, onların doğrusal olmayan doğasını ortaya koymaktadır.

Bu çalışma, literatüre farklı açılardan katkı sağlamayı hedeflemektedir. BIST ve NASDAQ borsalarından seçilen hisselerin, çeşitli zaman vadesi sonlarında alacakları değerlerin tahmini yapılmaya çalışılacaktır. Veriler, hem düşük hem de yüksek frekanslı değişkenleri içermektedir. Çalışmada hem MLP hem de LSTM modelleri kullanılmıştır. MLP, farklı aktivasyon fonksiyonları ve değişik sayıda algılayıcı ile test edilmiştir. Tahminler, hem regresyon analizi hem de gelecekteki yön tahmini açısından yapılmıştır. Regresyon analizi için MAPE, yön tahmini için ise kesinlik, duyarlılık, doğruluk ve F1 skoru gibi metrikler kullanılmıştır. Böylece, farklı YSA modellerinin hisse tahminindeki başarıları, çeşitli parametreler ve zaman vadelerinde karşılaştırılmıştır.

3. HİSSE SENEDİNİN TAHMİN EDİLEBİLİRLİĞİ

Hisse senetlerinin tahmin edilip edilemeyeceği konusunda çeşitli görüşler bulunmaktadır. Bu görüşlerden, hisselerin tahmin edilemeyeceğini savunan iki önemli hipotez, rastgele yürüyüş hipotezi ve etkin piyasalar hipotezidir [7, 8]. Bu görüşlere karşı olarak, hisselerin tahmin edilebileceğini savunan rastgele olmayan yürüyüş hipotezi ve yüz yılı aşkın süredir kullanılan Dow teorisi gibi farklı yaklaşımlar da mevcuttur [5, 6].

Hisselerin tahmin edilebileceğini savunanlar, genellikle temel analiz ve teknik analiz gibi iki geleneksel yöntemi kullanır [27, 28]. Bunun yanı sıra, son yıllarda YSA ile yapılan tahminler de giderek daha popüler hale gelmiştir [29].

Bu bölümde, hisse senetleri ve bunların işlem gördüğü borsalar açıklanmış, ayrıca bu çalışmada kullanılan ve hisse senetlerini etkileyen değişkenler sunulmuştur. Ayrıca, uzmanlar tarafından ortaya atılan rastgele yürüyüş hipotezi, etkin piyasalar hipotezi, rastgele olmayan yürüyüş hipotezi ve Dow teorisi gibi görüşlerden bahsedilmiştir. Bunun yanı sıra, teknik ve temel analiz yöntemleri de ele alınmış ve YSA ile karşılaştırılmıştır.

3.1 Borsa

Borsa, hisse senetleri ve diğer yatırım araçlarının belirli kurallar çerçevesinde alınıp satıldığı bir pazar olarak tanımlanabilir. Bu işlemler, genellikle yüz yüze yapılmaz; yerine aracı kurumlar devreye girer. Alım satım işlemleri, arz ve talep ilişkisine dayanır; yani bir ürünün satılabilmesi için alıcısının da olması gerekir. Borsa işlemleri çoğunlukla hisse senetleri üzerinden yapılır. Tipik bir borsa işlemi, hisse senedinin değerlenmeden önce alınıp, değerlendirildikten sonra satılarak kâr elde edilmesidir [4].

Hisse senetleri, sermaye şirketlerinin ortaklık hakkını ve pay oranını belgeleyen değerli kağıtlardır. Borsada en yaygın işlem gören yatırım araçlarıdır. Hisse senetlerinin değerleri, arz ve talebe bağlı olarak sürekli değişir. Kolayca alınıp satılabilen bu senetler, uzun, orta ve kısa vadeli yatırımlar için yatırımcılar tarafından tercih edilir [4].

Türkiye'de borsa faaliyetini yürüten tek kurum, 1985 yılında İstanbul Menkul Kıymetler Borsası olarak kurulmuş ve 2013 yılında adını Borsa İstanbul kısaca BIST olarak değiştirmiştir. Ana sahibi, yüzde 80'inden fazla hissesine sahip olan Türkiye Varlık

Fonu'dur. İşlemler Türk Lirası (TL) üzerinden yapılmakta olup, 2024 itibarıyla 668 şirket işlem görmektedir [30].

Borsada endeks, bir grup hisse senedinin performansını ölçmek amacıyla belirli kurallarla oluşturulmuş göstergelerdir. Endeks, hisse senedi fiyatlarında zaman içinde meydana gelen genel değişimi yansıtır. BIST bünyesinde BIST100, BIST50 ve BIST30 gibi farklı endeksler bulunmaktadır. BIST30 endeksi, Türkiye'deki en büyük ve en aktif 30 şirketin performansını ölçer. Bu liste her üç ayda bir güncellenir. BIST30'da yer alan şirketler, aynı zamanda BIST50 ve BIST100 endekslerinde de bulunur [30].

NASDAQ, dünya çapında en önemli borsalardan biridir. 1971 yılında dünyanın ilk elektronik borsası olarak faaliyete başlamış ve 1998 yılında Amerika'da online olarak işlem yapılabilen ilk borsa olmuştur. Bugün, dünyadaki en değerli şirketlerin çoğunun hisselerinin alınıp satıldığı ve yatırımcılar tarafından iyi bilinen bir borsadır. Merkezi New York'ta bulunan bu borsa, Amerika'da işlem hacmi bakımından en büyük borsa olup, işlem gören hisselerin piyasa değerine göre NYSE'den sonra ikinci sıradadır. NASDAQ, NASDAQ Nordic borsa ağı ve birkaç ABD merkezli hisse senedi ile opsiyon borsasının da sahibi olan Nasdaq Inc. şirketine aittir. Sağlık, finans, eğlence, perakende ve gıda sektörlerindeki şirketlerin hisseleri alınıp satılmakla birlikte, NASDAQ daha çok teknoloji şirketlerinin hisselerine odaklanır. Bu borsada işlem gören firmalar, çoğunlukla Amerika'dan olmakla birlikte, Çin ve İsrail gibi yabancı ülkelerden de birçok şirket yer almaktadır [4].

3.2 Hisse Senedini Etkileyen Faktörler

Hisse senetlerini etkileyen birçok faktör bulunmaktadır. Bunlar arasında politika, sosyal medyadaki trendler ve ünlü kişilerin yaşadığı olaylar bile yer alabilir. Bunun nedeni, ekonominin herkese bağlı olması ve bu nedenle birçok farklı faktörden etkilenmesidir. Bu faktörler çok çeşitliliktedir, ancak çoğu ekonomi açısından iki ana kategoriye ayrılabilir: makroekonomik ve mikroekonomik faktörler [32].

Makroekonomik faktörler, bir işletmenin dışından kaynaklanır ve tüm piyasadaki şirketleri etkiler. Bu faktörlere örnek olarak GSYİH, perakende satışlar, ekonomik büyüme oranı, işsizlik oranı, tüketici ve üretici enflasyonu (TÜFE ve ÜFE), tüketici güven endeksi, cari işlemler dengesi, bütçe dengesi, faiz oranları, enflasyon, tarım dışı istihdam, USD/TL kuru, para arzı, petrol fiyatları ve sanayi üretim endeksi gibi göstergeler verilebilir [33].

Mikroekonomik faktörler, bir işletmeye özgü etkenlerdir. Bunlara örnek olarak şirketin finansal durumu, yönetim yapısı, sermaye artırımı, kâr dağıtım politikası, içerden öğrenenlerin ticareti (insider trading), sektördeki yeri ve pazar payı, finansal tabloların kalitesi, şirketin hisse senetlerinin mali değerleri ve risk düzeyleri verilebilir [33, 34].

Bu faktörler, ekonomi analizi yaparken kullanılır. Dolayısıyla, hisse senedi tahmini yaparken bu faktörleri dikkate almak önemlidir. Daha fazla etkenin kullanılması, tahminin doğruluğunu artıracaktır. Bu nedenle, bu tez çalışmasında parametreler seçilirken söz konusu faktörler göz önünde bulundurulmuştur.

3.2.1 Çalışmada kullanılan makroekonomik veriler

Bu çalışmada kullanılan makroekonomik faktörler arasında GSYİH, faiz oranı, altın fiyatı, petrol fiyatı, ABD Dolar Endeksi, NASDAQ Endeksi, BIST100 Endeksi ve USD/TL kuru bulunmaktadır [4, 31, 32]. Bunları açıklamak gerekirse, öncelikle GSYİH, bir ülkede üç ay, altı ay veya bir yıl gibi belirli bir dönemde yerleşik üretim birimleri tarafından üretilen nihai mal ve hizmetlerin, cari fiyatlar üzerinden toplam piyasa değerini ifade eder. Bu çalışmada GSYİH verisi, ABD Merkez Bankası tarafından her üç ayda bir yayımlanan verilerden alınmıştır.

İkinci olarak, faiz oranı, ödünç verilen veya alınan ana para üzerinden belirli bir dönemde ödenmesi gereken faiz miktarını ifade eder. Ödünç alınan veya verilen paranın üzerindeki toplam faiz, ana paraya, faiz oranına, bileşik faiz sıklığına ve paranın ödünç verildiği veya alındığı süreye bağlıdır. Bu çalışmada kullanılan faiz oranı verisi ABD Merkez Bankası'nın aylık olarak yayımladığı verilere dayanmaktadır.

Üçüncü olarak, altın fiyatlarının ekonomi üzerindeki etkisi, tarih boyunca süregelen bir olgudur. Altının uzun süre bozulmadan kalabilmesi, diğer maddelerle etkileşime girmemesi bu durumu açıklayabilir. Ayrıca, ülkeler varlıklarını değerlendirirken sahip oldukları altın rezervlerini de göz önünde bulundururlar. Tüm bu nedenlerle, altın fiyatının haftalık değeri bu çalışmaya dahil edilmiştir [35].

Dördüncü sırada yer alan petrol fiyatı, genellikle bir varil (159 litre) ham petrolün spot fiyatını ifade eder. Bu fiyat, Batı Teksas Ara Petrolü, Brent Ham Petrolü, Dubai Ham Petrolü, OPEC Referans Sepeti gibi farklı petrol türleri için bir referans noktasıdır. Petrol fiyatları, bir ülkenin yerel üretiminden ziyade, küresel arz ve talep dengesine göre belirlenir. Bu çalışmada, bu değişkenin günlük verileri kullanılmıştır.

Beşinci olarak, ABD Dolar Endeksi, ABD dolarının bir grup yabancı para birimine karşı değerini gösteren bir endekstir. Bu endeks, ABD'nin ticaret ortaklarının para birimlerinin oluşturduğu bir sepeti temel alır. ABD doları diğer para birimlerine karşı değer kazandıkça, endeksin değeri yükselir. Bu çalışmada, bu değişkenin günlük verileri kullanılmıştır.

Altıncı sıradaki NASDAQ Endeksi, NASDAQ borsasında işlem gören neredeyse tüm hisse senetlerini kapsayan bir endekstir. Dow Jones Endüstri Ortalaması ve S&P 500 ile birlikte, Amerika Birleşik Devletleri'ndeki en çok izlenen üç borsa endeksinden biridir. Bu çalışmada, bu endeksin günlük verileri kullanılmıştır.

Bir sonraki değişken olan BIST100, Türkiye'deki en önemli hisse senedi endeksidir. BIST100, İstanbul Borsası'nda işlem gören en büyük 100 şirketin hisse senedi fiyatlarının performansını gösterir. Bu nedenle, Türkiye'nin en güvenilir hisselerini (blue chips) temsil eder ve Türk hisse senedi piyasasının lider endeksidir. Çalışmada bu endeksin günlük verileri kullanılmıştır.

Son olarak, USD/TL döviz kuru, Amerikan Doları ile Türk Lirası arasındaki değişim oranını gösterir. Bu oran, para birimlerinin uluslararası değeri hakkında bilgi verdiği için oldukça önemlidir. Çalışmada bu değişkenin günlük verileri kullanılmıştır.

3.2.2 Çalışmada kullanılan mikroekonomik faktörler

Bu çalışmada kullanılan makroekonomik faktörler arasında piyasa değeri, hasılat, brüt kâr, net gelir, FAVÖK, FVÖK, hisse başına kazanç, toplam varlıklar, toplam yükümlülükler, hissedar öz sermayesi, net nakit akışı ve hisse senedi bazlı tazminat yer almaktadır [4, 32-34]. Bu değişkenlerden sadece piyasa değeri günlük verilerle kullanılırken, diğer tüm veriler ilgili şirketler tarafından açıklanan çeyrek (üç aylık) raporlardan alınmıştır [36].

Bunları açıklamak gerekirse, öncelikle piyasa değeri bir şirketin veya varlığın finansal piyasada ne kadar değerli olduğunu gösteren bir ölçüttür. Bir şirketin piyasa değeri, dolaşımdaki hisse senedi sayısı mevcut hisse fiyatıyla çarpılarak bulunur. Bu değer piyasa katılımcıları tarafından belirlenir ve genellikle piyasa kapitalizasyonu olarak da adlandırılır. Bu veriler İş Bankası Yatırım'dan alınmıştır [37].

İkinci değişken olan hasılat, bir işletmenin temel faaliyetlerinden elde ettiği toplam geliri ifade eder. Genellikle "ticari gelir" veya "ciro" olarak adlandırılır ve mal ve hizmet satışlarından elde edilen gelirleri kapsar.

Brüt kâr, şirketin mal veya hizmet üretimi ve satışı için yaptığı maliyetleri çıkardıktan sonra kalan kazançtır. Brüt kâr, hasılattan satılan malların maliyetinin çıkarılmasıyla hesaplanır.

Net gelir, bir şirketin belirli bir dönem için, satışlardan elde ettiği gelirden, satılan malların maliyeti, giderler, amortisman, faiz ve vergiler düşüldükten sonra kalan tutardır.

Faiz, vergi, amortisman ve itfa öncesi kazanç veya FAVÖK, net gelire alternatif bir kârlılık ölçüsüdür. Amortisman, itfa, faiz ve vergiler hariç tutularak, şirketin operasyonlarından elde edilen nakit kârı temsil etmeye çalışır.

FVÖK, bir şirketin kârlılığını gösteren bir başka ölçüttür. Bu, vergi ve faiz giderleri hariç, gelirden yapılan giderlerin çıkarılmasıyla hesaplanır. FVÖK, işletme kazancı veya işletme kârı olarak da adlandırılabilir.

Hisse başına kazanç, şirketin net gelirinin, tedavüldeki toplam hisse senedi sayısına bölünmesiyle hesaplanan ve her bir hisse senedinin kazancını gösteren bir ölçüttür.

Toplam varlıklar, bir şirketin sahip olduğu tüm varlıkların defter değerlerinin toplamını ifade eder. Net değerli borç sözleşmelerinde sıkça kullanılır. Amortisman sonrası varlıkların değeri toplam varlıklara dahil edilir.

Toplam yükümlülükler, şirketin veya bireyin dış taraflara olan borçları ve yükümlülüklerini ifade eder. Bu, bilançoda toplam varlıklardan çıkarılarak öz sermaye hesaplanır.

Hissedar öz sermayesi, bir şirketin net değeridir ve şirket tasfiye edilip tüm borçları ödendikten sonra hissedarlara iade edilecek toplam tutar olarak tanımlanır. Hissedar öz sermayesi, toplam varlıklardan toplam yükümlülüklerin çıkarılmasıyla hesaplanır.

Net nakit akışı, bir şirketin belirli bir süre içinde, şirkete giren ve çıkan para arasındaki farktır.

Hisse senedi bazlı tazminat, şirketin çalışanlarına sermaye sahipliği hakları vererek, yönetim, hissedarlar ve çalışanların çıkarlarını uyumlu hale getirmek amacıyla yapılan ödülleri ifade eder [32].

3.3 Rastgele Yürüyüş Hipotezi

Rastgele yürüyüş hipotezi, varlık fiyatlarındaki değişikliklerin tamamen rastgele olduğunu savunur. Bu hipoteze göre, hisse senedi fiyatlarındaki hareketler, öngörülemeyen olaylardan kaynaklanır. Örneğin, piyasa bir olay karşısında beklenmedik bir şekilde tepki verir ve fiyat değişiklikleri, yatırımcıların bu olayı nasıl algıladığına bağlı olarak değişir; bu da rastgele bir durumdur. Bu görüş, hisse senedi fiyatlarının tahmin edilemeyecek şekilde hareket ettiğini ve geçmiş fiyatların, gelecekteki fiyatları tahmin etmek için kullanılmasının verimli olmayacağını olduğunu belirtir. Yani, kısa vadeli fiyat değişikliklerinin tahmin edilemez olduğunu ileri sürer. Ayrıca, rastgele yürüyüş teorisi, borsanın verimli olduğunu, yani piyasada mevcut olan tüm bilgilerin zaten fiyatlara yansıdığını kabul eder ve ek risk almadan piyasayı yenmenin mümkün olmadığını öne sürer. Bu nedenle, yatırımcıların aldığı kararların, genellikle hatalı analizler ve yanlış yorumlar nedeniyle güvenilir olmadığını savunur. Bu teori, yatırım danışmanlık hizmetlerinin bir yatırımcının portföyüne pek fayda sağlamadığını iddia eder. Hatta bir maymunun, uzmanlar tarafından seçilen bir portföy kadar iyi bir seçim yapabileceğini öne sürer [8].

Rastgele yürüyüş hipotezine karşı yapılan ana eleştiri, piyasa katılımcılarının davranışlarının ve eylemlerinin fiyatlar üzerindeki etkisinin göz ardı edilmesidir. Bu yaklaşım, finansal piyasaların karmaşıklığını aşırı derecede basitleştirir. Ayrıca, fiyatların yalnızca rastgele değil, faiz oranlarındaki değişiklikler, hükümet düzenlemeleri veya içeriden bilgi ticareti gibi daha az etik faktörlerden de etkilenebileceği dikkate alınmaz.

Rastgele yürüyüş teorisi, yatırımcıların piyasa hareketlerini tahmin etmek için teknik analiz yöntemlerini kullanabileceği görüşüne karşı çıkar. Ancak bazı yatırımcılar ve analistler, bu teoriyi eleştirir. Onlara göre, geçmiş fiyat hareketleri, gelecekteki fiyatlarla ilgili değerli bilgiler sunabilir. Teknik analiz savunucuları, piyasa psikolojisini ve tarihsel fiyat kalıplarını anlamanın, yatırımcıların gelecekteki piyasa hareketlerini tahmin etmelerine yardımcı olabileceğini belirtir. Diğer bir grup yatırımcı ise, başarılı yatırımcıların (örneğin Warren Buffett gibi) uzun vadeli analizler yaparak piyasadan sürekli olarak daha iyi sonuçlar elde ettiklerini göstererek bu hipotezi sorgulamışlardır.

3.4 Rastgele Olmayan Yürüyüş Hipotezi

Piyasanın bir ölçüde öngörülebilir olduğuna inanan bazı ekonomistler, profesörler ve yatırımcılar da bulunmaktadır. Bu kişiler, fiyatların belirli trendler doğrultusunda hareket

edebileceğini ve geçmiş fiyat hareketlerinin, gelecekteki fiyat hareketlerini tahmin etmek için kullanılabileceğini savunurlar. Bu görüşü destekleyen bazı ekonomik çalışmalar yapılmış ve iki ekonomi profesörü, rastgele yürüyüş hipotezinin yanlış olduğunu kanıtlamaya çalışan bir kitap yazmıştır [6].

Davranışsal finans konusunda önde gelen bir araştırmacı olan Martin Weber, borsadaki trendleri bulmak için bir dizi test ve çalışma gerçekleştirmiştir. Weber, borsayı on yıl boyunca incelemiş ve bu süre zarfında piyasa fiyatlarında belirgin trendler aramıştır. İlk beş yılda fiyatları hızlı artan hisse senetlerinin, sonraki beş yılda genellikle düşük performans gösterdiğini gözlemlemiştir. Weber ve rastgele yürüyüş hipotezini reddeden diğer araştırmacılar, bunun rastgele yürüyüş hipotezine karşı önemli bir çelişki oluşturduğunu belirtmektedir.

Weber'in rastgele yürüyüş hipotezini çürüten bir diğer önemli testi, kazanç tahminleri yukarı yönlü revize edilen hisse senetlerinin, sonraki altı ayda diğer hisse senetlerine göre daha iyi performans göstermesiyle ilgili bulgusudur. Bu bilgi, yatırımcıların hangi hisse senetlerinin değer kazanacağını ve hangi hisse senetlerinin değer kaybedeceğini tahmin etmede bir avantaj sağlayabilir. Ayrıca, MIT Sloan School of Management ve Pennsylvania Üniversitesi'nde Finans profesörü olan Andrew W. Lo ve Archie Craig MacKinlay de rastgele yürüyüş hipotezinin yanlış olduğunu düşündükleri kanıtları sunmuşlardır [6].

3.5 Etkin Piyasalar Hipotezi

EMH, hisse senedi fiyatlarının mevcut tüm bilgileri doğru bir şekilde yansıttığını ve bu nedenle sürekli olarak karlı bir yatırım stratejisi geliştirmenin mümkün olmadığını savunan bir teoridir. EMH'ye göre, hisse senetleri her zaman doğru değerleri üzerinden işlem görür, bu da yatırımcıların değerinin altında hisse senedi alıp, aşırı fiyatlandırılmış olanları satmalarını engeller. Bu nedenle, piyasayı yenmek için uzman seçiminden veya piyasa zamanlamasından fayda sağlamak imkansızdır. Yatırımcıların daha yüksek getiri elde etmeleri yalnızca daha fazla risk almalarıyla mümkündür.

Modern finans teorisinin temel taşlarından biri olmasına rağmen, EMH oldukça tartışmalı ve sıklıkla eleştirilen bir kavramdır. EMH'nin savunucuları, temel veya teknik analizle değerinin altında hisse senedi aramanın ya da piyasanın eğilimlerini tahmin etmenin anlamsız olduğunu savunurlar. Çünkü hisse senedi fiyatları zaten mevcut tüm bilgileri yansıttığı için, en iyi yatırım stratejisi düşük maliyetli ve pasif bir portföy oluşturmak

olacaktır. Teorik olarak, ne teknik ne de temel analiz sürekli olarak aşırı getiri (alfa) sağlamaz; ancak içeriden alınan bilgiler, aşırı riskle daha yüksek getiri sağlayabilir.

Akademisyenler, EMH'yi destekleyen birçok kanıt bulunduğunu belirtirken, buna karşı çıkanlar da mevcuttur. Örneğin, Warren Buffett gibi yatırımcılar, EMH'ye göre tanımı gereği verimsiz olan bir şekilde uzun vadede piyasayı sürekli olarak yenecek şekilde yatırım yapmıştır. EMH'ye karşı çıkanlar, 1987 borsa çöküşü ve varlık balonları gibi olayların, hisse senedi fiyatlarının adil değerlerinden sapabileceğini gösteren örnekler olduğunu savunurlar.

EMH, piyasaların verimli olduğunu savunur, ancak bazı piyasaların açıkça daha verimsiz olduğu gözlemlenmiştir. Verimsiz piyasa, bir varlığın fiyatının gerçek değerini doğru şekilde yansıtmadığı bir durumu tanımlar. Bu verimsizlikler, bilgi eşitsizlikleri, düşük likidite, yüksek işlem maliyetleri, piyasa psikolojisi ve duygusal tepkiler gibi nedenlerden kaynaklanabilir. Bu durumlar, piyasa başarısızlıklarına yol açabilir.

EMH genellikle üç şekilde ele alınır: zayıf, yarı güçlü ve güçlü. Her bir form, piyasaların ne kadar verimli olduğunu farklı bir şekilde açıklar.

Güçlü biçim, tüm bilginin, hatta yayınlanmamış bilgilerin bile piyasa fiyatlarına yansıdığını söyler. Bu durumda, içeriden bilgiye sahip olsanız bile, bununla işlem yaparak kazanç sağlamanız mümkün değildir.

Yarı güçlü biçim, piyasada yayımlanmış tüm bilgilerin fiyatlara yansıdığı için, bu bilgileri kullanarak gelecekteki fiyatları tahmin etmenin verimsiz olduğunu savunur.

Zayıf biçim, geçmiş fiyatlara dayanarak gelecekteki fiyatları tahmin etmenin mümkün olmadığını belirtir. Bu, teknik analize yönelik bir eleştiridir, çünkü geçmiş fiyat hareketleri gelecekteki fiyatları tahmin etmeye yardımcı olmaz.

Rastgele yürüyüş ve etkin piyasalar hipotezleri farklı varsayımlar üzerinde olsa da, her ikisi de piyasayı sürekli olarak yenmenin neredeyse imkansız olduğunu savunur. Bu nedenle, aktif yönetim yerine pasif yatırım stratejilerini desteklerler [7].

3.6 Dow Teorisi

Rastgele yürüyüşe karşı bir alternatif teori, Dow Teorisi olarak bilinir. Dow Teorisi, hisse senedi fiyatlarının belirli bir trend (yükselen ya da düşen) içinde hareket ettiğini ve bu trendlerin farklı aşamalara (birikim, kar ve dağıtım) sahip olduğunu öne sürer. Ayrıca, fiyat hareketlerinin gücünü belirlemek için işlem hacminin önemli bir gösterge olduğuna inanılır.

19. yüzyılın sonlarında, Dow Jones & Co. ve The Wall Street Journal'ın kurucusu Charles Dow tarafından geliştirilen bu teori, hisse senedi fiyatlarının mevcut trendlere dayanarak gelecekteki hareketlerinin tahmin edilebileceğini savunur.

Dow Teorisi, rastgele yürüyüş hipotezinin aksine, hisse senedi fiyatlarının öngörülemez olduğu ve yatırımcıların piyasayı sürekli olarak geride bırakamayacağı fikrine karşı çıkar. Dow Teorisi, kısa vadeli fiyat dalgalanmalarının rastgele olabileceğini kabul eder, ancak uzun vadeli fiyat hareketlerinin ekonomik temellere dayandığını ve bu trendlerin teknik analizle tespit edilebileceğini savunur [5].

3.7 Hisse Senetlerini Tahmin Etmede Kullanılan Bazı Geleneksel Yöntemler

Hisse senedi tahminleri, araştırmacıların ilgisini her zaman çekmiştir ve bu nedenle çeşitli yöntemler geliştirilmiştir. Geleneksel olarak kullanılan iki ana yöntem temel analiz ve teknik analizdir [27, 28]. Ancak günümüzde teknolojinin gelişmesiyle birlikte daha ileri düzey yöntemler de ortaya çıkmıştır. Bu yöntemlerden biri, YSA olarak öne çıkmaktadır. Bu bölümde, temel analiz ve teknik analiz yöntemleri açıklanacak ve YSA yöntemi ile karşılaştırılacaktır.

3.7.1 Temel analiz

Temel analiz, bir hisse senedinin gerçek değerini belirlemeyi amaçlar. Bu nedenle, hisse senedinin fiyatını etkileyen karlılık, likidite ve mali yapı gibi temel faktörleri inceler. Gerçek değeri ortaya koyarak, borsada değeri düşük kalan hisselerin alınmasını, yüksek değeri olanların ise satılmasını hedefler. Temel analiz genellikle uzun vadeli tahminler yapmayı amaçlar. Bu analiz üç aşamadan oluşur: genel ekonomik konjonktür analizi, endüstri analizi ve işletme analizi.

Genel ekonomik konjonktür analizi yapılırken, hangi hisse senedinin alınacağı değil, hangi ülke borsalarında ve hangi zaman diliminde işlem yapılacağı gibi sorulara odaklanılır. Bu analizde makroekonomik faktörler kullanılır. Endüstri analizinde, şirketlerin faaliyet gösterdiği sektörler incelenir. Çünkü sektör, bir şirketin değerini önemli ölçüde etkiler. Örneğin, hammadde temini kolay ve rekabetin düşük olduğu sektörlerdeki şirketlerin hisseleri, diğerlerinden daha hızlı artış gösterebilir. Son olarak, işletme analizinde şirketin mali tabloları incelenir, hisse değerinin doğru fiyatı belirlenir ve şirketin izlediği politikalar analiz edilir. Mali tablolar, şirketin varlıkları, borçları ve diğer finansal bilgileri içeren raporlardır. Bu raporlardan elde edilen verilerle şirketin durumu değerlendirilir ve borsadaki

hisselerinin değeri, gerçek değerine göre düşük veya yüksek olduğu belirlenir. İşletme politikaları analizi, şirketin gelecekteki performansını tahmin etmeye yönelik yapılır. İşletme analizi mikroekonomik verileri kullanırken, temel analiz hem makroekonomik hem de mikroekonomik verileri incelemektedir [27, 28].

3.7.2 Teknik analiz

Teknik analiz, hisse senedi değerlerinin belirli bir trend içinde hareket ettiğini kabul eder ve geçmiş fiyat hareketlerine dayanarak gelecekteki fiyat hareketlerini tahmin etmeye çalışır. Bu analiz, temel analizden daha basittir ve kısa ile orta vadede hisse senedinin gelecekteki yönünü tahmin etmeye odaklanır. Teknik analizde işlem hacmi, hisse senedi hareketleri ve trendler gibi veriler kullanılır. Bugün kullanılan teknik analiz yöntemlerinin temeli, 1900'lü yıllarda Charles Dow tarafından geliştirdiği Dow Teorisi'ne dayanır. Bu teori ile analistler, trendleri, kırılma ve destek seviyelerini, hacim yansıma noktalarını ve diğer önemli noktaları belirlemeye çalışırlar. Teknik analizde, hisse senedinin önceki değerleri ile mevcut değerleri karşılaştırıldığından, yeterli miktarda geçmiş veri toplanması önemlidir [27, 28].

3.7.3 YSA ile geleneksel yöntemlerin karşılaştırması

İstatistiksel bilgi, YSA için harika bir temel oluşturur. İstatistiksel analiz yapılırken, veri hakkında çeşitli varsayımlarda bulunmak gerekebilir ve bazen analiz, sadece sınırlı sayıda olasılık üzerinden yapılmak zorunda kalabilir. Ancak YSA kullanıldığında, daha fazla değişkenin etkileşimleri incelenebilir. YSA, bu değişkenleri kullanarak doğru sonuçları tahmin etmeye çalışır. Ayrıca, daha fazla veri ile daha karmaşık ve birbirine bağlı analizler yapılabilir. Bu nedenle, YSA, geleneksel yöntemlere göre daha avantajlıdır [29].

4. YAPAY SİNİR AĞLARI

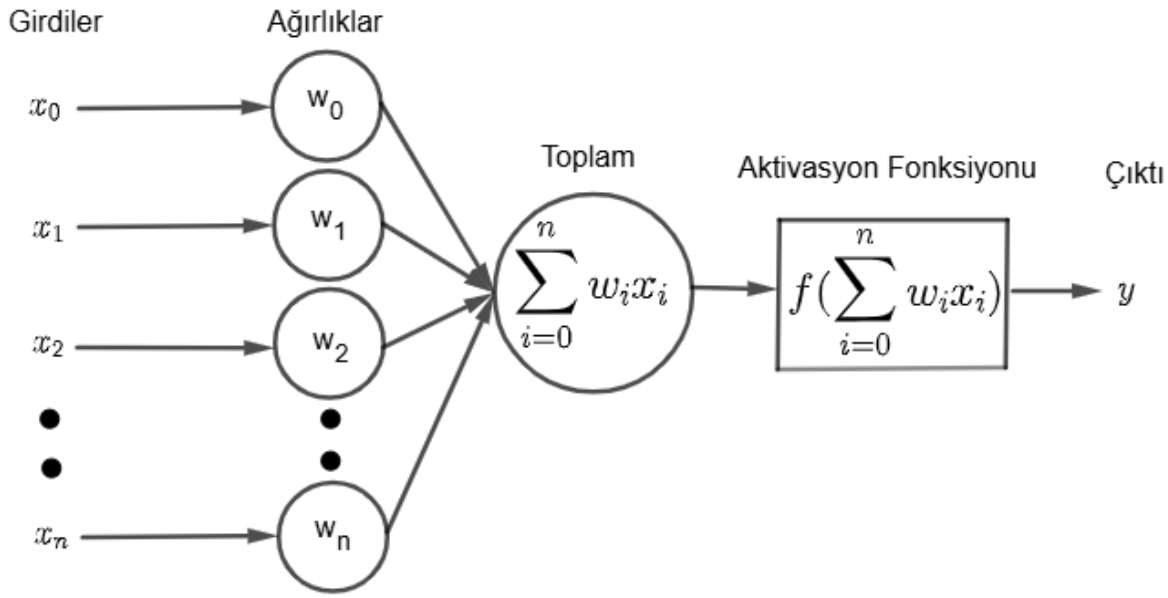
Yapay zekâ, insan zekâsını taklit ederek karmaşık problemleri çözebilen makineler geliştirmeyi amaçlayan bir bilim dalıdır. YSA ise, yapay zekânın bir alt dalı olan makine öğrenmesinin bir türüdür [9]. Bu bölümde YSA açıklanmıştır.

4.1 Yapay Sinir Ağlarının Tanımı

YSA, insan beynindeki biyolojik nöronların çalışma prensiplerinden esinlenerek tasarlanmıştır. Ağların temel yapı taşı olan nöronlar (veya algılayıcılar, düğümler), dışarıdan gelen bilgileri alır, her girdiyi kendi ağırlığı ile çarpar ve bu çarpımlar toplanarak aktivasyon fonksiyonuna uygulanır. Sonuç olarak elde edilen değer, o nöronun çıktısını oluşturur ve bu çıkış, diğer nöronlara girdi olarak verilir. Birçok sinir hücresinin bir araya gelmesiyle YSA oluşur. Bu ağların öğrenebilmesi için uygun veri setleriyle eğitilmeleri gerekir. Öğrenme tamamlandığında, ağlar istenilen uygulamalar için kullanılabilir. Görüntü işleme, öneri sistemleri, konuşma tanıma, dil çevirisi ve zaman serileri tahmini gibi alanlarda yaygın olarak kullanılırlar [38, 39].

4.2 Yapay Sinir Hücresinin Temel Elemanları

Yapay sinir hücrelerinde kullanılan temel bileşenlere işlem elemanları denir. Bu işlem elemanları beş ana bileşenden oluşur: girdiler, ağırlıklar, toplam fonksiyonu, aktivasyon fonksiyonu ve hücre çıktısı [40-42]. Bir yapay sinir hücresinin şematik gösterimi Şekil 4.1'de bulunmaktadır. Bu bölümde, bu işlem elemanları ayrıntılı olarak açıklanmıştır.



Şekil 4.1 : Yapay sinir hücresinin sistematik gösterimi

Yapay sinir hücresi, bilgileri dış dünyadan kendisine bağlanan kanallar aracılığıyla alır. Bu bilgiler, hücre için girdiler olarak kabul edilir ve işleme tabi tutulur. Girdiler, dış dünyadan, sinir hücresinin kendisinden veya diğer sinir hücrelerinden gelebilir [43]. Şekil 4.1'de gösterilen yapay sinir hücresinde x_1, x_2, ..., x_n girdileri temsil eder. Hücreye eklenen bias yani sabit değeri x_0 ile gösterilmiştir. Bias, özelliklerin ve ağırlıkların çarpımına eklenen bir sabittir ve sonucu dengelemeye yardımcı olur. Bias, modelin aktivasyon fonksiyonunun pozitif veya negatif yönde kaymasına destek olur.

YSA'ya verilen girdiler, sisteme girmeden önce normalleştirilir. Normalizasyon, girdilerin benzer bir ölçeğe sahip olmasını sağlar ve bu da eğitim sürecinin daha istikrarlı ve hızlı olmasına yardımcı olur [44]. Normalizasyonun faydaları şunlardır:

- Tüm girdi özelliklerinin benzer ölçeklere sahip olmasını sağlar.
- Belirli özelliklerin öğrenme sürecine hakim olmasını önler.
- Öğrenme sürecinde daha hızlı yakınsamayı teşvik eder.
- En uygun çözüme daha verimli bir şekilde ulaşılmasını sağlar.
- Eğim inişi algoritmalarının kayıp fonksiyonu üzerinde daha etkili gezinmesine yardımcı olur.
- Ağırlık başlatma seçimlerine duyarlılığı azaltır.
- Katmanlar arasında daha iyi ağırlık güncellemeleri sağlar.

- Modelin yeni verilere genelleme yapma yeteneğini geliştirir.
- Öğrenme sürecini daha sağlam hale getirerek aşırı uyumu azaltır.
- Aktivasyon fonksiyonlarının tutarlı bir aralıkta çalışmasını sağlar.
- Doygunluğu önler ve belirli aktivasyon fonksiyonları için öğrenmeyi hızlandırır.
- Toplu normalizasyon gibi tekniklerle iyi sinerji oluşturur.
- Eğitim sürecinin genel kararlılığına ve verimliliğine katkıda bulunur.

Özetle, YSA için girdilerin normalleştirilmesi, kararlı ve verimli bir öğrenme süreci için önemli bir adımdır. Bu işlem, yakınsamanın daha hızlı olmasını sağlar ve modelin farklı veri kümelerine etkili şekilde genellemesini mümkün kılar. Ayrıca, gradyan inişinin kararlılığına katkı sağlar ve ağırlık başlatma ile aktivasyon fonksiyonu özellikleriyle ilgili sorunları azaltır. Genel olarak, normalizasyon, sinir ağlarının performansını artırarak modern derin öğrenme uygulamalarında temel bir rol oynamaktadır [44].

Ağırlıklar her girdi için ayrı olup o girdinin önemini belirtir. Her nöral bağlantının kendine özgü ağırlığı mevcuttur. Giriş sinyali nörona verilmeden, bağlantısına özgü olan ağırlıkla çarpılır. Ağırlığın büyüklüğü ya da küçüklüğü, gelen sinyalin yükseltileceğini veya düşürüleceğini ifade eder. Ağırlığın işareti ise nörona pozitif ya da negatif etki edeceğini gösterir [45]. Şekil 4.2'de verilen temel yapay sinir hücresinde w_1, w_2 .., w_n girdilerin çarpıldığı ağırlıkları temsil etmektedir. Bias değerinin çarpıldığı ağırlık ise w_0 ile gösterilmiştir. Bu ağırlıkla bias da eğitim esnasında güncellenebilmektedir.

Toplam fonksiyonu aracılığıyla ilgili nörona gelen net girdi değeri hesaplanır ve sonuçta çıkan değer aktivasyon fonksiyonu ile değerlendirilir. En yaygın olarak kullanılan toplam fonksiyonu her girdi değerinin ve bias değerini, bunlara özgü ağırlıklar ile çarpımlarının toplanmasıdır [43]. Bu toplam (4.1)'de gösterilmiştir.

$$Toplam = \sum_{i=0}^{n} x_i w_i \tag{4.1}$$

Aktivasyon fonksiyonu hücreye gelen net girdi değerini işleyerek hücrenin çıktı değerini belirlemekte kullanılır. Farklı türleri bulunmaktadır. Sigmoid, *tanh,* ReLU ve lineer gibi fonksiyonlar örnek verilebilir. Seçim yapılırken ağın yapısı ve performansı dikkate alınır [46]. Bir tane algılayıcının çıktı değeri hesaplaması (4.2)'de verilmiştir. Burada *y* çıktı değerini gösterirken f fonksiyonu aktivasyon fonksiyonunu göstermektedir. Buradan göründüğü üzere YSA'daki en önemli elemanlardan birisidir çünkü doğrusal olmayan yapısı

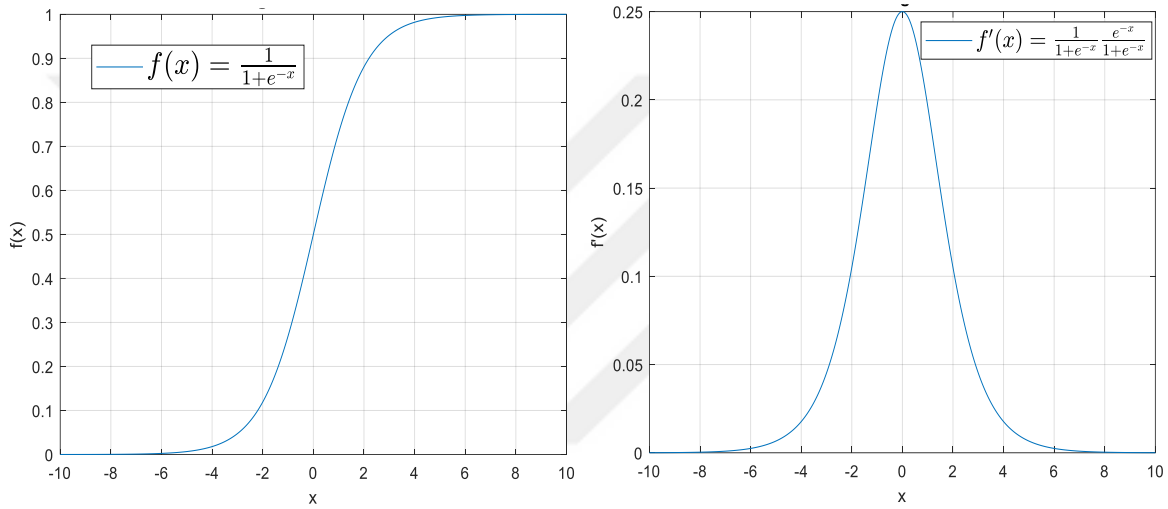
sayesinde YSA'nın verideki kompleks bağlantıları eğitim esnasında öğrenmesini mümkün kılar.

$$y = f(\sum_{i=0}^{n} x_i w_i) \tag{4.2}$$

Sigmoid çoğunlukla ileri beslemeli sinir ağlarında kullanılan doğrusal olmayan türevlenebilir ve kolay anlaşılır bir aktivasyon fonksiyonu olup, denklem (4.3) teki gibi ifade edilmektedir.

$$f(x) = \frac{1}{1+ e^{-x}} \tag{4.3}$$

Sigmoid fonksiyon ve türevinin grafikleri Şekil 4.2'de verilmiştir.



Şekil 4.2 : Sigmoid fonksiyonu ve türevi

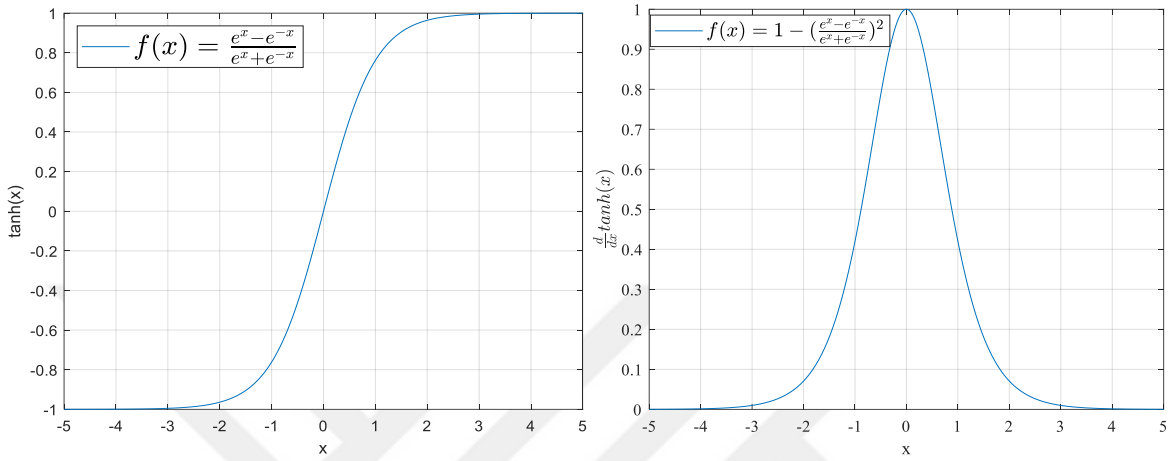
Sigmoid, giriş sinyalini alarak [0-1] aralığında olasılıksal bir çıktı üretmektedir. Bu fonksiyon, özellikle ikili sınıflandırma problemlerinde ve lojistik regresyon gibi görevlerde etkili bir şekilde uygulanmaktadır.

Ancak sigmoid, daha derin gizli katmanlardan giriş katmanlarına geri yayılım sırasında kaybolan gradyanlar, gradyan doygunluğu, yavaş yakınsama ve sıfır olmayan merkezli çıkışlar gibi önemli dezavantajlara sahiptir ve bu da gradyan güncellemelerinin farklı yönlerde yayılmasına neden olur. Hiperbolik tanjant fonksiyonu da dahil olmak üzere diğer aktivasyon fonksiyonu formları, sigmoidin bu dezavantajlardan bazılarını gidermek için önerilmiştir. Sigmoid fonksiyonunun türevinin denklemi (4.4)'te verilmiştir. Bu türevinin şekli ise Şekil 4.2'de verilmiştir.

$$f'(x) = \left(\frac{1}{1+ e^{-x}}\right)\left(\frac{e^{-x}}{1+ e^{-x}}\right) = f(x)(1 - f(x)) \tag{4.4}$$

YSA'da kullanılan diğer aktivasyon fonksiyonlardan birisi de *tanh* fonksiyonudur. Aralığı -1 ile 1 arasında olan daha düzgün sıfır merkezli türevlenebilir bir fonksiyon olup (4.5) de verilen denklem ile ifade edilmektedir. Bahsedilen *tanh* fonksiyonu ve türev grafikleri Şekil 4.3'te verilmiştir.

$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{4.5}$$



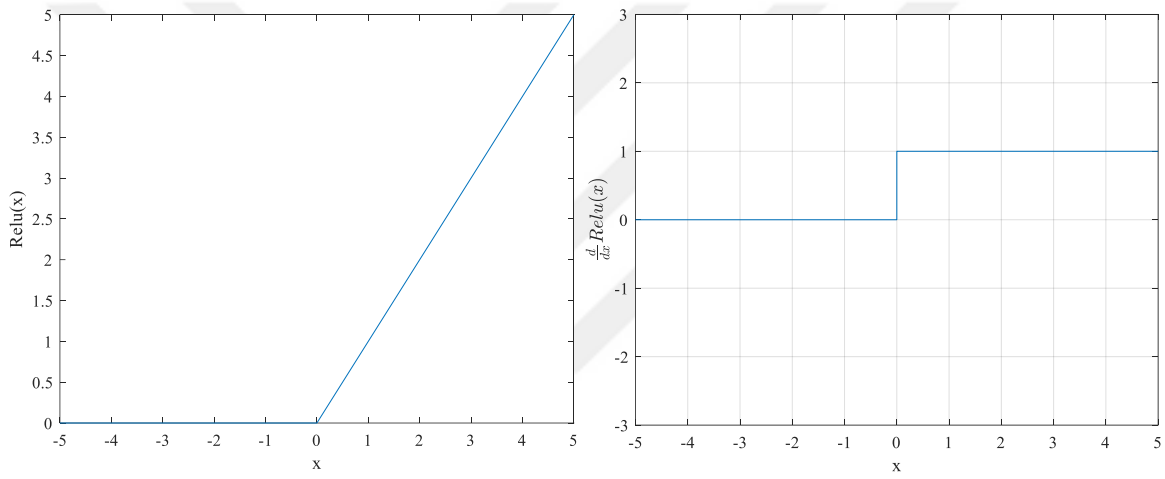
Şekil 4.3 : *Tanh* fonksiyonu ve türevi

Bu fonksiyonun türevi sigmoid fonksiyonunun türevinden daha büyük olduğu için çok katmanlı sinir ağlarında daha iyi eğitim performansı sunmaktadır. Bu sebeple sıklıkla sigmoid fonksiyonu yerine tercih edilmektedir. Ancak, *tanh* fonksiyonu da sigmoid gibi kaybolan gradyan problemiyle başa çıkamamaktadır. Bu fonksiyonun en önemli avantajı, sıfır merkezli çıktılar üretmesidir. Bu özellik, ağın geri yayılım sürecini kolaylaştırarak ağırlık güncellemelerinin daha etkili bir şekilde yapılmasına katkı sağlar. Başka bir özelliği de, yalnızca girdinin değeri 0 olduğunda 1 değerinde bir eğime ulaşabilmesidir, ki bu durum, *tanh* fonksiyonun hesaplama sırasında bazı ölü nöronlar üretmesine neden olabilmektedir. Ölü nöron, aktivasyon ağırlığının sıfır eğim sonucu nadiren kullanıldığı bir durumdur. *Tanh* fonksiyonunun bu sınırlaması, sorunu çözmek için aktivasyon fonksiyonları üzerinde daha fazla araştırma yapılmasını teşvik etmiştir ve ReLU aktivasyon fonksiyonunu doğurmuştur. *Tanh* çoğunlukla doğal dil işleme ve konuşma tanıma görevleri için RNN'de kullanılmaktadır [47]. Bu fonksiyonun türevinin denklemi (4.6)'da verilmiş ve buna ait grafik Şekil 4.3'te verilmiştir.

$$f'(x) = 1 - \left(\frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}\right)^{2} = 1 - f(x)^{2} \tag{4.6}$$

ReLU aktivasyon fonksiyonu, 2010 yılında Nair ve Hinton tarafından önerilmiş olup, günümüzde derin öğrenme uygulamalarında en sık tercih edilen aktivasyon fonksiyonu haline gelmiştir. Daha hızlı öğrenme kapasitesiyle dikkat çeken ReLU, derin öğrenme modellerinde yaygın olarak kullanılan ve başarı oranı yüksek bir fonksiyon olarak öne çıkmaktadır. Sigmoid ve *tanh* gibi diğer aktivasyon fonksiyonlarına kıyasla daha iyi performans ve genelleme yeteneği sunar. ReLU, neredeyse doğrusal bir yapıya sahip olduğu için eğim iniş yöntemleriyle doğrusal modellerin optimize edilmesini kolaylaştıran avantajlarını korur [47]. ReLU fonksiyonu (4.7) deki gibi ifade edilmektedir. RELU fonksiyonu ve türevinin grafikleri Şekil 4.4'te verilmiştir.

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \tag{4.7}$$



Şekil 4.4 : ReLU ve türevi

Bu fonksiyon, sıfırdan küçük girdilerin değerlerini düzeltir, böylece bunları sıfıra zorlar ve daha önceki aktivasyon fonksiyonu tiplerinde gözlemlenen kaybolan gradyan sorununu ortadan kaldırır. ReLU, derin sinir ağlarının gizli birimlerinde başka bir aktivasyon fonksiyonu ile birlikte, nesne sınıflandırmasında ve konuşma tanıma uygulamalarında bulunan tipik örneklerle ağın çıkış katmanlarında kullanılmıştır.

YSA'da sıklıkla tercih edilmesinin nedeni, katman sayısı arttığında alınan türevlerin *tanh* ve sigmoid fonksiyonlarında sönümlenirken ReLU'da değişmemesidir. Çünkü türevi bir olacaktır. Türevinin denklemi (4.8)'de, grafiği ise Şekil 4.4'te verilmiştir.

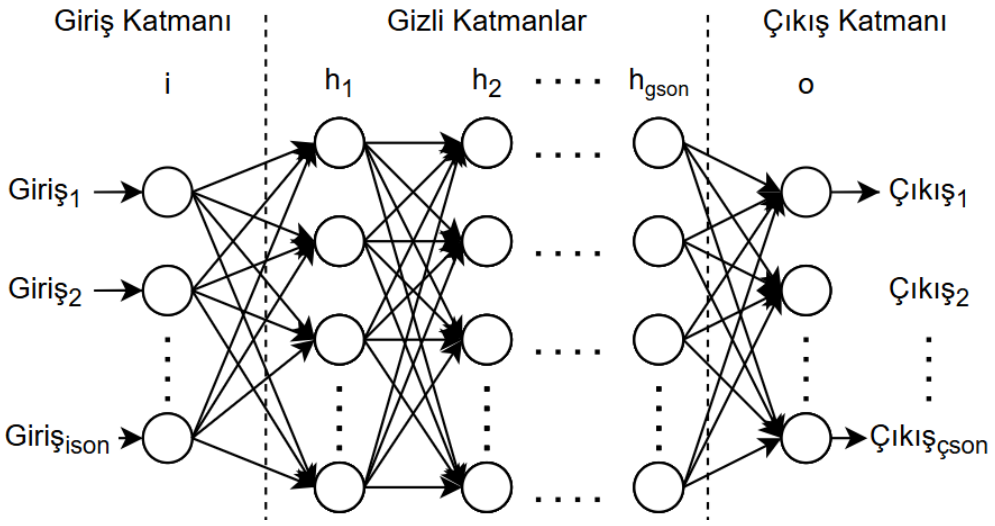
$$f'(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases} \tag{4.8}$$

Düzeltilmiş doğrusal birimlerin hesaplamada kullanılmasının temel avantajı, üssel ve bölmeleri hesaplamadığı için daha hızlı hesaplamayı garanti etmesi ve genel hesaplama

hızının artmasıdır. ReLU'nun bir diğer özelliği ise sıfırdan maksimuma kadar değerleri sıkıştırdığı için gizli birimlerde seyreklik yaratmasıdır. Ancak, ReLU'nun sigmoid fonksiyonuna kıyasla kolayca aşırı uyum sağlaması gibi bir sınırlaması mevcuttur. ReLU basitliği ve güvenilirliği nedeniyle farklı derin öğrenme mimarilerinde kullanılmıştır [47].

YSA'lar farklı mimarilerde tasarlanabilir. Her YSA mimarisinin genel amacı tahminleme veya sınıflama performansını yükseltme ve hatayı en aza indirmektir. Bu sebeple elde edilen tahminlerle gerçek değerler arasındaki farka göre kendi içindeki ağırlıklar gibi elemanlarını güncellemesi gerekmektedir. Bu farkı hesaplarken kayıp fonksiyonunu kullanır. Regresyon yaparken MSE, MAE gibi fonksiyonlar kullanılırken sınıflandırma yaparken ikili çapraz entropi, kategorik çapraz entropi gibi kayıp fonksiyonları kullanılır [9, 38].

Genel manada YSA mimarileri giriş katmanı, gizli katmanlar ve çıkış katmanı olmak üzere üç katmandan oluşmaktadır [9, 38, 48]. Şekil 4.5'te çok katmanlı YSA'nın genel gösterimi verilmiştir. Şekilden de görüldüğü gibi YSA'nın her katmandaki bir düğüm diğer katmandaki düğümler ile bağlantılandırılarak ileri beslemeli etkileşim sağlanmaktadır. Bu türden çok katmanlı YSA mimarilerine genellikle çok katmanlı algılayıcı (multi layer perceptron; MLP) denir. Bu sistemde çıktı değeri ileriye doğru üretilirken ve ağırlıklar geriye doğru güncellenir. Bundan dolayı ileri beslemeli geri yayılımlı YSA (feedforward backpropagated ANN) olarak da anılırlar.



Şekil 4.5 : Çok katmanlı YSA'nın genel gösterimi

Her YSA, MLP gibi düzgün sıralanmış ve bütün bağlantıları olan düğümlerden oluşmayabilir. Bazen bazı düğümler arasındaki bağlantılar olmayabilir veya kendinden hemen sonraki katman yerine bir sonraki katmandaki düğümlerle bağlantı kurabilirler veya

önceki katmanın girdi değerini kendisini beslemede de kullanılabilir. MLP gibi genel mimariler olsa da düğümler herhangi bir şekilde birbirlerine bağlanılarak herhangi bir mimari üretmek mümkündür.

Gizli katmanlardaki düğümler her zaman normal sinir hücrelerinden oluşmak zorunda değildir. Buna örnek olarak yenilemeli RNN hücresi ve LSTM hücresi verilebilir. Bu yapay hücreler gerçek biyolojik sinir hücreleri örnek alınarak geçmiş zamandaki deneyimin katkısını da dikkate alarak sonuç üretmektedir.

Herhangi bir YSA tasarımında katmanlardaki düğüm sayıları belirlenirken dikkatli olunmalıdır. Çok sayıda düğüm kullanımının aşırı uyuma veya az sayıda düğüm sayısı yetersiz uyuma sebep olabilmektedir. Bu sebeple bu seçim yapılırken farklı denemeler yapılarak en uygun yapıya karar verilmektedir [49]. Bu konuda sıklıkla kullanılan ve kabul gören yaklaşık hesap kuralı (rule-of-thumb) hakkında bilgi verilebilir. Bu kurala göre herhangi bir gizli katmanın düğüm sayısı kendinden önceki katmanın algılayıcı sayısının yaklaşık 2/3'ü kadar olacak şekilde ayarlanabileceği literatürde belirtilmektedir [50].

YSA sisteminde verinin ilenmesi ve değerlendirmesi sürecinde YSA'nın giriş katmanı ile tanımlanan her hücre sisteme giriş olarak belirlenen verileri belli bir ağırlıkla gizli katmandaki hücrelere iletilir [9, 38]. Gizli katmanlarda yerleştirilen hücreler belirlenen bir aktivasyon fonksiyonu ile sonuç üretip belli bir ağırlıkla bir sonraki katmanda bulunan hücrelere aktarılır. Çıkış katmanında ise gizli katmandan gelen bilgileri işleyerek çıkış değeri veya değerleri üretilir. Bu karmanda bulunan hücre sayısı istenen kullanıma göre değişme bilmektedir [9, 38]. Örneğin, regresyon amaçlı yapılarda kaç tane değişkenin tahmin edilmesi isteniyorsa çıkış katmanı o kadar düğümden oluşturulur. Benzer şekilde, sınıflandırma amaçlı yapılarda verinin kaç sınıfa ayrıştırılması isteniyorsa çıkış katmanı o kadar düğümden oluşturulur.

4.3 MLP'nin Geri Yayılım Algoritması ile Eğitilmesi

İdeal şartlarda YSA ile elde edilen sonuç ile istenilen sonuç ile aynı olması beklenmektedir. Ancak gerçek şartlarda hesaplanan sonuç ile istenilen sonuç arasında bir fark olmaktadır. Bu fark hata olarak ifade edilir ve kayıp fonksiyonu kullanılarak hesaplanır. Makine öğrenmesi kavramı da giriş ile çıkış arasında işlem gören tüm katmanlardaki ağırlıkların tekrar tekrar güncellenerek hesaplanan sonuç ile istenilen sonuç arasındaki farkı sıfırlama ya da sıfıra yakınsama işlemidir. Ağırlıklarda yapılan bu güncellemenin de standart

bir kurala göre yapılması gerekmektedir ki bu güncelleme genellikle “Gradient Descent” (gradyan inişi) metodu ile yapılmaktadır.

Zaman serisi verisi, belirli zaman damgalarıyla ilişkilendirilmiş bir dizi değerden oluşur. Eğitim işlemi esnasında, farklı zaman damgalarından (time stamp) oluşan eğitim verisi gruplara ayrılır. Bu gruplar içerdikleri zaman damgalarının sayısına göre tekli veya çoklu olabilirler. Oluşturulan bu grupların her birisi ayrı ayrı sıralı bir şekilde kullanılarak sinir ağı güncellenir. Eğer eğitim esnasında ağırlık güncellemeleri yapılırken tek bir zaman damgası içeren gruplar kullanılmışsa buna stotastik gradyan inişi (stochastic gradient descent) denir. Eğer eğitim verisi gruplandırılırken ayrılmamışsa ve bu verideki bütün zaman damgaları ağırlık güncellemesi yapılırken kullanılıyorsa bu durum toplu gradyan inişi (batch gradient descent) olacak adlandırılır. Eğer tekli veya eğitim verisindeki bütün zaman damgalarını içermeyen, bunların arasındaki miktarda zaman damgası içeren gruplar kullanılarak güncelleme yapılıyorsa bu duruma mini-batch gradyan inişi (mini-batch gradient descent) denmektedir [51, 52].

Güncelleme yapılırken eğitim verisindeki bütün zaman damgalarının hepsinin kullanıldığı her sefere bir epoch (devir) denilir. Örnek olarak eğer stotastik gradyan inişi ile eğitim yapılıyorsa, eğitim verisindeki her zaman damgası için sinir ağı tek tek güncellenir. Eğitim verisindeki bütün zaman damgaları kullanıldığında başa dönülür ve buna bir devir denir. Maksimum epoch sayısı eğitimden önce belirlenerek eğitimin sonsuza kadar sürmesinin önüne geçilir.

Oluşturulan her bir grup için sinir ağının hatası hesaplanır. Bu amaçla hata oranı hesaplanmak istenen grubun içindeki zaman damgaları tek tek sinir ağına girdi olarak verilerek her zaman damgası için çıktı elde edilir. Şekil 4.5’te göründüğü üzere her katmandaki algılayıcı, bir önceki katmandaki bütün algılayıcıların çıkış değerini girdi olarak almaktadır. Sinir ağında bulunan herhangi bir düğümün çıktısını hesaplarken önce bu düğümün bulunduğu katmanın bir önceki katmanında bulunan düğümlerin çıktıları ile yine bir önceki katmandaki düğümlerden ilgili düğüme gelen bütün ağırlıklar ile çarpılır. Elde edilen sonuç o düğümün kullandığı aktivasyon fonksiyonuna verilerek o düğümün çıktısı hesaplanır. Bu işlem ilk gizli katmandaki düğümlerden başlayarak çıktı katmanındaki düğümlere kadar yapılarak çıktılar elde edilir. Bu işleme ileriye doğru yapıldığı için ileri besleme denir. Burada çıktı değerlerinin kendinden önceki katmanlardaki bütün düğümlere bağlı olduğu görülmektedir.

Bu çıktılar gerçek değerlerle kayıp fonksiyonu kullanılarak karşılaştırılır. O gruptaki her zaman damgası için elde edilen kayıp fonksiyonu çıktıları toplanarak güncelleme için kullanılacak hata hesaplanır. Örneğin güncellemenin Q tane zaman damgası içeren bir grup kullanılarak mini-yığın gradyan inişi ile yapıldığını farz edilirse ve eğitim esnasındaki bu yığının q'nuncu zaman damgasının sinir ağına girdi olarak verildiğinde elde edilen kayıp fonksiyonu çıktısına E^q denirse, bu değerin hesaplanması (4.9)'da verildiği şekilde yapılacaktır.

$$E^q = \frac{1}{N}\sum_{i=1}^{N}(o_i^q - y_i^q)^2 \tag{4.9}$$

Verilen denklemde sinir ağının çıkış katmanında toplam N tane düğümden i'ninci düğümün çıktısı y_i^q olarak gösterilmiştir. İlgili çıktının gerçek değeri yani bulunması istenen sonuç o_i^q olarak gösterilmiştir. Regresyon analizinde sıklıkla kullanılan MSE ile bu ikisi arasındaki fark hesaplanmıştır. Bu işlemde aralarındaki farkın karesi her düğüm için ayrı ayrı hesaplanmış ve toplanmış, bulunan toplam bütün düğümlerin sayısına bölünmüştür. Böylece ilgili zaman damgasına ait kayıp fonksiyonu çıktısı hesaplanmıştır. Güncellemede kullanılan grubun hatası L ile gösterilse, (4.9)'da verilen kayıp fonksiyonu çıktısı ile bu hatanın hesaplanması (4.10)'da verilmiştir.

$$L = \sum_{q=1}^{Q} E^q \tag{4.10}$$

Hesaplanan bu hatanın güncellenmek istenen sinir ağı elemanına göre türevi alınır. Örnek olarak sinir ağında bulunan herhangi bir w ağırlığının güncellenmesi yapılmak istenirse L'nin bu ağırlığa göre türevinin alınması gerekmektedir. L'nin ilgili ağırlığa göre türevi için L'den başlayarak geriye doğru önce sinir ağının çıktı değerlerine göre türevi alınır. Buradan itibaren her seferinde bir önceki katmandaki düğümlere göre türevleri istenen ağırlığa gelene kadar her yol hesaba katılarak yapılır. Bu işlem zincir kuralı (chain rule) ile yapılır. Bulunan sonuç, belirlenmiş olan öğrenme katsayısı (η) ile çarpılarak w'dan çıkartılır. Böylece w o grup kullanılarak güncellenmiş olur. Bu işlemde w'nun güncellenmeden önceki değeri w_{eski} ile ve güncellendikten sonraki değeri w_{yeni} ile simgelenirse ilgili işlemler (4.11)'de verildiği gibi gösterilebilir.

$$w_{yeni} = w_{eski} - \eta * \frac{\delta L}{\delta w} \tag{4.11}$$

Güncelleme işlemi, her grup için o gruptaki bütün zaman damgalarının bitiminde uygulanır. Bu işlem, istenilen hata oranının altına inildiğinde veya önceden belirlenmiş olan

maksimum devir sayısına gelindiğinde sona erer. İlgili sürecin tamamına ileri besleme ve YSA'nın geri yayılım algoritması kullanarak eğitilmesi denmektedir [51-53].

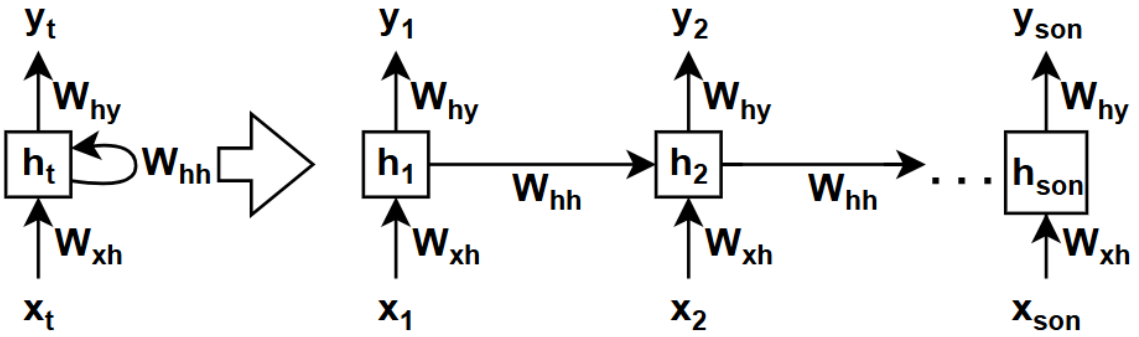
Ağdaki katman sayısı arttıkça, hatanın türevinin hesaplanmasında kullanılan zincir de uzar. Bu da gradient sönümlenmesi ya da gradient patlaması gibi sorunlara yol açabilir [9]. Zincirdeki her türev alınırken aynı zamanda aktivasyon fonksiyonunun türevi de alınır. Bu nedenle, kullanılan aktivasyon fonksiyonlarının türevlerinin büyüklüğü büyük önem taşır. Örneğin, sigmoid fonksiyonunun türevi en fazla 0.25 değerini alır. Bu durumda katman sayısı arttıkça, türevler giderek daha küçük bir değere yaklaşır. Bu durum, ağın öğrenmesini zorlaştırarak, ağırlıkların doğru şekilde güncellenmesini engelleyebilir. Diğer yandan, türevi 1'den büyük bir aktivasyon fonksiyonu kullanılırsa, türevler her katmanda artar ve bu da ağırlıkların aşırı güncellenmesine neden olabilir. Bu nedenlerle, YSA tasarımında kullanılacak aktivasyon fonksiyonu, katman sayısı ve her katmandaki nöron sayısı dikkatlice seçilmelidir.

4.4 Zaman Serisi Tahmininde Yaygın Kullanılan Bazı YSA Modelleri

Zaman serisi tahmini için kullanılan birçok YSA modeli mevcuttur. Bu bölümde bu modellerin en önemlilerinden RNN ile onun alt dallarından biri olan LSTM'e kısa bir genel bakış sunulmaktadır.

4.4.1 RNN

MLP'nin çıktı üretirken yalnızca bir tane zaman damgası kullanması, bu zaman damgasından öncekileri hafızada tutamaması ve kullanamaması, onların zaman serileri, metin serileri gibi ardışık ayrık sıralı zaman damgalarından oluşan verilerin analizinde daha az kullanışlı olmasına sebep olmaktadır. Bu soruna çözüm olarak yinelemeli sinir ağları tasarlanmıştır. Bu modelin temel elemanına yenilemeli sinir hücresi denir. Bu hücre, kendine özgün bir şekilde daha önceki ardışık zaman damgası girdi verilerinin oluşturduğu bir gizli durum verisi vektörünü de hesaplamak istediği zaman damgası verisinin yanına ekleyerek girdi olarak hesaba katar. Böylece işlem yaparken önceki zaman damgalarını da işleme katmış olur. Yinelemeli sinir hücresinin katlanmış ve açık haldeki gösterimi Şekil 4.6'da verilmiştir.



Şekil 4.6 : Yinelemeli sinir hücresinin katlanmış ve açık halde gösterimi

Şekilde $\boldsymbol{x}_t$, t'ninci zaman damgasında sisteme verilen girdi vektörünü ifade eder. Girdi değerlerine uygulanan ağırlık vektörü $\boldsymbol{w}_{xh}$, bir önceki zaman damgasından gelen gizli durum vektörü $\boldsymbol{h}_{t-1}$, bu vektörün çarpıldığı ağırlık vektörü $\boldsymbol{w}_{hh}$, mevcut zaman damgasında üretilen gizli durum vektörü $\boldsymbol{h}_t$, bu vektör ile çarpılarak mevcut zaman damgasının çıktısı olan $\boldsymbol{y}_t$ vektörünü oluştururken kullanılan ağırlık vektörü $\boldsymbol{w}_{hy}$ ile gösterilmiştir.

Toplam girdi ve çıktı sayılarına göre RNN dört gruba ayrılır. Bunlar çoklu giriş çoklu çıkış, çoklu giriş tekli çıkış, tekli giriş çoklu çıkış, tekli giriş tekli çıkış olarak isimlendirilmiştir. Her bir grup farklı durumlarda kullanılmaktadır. Örnek olarak tekli giriş çoklu çıkış içinde birden fazla kelime barındıran fotoğrafın içindeki kelimelerin çıkarılmasında kullanılabilir ve çoklu giriş tekli çıkış ise ardışık kelimelerin verilip bunların belirttiği duygunun çıkarılmasında kullanılabilir [9, 38].

Avantajlarına rağmen RNN'nin büyük bir dezavantajı bulunmaktadır. Bu da onun gradient problemidir. Çok sayıda zaman damgasından oluşan veri ile işlem yapıldığında gradient değeri yüksek miktarda artabilir veya sönümlenip bir etkisi kalmayabilir. Örnek vermek gerekirse sigmoid aktivasyon fonksiyonu kullanan bir RNN'de algılayıcının çıktı değerinin sonraki sette besleme olarak verildiği bir durumda, t'ninci zaman damgasındaki gizli durum vektörünün hesaplanması (4.12) denkleminde verilmiştir.

$$\boldsymbol{h}_t = \sigma(\boldsymbol{w}_{xh}\boldsymbol{x}_t + \boldsymbol{w}_{hh}\boldsymbol{h}_{t-1}) \tag{4.12}$$

Denklemde sigmoid fonksiyonu σ ile gösterilmiştir. Bulunan t'ninci zaman damgasının gizli durum vektörü ile $\boldsymbol{w}_{hy}$ ağırlık matrisi çarpılarak t'ninci zaman damgasındaki çıktı vektörü bulunur. İlgili işlem (4.13)'te verilmiştir.

$$\boldsymbol{y}_t = \boldsymbol{w}_{hy}\boldsymbol{h}_t \tag{4.13}$$

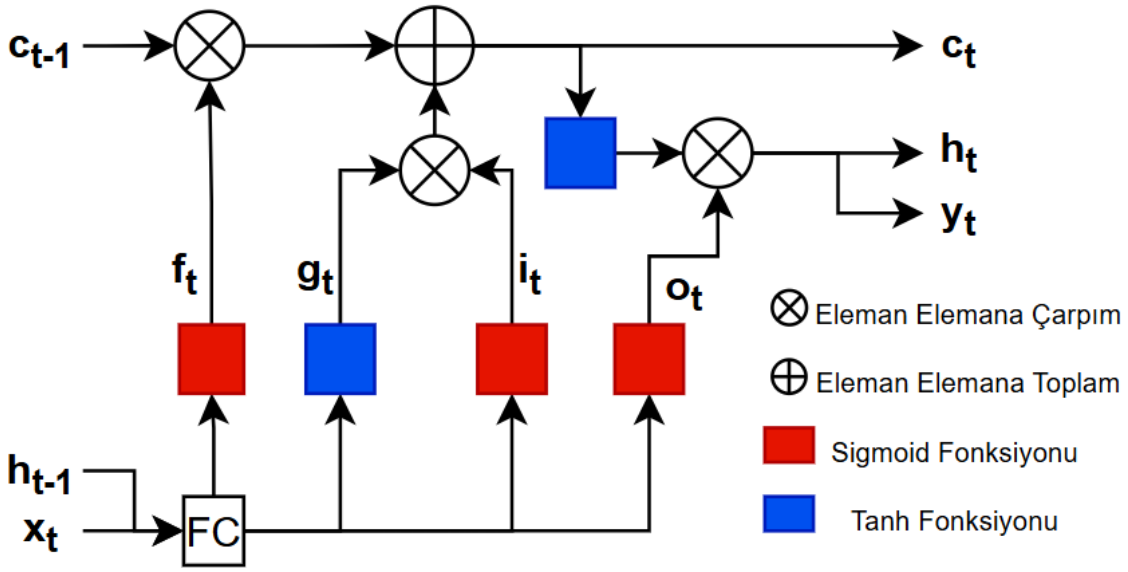
Sadece üç zaman damgası içeren bir veri kullanıldığı varsayılırsa son zaman damgasının gizli durum vektörü olan $\boldsymbol{h}_3$ hesaplanırken kullanılan denklem iç içe geçen sigmoid fonksiyonları içerecektir ve (4.14)'te verildiği gibi olacaktır.

$$\boldsymbol{h}_3 = \sigma(\boldsymbol{w}_{xh}\boldsymbol{x}_3 + \boldsymbol{w}_{hh}(\sigma(\boldsymbol{w}_{xh}\boldsymbol{x}_2 + \boldsymbol{w}_{hh}\sigma(\boldsymbol{w}_{xh}\boldsymbol{x}_1)))) \tag{4.14}$$

Ağırlıkları güncellemekte kullanılan hata bulunurken, hesaplanan tahminleri gösteren $\boldsymbol{y}_t$ çıktı vektörü, gerçekte olması gereken değerlerlerle karşılaştırılır. Dolayısıyla hatanın türevi $\boldsymbol{y}_t$ vektörüne göre alınır. Kullanılan $\boldsymbol{y}_t$ vektörü yalnızca $\boldsymbol{h}_t$ vektörüne bağlı olduğu (4.13)'te görülmektedir. Bu sebeple $\boldsymbol{y}_t$'nin türevi yerine $\boldsymbol{h}_t$'nin türevi alınabilir. (4.14)'te verilen denklemde de göründüğü üzere kullanılan zaman damgası sayısı arttıkça $\boldsymbol{h}_t$ vektörünü oluşturan iç içe geçen sigmoid fonksiyonları sayısı da artmaktadır. Bu sebeple istenen ağırlığa göre türevi alınmak istendiğinde bu türev iç içe geçen ve kullanılan zaman damgalarının sayısıyla beraber gittikçe uzayan zincire sahip olacaktır. Bu kısımda aktivasyon fonksiyonunun da zincirin her elemanı ile türevinin alındığı göz ardı edilmemelidir. Sigmoid gibi fonksiyonların türevleri 1'den düşük olduğu için bunların aktivasyon fonksiyonu olarak kullanımı, işlem esnasında türevin sürekli azalmasına ve en sonunda da sönümlenmesine yol açmaktadır. Böylece sonradan gelen zaman damgalarının istenen değişkenin güncellenmesine etkisi neredeyse kalmamakta ve sinir ağının buradaki verileri öğrenememesine sebep olmaktadır. Başka aktivasyon fonksiyonları da kullanılsa böyle bir durum olasılığını koruyacaktır [51].

4.4.1.1 LSTM

LSTM, RNN'deki gradient sorununa alternatif olarak geliştirilmiştir. Bu model, çok sayıda zaman damgası içeren veriler için daha uygundur. RNN'den farklı olarak, kendi içinde bir önceki zaman damgasının gizli durum vektörüne ek olarak hücre durumu vektörünü de aktarır [54]. Bir LSTM hücresinin içinde giriş, unutma ve çıkış olmak üzere üç kapı bulunmaktadır. LSTM hücresinin sembolik gösterimi Şekil 4.7'de verilmiştir.



Şekil 4.7 : LSTM hücresi sembolik gösterimi

Şekilde verilen değişkenlerin açıklaması bu kısımda yapılmıştır. Bütün girdi verileri, artan sırada zaman damgalarına göre sıralanıp eşit uzunluktan oluşan vektörler halinde sırayla LSTM hücresine girdi olarak verilirler. Bu verilerden t'ninci zaman damgasındaki girdi vektörü $\boldsymbol{x}_t$ ile gösterilir. Şekilde (t-1)'inci zaman damgasından gelen hücre durumu vektörü $\boldsymbol{c}_{t-1}$ ile gösterirken, t'ninci zaman damgasının hücre durumu çıktısı $\boldsymbol{c}_t$ ile gösterilmiştir. Benzer şekilde $\boldsymbol{h}_{t-1}$ (t-1)'inci zaman damgasından gelen gizli durumu vektörünü belirtirken $\boldsymbol{h}_t$ t'ninci zaman damgasında elde edilen gizli durum vektörünü göstermektedir. Hesaplanan t'ninci zaman damgasındaki unutma kapısının çıktı vektörü $\boldsymbol{f}_t$, giriş kapısının çıktı vektörü $\boldsymbol{i}_t$ ve bu vektör ile çapılarak giriş kapısının etkisini ayarlamakta kullanılan vektör $\boldsymbol{g}_t$, çıkış kapısının çıktı vektörü $\boldsymbol{o}_t$ ile gösterilmiştir. FC simgesi full-connected (bütünüyle bağlanmış) demektir ve $\boldsymbol{x}_t$ ile $\boldsymbol{h}_{t-1}$ vektörlerinin art arda eklenip bir vektör oluşturulmasını ifade etmektedir. Kırmızı renkle gösterilen logistic olarak isimlendirilen sembol sigmoid fonksiyonunu gösterirken mavi renkte gösterilen simge *tanh* aktivasyon fonksiyonunu göstermektedir. Çarpı simgesi gelen matrislerin eleman elemana çarpılacağını, toplam sembolü ise toplanacağını göstermektedir.

Unutma kapısında, bir önceki zaman damgasından mevcut zaman damgasına girdi olarak gelen hücre durumunun unutulup unutulmayacağına karar verilir. Burada oluşturulmuş girdi vektörü, $\boldsymbol{w}_f$ ağırlık matrisine tabi tutulur. Elde edilen değer b_f bias değeri ile toplanarak sigmoid fonksiyonuna girdi olarak verilir. Böylece $\boldsymbol{f}_t$ çıktı vektörü elde edilir ve bu değer önceki hücre durumu vektörü ile çarpılır. Yapılan işlemlerin denklemi (4.15)'te verilmiştir.

$$\boldsymbol{f}_t = \sigma(\boldsymbol{w}_f[\boldsymbol{h}_{t-1},\ \boldsymbol{x}_t] + \ b_f) \tag{4.15}$$

Giriş kapısında, gelen girdi değerleri ile önceki hücreden alınan gizli durum bilgileri kullanılarak yeni hücre durumuna eklenecek değerlerin belirlenmesi yapılır. Öncelikle önceki hücreden gelen gizli durum bilgisi ile mevcut girdiler birleştirilerek $\boldsymbol{w}_i$ ağırlık vektörü ile çarpılır. Elde edilen sonuç, b_i bias değeri ile toplanarak sigmoid fonksiyonuna tabi tutulur. Bu işlem sonucu $\boldsymbol{i}_t$ elde edilir. Yapılan işlem (4.16)'da verilmiştir. Sonrasında, (4.17)'de göründüğü üzere, yine önceden kullanılan birleştirilmiş veriler bu sefer $\boldsymbol{w}_g$ ağırlık vektörü ile çarpılıp b_g bias değeri ile toplanarak *tanh* fonksiyonuna girdi olarak verilir. Bu işlemin sonucunda ise $\boldsymbol{g}_t$ vektörü elde edilir.

$$\boldsymbol{i}_t = \ \sigma\ (\boldsymbol{w}_i[\boldsymbol{h}_{t-1},\ \boldsymbol{x}_t] + \ b_i) \tag{4.16}$$

$$\boldsymbol{g}_t = tanh(\boldsymbol{w}_g[\boldsymbol{h}_{t-1},\ \boldsymbol{x}_t] + b_g) \tag{4.17}$$

Bu noktada daha önce elde edilen unutma kapısının çıktı vektörü ile önceki hücreden gelen hücre durum bilgisi vektörü çarpılır. Bulunan değere giriş kapısında elde edilen iki vektörün birbiriyle çarpımı eklenerek yeni hücre durumu (4.18)'de göründüğü gibi elde edilir. İşlemdeki $\odot$ eleman elemana çarpımı sembollemektedir.

$$\boldsymbol{c}_t = \ \boldsymbol{f}_t \odot \boldsymbol{c}_{t-1} + \ \boldsymbol{i}_t \odot \boldsymbol{g}_t \tag{4.18}$$

Çıkış kapısında oluşturulmuş girdi vektörü $\boldsymbol{w}_o$ ağırlık vektörü ile çarpılır ve b_o bias değeri ile toplanır. Elde edilen toplam sigmoid fonksiyonuna (4.19)'da verildiği gibi tabi tutulur. Elde edilen $\boldsymbol{o}_t$ vektörü ile henüz hesaplanmış olan mevcut zaman damgasının hücre durumu vektörünün *tanh* fonksiyonuna tabi tutulmasının çıktısı çarpılarak mevcut zaman damgasının gizli durumu vektörü elde edilir. Sürecin denklemi (4.20)'de verilmiştir.

$$\boldsymbol{o}_t = \ \sigma(\boldsymbol{w}_o[\boldsymbol{h}_{t-1},\ \boldsymbol{x}_t] + \ b_o) \tag{4.19}$$

$$\boldsymbol{h}_t = \ \boldsymbol{o}_t tanh(\boldsymbol{c}_t) \tag{4.20}$$

Bu noktada önceden bahsedilen RNN gradient probleminin LSTM kullanılarak nasıl çözüldüğü gösterilebilir. Eğer (4.18)'de hesaplanan hücre durumunun bir önceki hücre durumuna göre türevi alınırsa yalnızca unutma kapısının sonuç vektörü (4.21)'de göründüğü gibi elde edilecektir.

$$\frac{\partial \boldsymbol{c}_t}{\partial \boldsymbol{c}_{t-1}} = \boldsymbol{f}_t \tag{4.21}$$

Buradan sonuçla eğer $\boldsymbol{f}_t$ vektöründeki değerler 1 ise gradientte bir azalma olmayacağı görülmektedir. Öğrenmenin başlangıcında $\boldsymbol{f}_t$ vektörü değeri hesaplanırken kullanılan b_f bias değeri sonucun 1 çıkması için kasten büyük alınır. Bu da başlangıçtaki bütün zaman damgalarının girdi değerlerinin hatırlanmasını sağlar. Öğrenmenin devamında ilerleyen zaman damgalarının kullanımı ile unutma kapısı, önceki zaman damgalarının etkilerinin unutulmasına başlar [51].

5. UYGULAMA

Önceki bölümlerde de belirtildiği gibi, YSA içerisinde birçok sistematik model literatürde mevcuttur ve her modelin kendine özgü avantajları ve dezavantajları bulunmaktadır. Birçok değişkene bağlı olarak yapılan rastgele yürüyüş tahminlerinde, herhangi bir modelin her durumda diğerlerinden daha üstün olduğunu söylemek mümkün değildir. Her model, çeşitli parametrelerden oluşur ve bu parametrelerin doğru seçilmesi, analiz edilen problemin yapısına göre büyük önem taşır. Bu nedenle, bir probleme optimum çözüm bulabilmek için farklı model ve parametrelerin denenmesi gerekmektedir.

Bu bölümde, seçilen model ve parametrelerin çıktının doğruluğu üzerindeki etkileri karşılaştırmalı olarak ele alınmıştır. Bu etkileri daha net bir şekilde gözlemlemek için iki ayrı örnek üzerinde denemeler yapılarak performans bakımından karşılaştırma yapılmıştır. Birinci örnekte, BIST'te işlem gören dört farklı hisse senedinin gelecekteki gün sonu kapanış değerleri, iki gizli katmandan oluşan MLP kullanılarak tahmin edilmeye çalışılmıştır. Bu gizli katmanlarda, ReLU ve Sigmoid aktivasyon fonksiyonları taşıyan farklı sayıda düğümler kullanılarak çeşitli denemeler yapılmıştır. Böylece farklı aktivasyon fonksiyonları ve düğüm sayılarının gelecek değeri tahmin etme performansları karşılaştırılmıştır. Bu örnekte kullanılan veriler, günlük frekansa sahip 8 farklı değişkenden oluşmaktadır.

İkinci örnekte ise NASDAQ'tan alınan 7 farklı hisse senedinin, farklı zaman periyotları sonundaki değerlerini tahmin etmek için hem LSTM hem de MLP modelleri ile denemeler yapılmıştır. Bu örnekte farklı frekanslara sahip 23 değişkenlerden oluşan veri kullanılarak farklı modellerin hisselerin farklı zaman çerçeveleri sonunda alacakları değeri tahmin etme performansları karşılaştırılmıştır. Farklı zaman çerçeveleri, ayrık iş günlerinden oluşan zaman serisi verileri farklı sayılarda günler halinde gruplandırılarak oluşturulmuştur. Bu iki örnek üzerinde yapılan çalışmanın detayları bu bölümde açıklanmıştır.

5.1 MLP ile Farklı Düğüm Sayıları ve Aktivasyon Fonksiyonlarının Denenmesi

Daha önce de belirtildiği gibi, sinir ağındaki katmanlarda bulunan düğüm sayıları ve aktivasyon fonksiyonları oldukça önemlidir. Genellikle her bir kombinasyonun ayrı ayrı denenip en iyi sonuç verenin seçilmesiyle bu parametrelerin doğru bir şekilde belirlenmesi

sağlanabilir. Bu örnekte, SOKM, OYAKC, ECILC ve VAKBN kodlu dört BIST hisse senedinin gelecek gün sonu değerleri, farklı düğüm sayıları ve aktivasyon fonksiyonları içeren MLP ile tahmin edilmeye çalışılmıştır. Bu tahminler yapıldıktan sonra sonuçlar yorumlanarak, bu parametrelerin sonuçlar üzerindeki etkileri açıklanmak istenmiştir. Bu dört hisse, BIST içerisinden rastgele seçilmiştir. Veriler, 7 Nisan 2022 ile 4 Nisan 2024 arasındaki 500 iş gününe ait olup ve İş Bankası Yatırım'ın web sitesinden alınmıştır [37]. Bu veriler, günlük frekansa sahip 8 farklı değişkenden oluşmaktaydı. Bu değişkenler Çizelge 5.1'de verilmiştir.

Çizelge 5.1 : Farklı aktivasyon fonksiyonları ve düğüm sayıları içeren MLP ile yapılan denemelerde kullanılan verilerin değişkenleri ve tanımları

DEĞİŞKEN İSMİ	TANIMI
Hisse Kapanış Değeri	İlgili borsa iş günü sonundaki hissenin kapanış değeri
Hisse En Düşük Değeri	İlgili borsa iş günü içerisinde hissenin aldığı en düşük değer
Hisse En Yüksek Değeri	İlgili borsa iş günü içerisinde hissenin aldığı en büyük değer
Ağırlıklı Ortalama Fiyat	$\frac{\sum işlem\ fiyatı\ x\ işlem\ miktarı}{Toplam\ İşlem\ Sayısı}$
Hacim	$\sum işlem\ fiyatı\ x\ işlem\ miktarı$
USD/TL Oranı	1 USD'nin Kaç TL Olduğunu Gösterir
BIST100 Endeks	Türkiye'nin En Önemli Hisse Senedi Endeksidir
Piyasa Değeri (USD)	(Şirketin Borsadaki Aktif İşlem Gören Toplam Hisse Sayısı) x (Hisse Senedi Fiyatı)

Bu değişkenler, önceki kısımlarda bahsedilmiş olan teknik analizde kullanılan değişkenlerin büyük kısmını kapsamaktadır. Bununla birlikte bazı önemli makroekonomik ve mikroekonomik faktörleri de içermektedir [4, 32-34]. Ağırlıklı ortalama fiyat, belirli bir hisse senedinin işlem gördüğü iş günü içinde yapılan tüm işlemlerin fiyatlarının, her bir işlemdeki toplam işlem miktarıyla çarpılarak elde edilen değerlerinin toplamının, o günkü toplam işlem miktarına bölünmesiyle hesaplanır. Diğer bir değişken olan hacim, ağırlıklı ortalama fiyatın toplam işlem miktarıyla çarpılmasıyla elde edilir. Diğer değişkenlerin açıklamaları üstteki çizelgede açıkça yapılmıştır.

Her hisse verisinin son 10 günü test için, geri kalan 490 günü ise eğitim amacıyla kullanılmıştır. Hisse senedi tahminlerinde sıklıkla kullanılan OOS yönteminin örnek alınmış olması nedeniyle verilerin bu biçimde ayrılması tercih edilmiştir. Hisse senedi verileri arasında tarihsel bir bağlantı bulunduğundan, K-fold çapraz doğrulama gibi tekniklerin kullanımı sınırlıdır. Çünkü tarihsel bütünlük bozulursa, önemli değişikliklerin sinir ağı

tarafından öğrenilmesi engellenmiş olur. Bu durumu daha iyi açıklamak için, sürekli artacak şekilde sıralanmış zaman damgalarından oluşan bir veriyi, eşit beş dilime böldüğümüzü varsayalım. İlk iki ve son iki dilimin eğitimde, ortadaki dilimin ise testte kullanıldığını farz edelim. Ortadaki dilimde, çok önemli bir politik olay veya şirketle ilgili büyük bir gelişme yaşanmış olabilir. Bu durumda, sinir ağı bu olayın yaşandığı kısmı eğitiminde kullanamadığından öğrenemeyecektir. Bu durum, ilgili eğitimden sonra elde edilecek tahminlerin tutarsız olmasına sebep verecektir. Bu, doğrusal olmayan diğer zaman serileri için de geçerlidir. Bu nedenle, bu tür serilerin eğitimi yapılırken verinin bütünlüğü korunarak, OOS yöntemiyle sadece serinin son kısmındaki birkaç zaman damgasından oluşan küçük bir dilim ayrılarak işlem yapılır. Test edilecek zaman damgalarının sayısı arttıkça önemli olayların meydana gelme olasılığı yükseldiğinden, çok fazla sayıda ayrılması tercih edilmemektedir. Diğer yandan, çok az sayıda zaman damgasıyla test yapmak ise gürültü seviyesini arttırmaktadır. Bu örnekte gürültüden en az derecede etkilenerek analiz yapılmasını sağlamak ve önemli olayların eğitim verisi dışında kalmasının önüne geçmek için son 10 zaman damgası test için ayrılmıştır [55].

Bütün veriler, 0 ile 1 arasında normalize edilmiştir. Bu işlemin yapılma amacı, önceki bölümlerde de açıklandığı gibi, veriler arasındaki ölçü farklarından kaynaklanabilecek hataları engellemek ve işlemleri daha verimli hale getirmektir. Ayrıca, kullanılan sigmoid aktivasyon fonksiyonunun belirli bir çalışma aralığına sahip olması nedeniyle bu işlem zorunlu hale gelmektedir [44]. Normalize etme işleminin diğer faydaları ise önceki bölümlerde detaylı bir şekilde açıklanmıştır.

Her biri 2 ile 16 arasında değişen algılayıcıya sahip iki gizli katmandan oluşan MLP, ReLU ve sigmoid aktivasyon fonksiyonları kullanılarak her hisse için ayrı ayrı eğitilip test edilmiştir. Toplam olarak 15*15*2=450 farklı senaryo test edilmiştir. Eğitilen model, ilgili hisseye ait hem eğitim hem de test verileriyle beslenerek tahminler elde edilmiştir. Eğitim ayarlarında maksimum devir sayısı 250, başlangıç öğrenme katsayısı 0.1, mini-batch büyüklüğü 16 ve optimizasyon algoritması olarak ADAM tercih edilmiştir.

Elde edilen tahminler, gerçek değerlerle MAPE ve EDOC metrikleri kullanılarak karşılaştırılmıştır. MAPE denklemi ise (5.1)'de verilmiştir [51].

$$MAPE = 100 * \frac{1}{N} \sum_{i=1}^{N} \frac{|A_i - P_i|}{A_i} \tag{5.1}$$

Bu denklemde, A_i zaman serisinin i'nci sırasındaki gerçek (actual) değeri, P_i ise o değere yapılan tahmin (prediction) ifade eder. Elde edilen tahminlerin gerçek değerden

farklarının, yine o gerçek değere oranları alınıp toplanır. Eğer toplamda *N* tahmin yapıldığı varsayılırsa, bu toplam *N*'e bölünerek MAPE elde edilir. MAPE, tahminlerin gerçek değerlere ne kadar yakın olduğunu ölçmek için kullanılır.

EDOC, tahmin veya gerçek değerlerin bulundukları zaman damgasındaki yönünü, bir önceki zaman damgasında gerçekleşen değere göre sınıflandırmak için kullanılır. Bu sınıflandırma, değişim oranı kullanılarak yapılmıştır. Değişim oranı, (5.2)'de verilen denkleme göre hesaplanmıştır.

$$D_t = \frac{A_t - A_{t-1}}{A_{t-1}} \tag{5.2}$$

Bu denklemde D_t, zaman serisinin t'ninci sırasındaki değişim oranını göstermektedir. Denklemdeki diğer elemanlar olan A_t, t'ninci sıradaki zaman damgasının tahmin edilen veya gerçek değerini, A_{t-1} ise (t-1)'inci sıradaki zaman damgasındaki gerçekleşen değeri ifade etmektedir. Hesaplanan değişim oranı kullanılarak, (5.3)'te gösterilen şekilde EDOC sınıflandırması yapılmıştır.

$$EDOC = \begin{Bmatrix} Artış, & D_t > 0.001 \\ Azalış, & D_t < -0.001 \\ Değişmemiş, & Else \end{Bmatrix} \tag{5.3}$$

Bu çalışmada, üç EDOC sınıfı olduğu varsayılmıştır. Değişim miktarı 0.001'in üzerinde ise 'Artış' (increase), -0.001'in altında ise 'Azalış' (decrease), bu iki değer arasında ise 'Değişmemiş' (no change) sınıfına atanmıştır. Böylece, hissenin bir önceki zaman damgasına göre değişim yönü sınıflandırılmıştır. Gerçekleşen değişim yönü ile tahmin edilen değişim yönleri karşılaştırılarak, kullanılan sistemin doğruluğu değerlendirilmiştir.

Her deneme için doğru tahmin edilen EDOC sayısı, o denemedeki toplam tahmin sayısına bölünerek doğruluk oranı hesaplanmıştır. Bu oran ve MAPE değerleri kullanılarak tüm denemeler sıralanmış ve en iyi performans gösteren deneme seçilmiştir. Seçilen denemede kullanılan parametrelerle oluşturulan modelin her hisse üzerindeki sonuçları kullanılarak karışıklık (confusion) matrisleri oluşturulmuştur. Bu matrisler kullanılarak kesinlik, duyarlılık, doğruluk ve F1 skoru hesaplanmıştır [51, 56].

5.2 MLP ve LSTM'in Değişen Zaman Çerçevelerinde Farklı Frekanstaki Değişkenlerden Oluşan Veriler Kullanılarak Denenmesi

YSA içinde birçok farklı model bulunduğu daha önce belirtilmiştir. MLP, tahmin üretirken yalnızca mevcut zaman damgasındaki değişkenleri dikkate alarak çalıştığı için

geçmiş zaman damgaları göz önünde bulunduramaz. Buna karşın, RNN önceki zaman damgalarına ait verileri de kullanarak tahminler üretir. Bu nedenle, RNN'nin zaman serisi tahminlerinde daha üstün bir performans sergileyebileceği düşünülmektedir. Ancak, bu her durumda geçerli olmayabilir, çünkü geçmiş veriler her zaman mevcut zaman damgasındaki sonucu etkilemeyebilir. Bu nedenle, önceki örnekte olduğu gibi yeni denemeler yapılması gerekmektedir. Bu örnekte, MLP'ye ek olarak daha önce detaylıca açıklanan bir tür RNN olan LSTM de kullanılmıştır. Ayrıca, farklı zaman çerçevelerinde tahminlerde nasıl bir sonuç elde edileceği araştırılmıştır. İlgili hisse senetleri hakkında daha fazla bilgi edinmek amacıyla, farklı frekanslara sahip değişkenler de kullanılmıştır.

İkinci örnekte, ilkinde kullanılan hisselerden ve borsadan farklı olarak, YSA'nın diğer hisseler üzerindeki genellenebilirliğini gözlemlemek amacıyla NASDAQ'ta işlem gören AAPL, AMD, AMZN, GOOGL, JNJ, MSFT ve NVDA kodlu hisselerin gelecek değerleri tahmin edilmeye çalışılmıştır. Bu hisseler rastgele seçilmiştir ve veriler 30 Haziran 2009'dan 3 Nisan 2024'e kadar olan borsa iş günlerini kapsamaktadır. Önceki örnekten farklı olarak, bu örnekte çeşitli frekansta etkenler kullanılmak suretiyle, ilgili hisseler hakkında daha fazla bilgi edinilmeye çalışılmıştır. Kullanılan değişkenler, 3 aylık, aylık, haftalık ve günlük frekansta sınıflandırılmıştır. Bu etkenler makroekonomik, mikroekonomik ve günlük hisse verileri olarak üç ana gruba ayrılabilir [4, 32-34].

Mikroekonomik etkenler arasında ciro, brüt kâr, net gelir, FAVÖK, FVÖK, hisse başına kazanç, toplam varlıklar, toplam yükümlülükler, özkaynaklar, net nakit akışı ve hisse senedine dayalı tazminatlar yer almakta olup her şirketin çeyreksel raporlarından alınmıştır. Ekonomide çeyrek bir yılı oluşturan on iki ayın her üç aylık dilimine denir. Bu sebeple çalışmada kullanılan mikroekonomik etkenlerin frekansı 3 aydır [32]. Günlük hisse verileri ise açılış, en yüksek, en düşük, kapanış, ayarlanmış kapanış fiyatları ve hacmi içermektedir, ve bu veriler günlük frekansa sahiptir. Makroekonomik etkenler ise GSYİH, Amerikan Merkez Bankası'nın aylık faiz oranları, altın fiyatları, benzin fiyatları, ABD Doları endeksi ve NASDAQ endeksi gibi göstergeleri içermektedir. Bu etkenlerin frekansı ise Çizelge 5.2'de verilmiştir.

Çizelge 5.2 : LSTM ve MLP'nin farklı zaman çerçevelerinde denenmelerinde kullanılan verilerin makroekonomik değişkenleri ve frekansları

Değişken İsmi	**Frekans**
GSYİH	Çeyrek Dönem
Faiz Oranı	Aylık
Altın Fiyatı	Haftalık
Benzin Fiyatı	Günlük
Amerikan Dolar Endeksi	Günlük
NASDAQ Endeksi	Günlük

Toplam olarak 23 değişken kullanılmıştır. Farklı frekansa sahip bu değişkenlerin tamamının günlük frekansa uyarlanabilmesi için, yüksek frekansa sahip değişkenlerin, yayımlandıkları iş gününden bir sonraki yayımlanma gününe kadar olan tüm günlerde aynı değeri aldığı varsayılmıştır. Bu işlem sonucunda tüm değişkenler günlük frekansa göre tekrar eden verilere dönüştürülmüştür. Elde edilen bu günlük zaman serisi verisi, 2, 3, 4 ve 5 günlük gruplar halinde yeniden düzenlenerek, farklı zaman çerçevelerindeki tahminler de üretilmeye çalışılmıştır. Gruplandırma işlemi sırasında, önce bir günlük toplam veri sayısı istenen grup büyüklüğüne bölünmüş ve kalan gün sayısı, verilerin başından çıkartılarak yeniden düzenlenmiştir. Bu işlem sonucunda, her grup için istenen gün sayısına göre gruplandırılmış yeni bir seri elde edilmiştir.

Gruplandırma işlemi sırasında, her grubun ilk günündeki makroekonomik ve mikroekonomik değişkenler olduğu gibi alınmıştır. Günlük hisse verileri içinse, her grubun ilk gününün açılış değeri aynen kullanılmıştır. Kapanış ve ayarlanmış kapanış değerleri, her grubun son günündeki verilerle eşleştirilmiş, en yüksek ve en düşük değerler ise grup içindeki tüm günlerin en yüksek ve en düşük değerlerinden seçilmiştir. Hacim, ilgili grup içindeki tüm hacimlerin toplamı olarak hesaplanmıştır. Böylece, istenen gün sayısına göre gruplanmış yeni zaman serileri elde edilmiş ve farklı zaman çerçevelerinde test edilmiştir.

Elde edilen gruplandırılmış zaman serileri, eğitim ve test verilerine ayrılırken önceki örnekte de kullanılan OOS yaklaşımı benimsenmiştir. Bu yöntemle 1, 2, 3, 4 ve 5 günlük gruplar veya zaman çerçeveleri için sırasıyla son 20, 10, 7, 5 ve 4 zaman damgası test seti olarak ayrılmış, geri kalan veriler ise eğitim seti olarak kullanılmıştır. Ayrıca, tüm veriler kullanılmadan önce 0 ile 1 arasına normalize edilmiştir.

Bu veriler, iki gizli katmana sahip MLP ve bir gizli katmana sahip LSTM modellerinin eğitilmesinde kullanılmıştır. MLP'nin katmanlarındaki algılayıcı sayıları sırasıyla 23-15-8-1, LSTM'nin ise 23-12-1 olarak belirlenmiştir. Gizli katmanlardaki düğüm

sayıları, yaklaşık hesap kuralına dayalı olarak seçilmiştir [41, 42]. MLP'de aktivasyon fonksiyonu olarak ReLU tercih edilmiş ve her iki modelde de kayıp fonksiyonu olarak MSE kullanılmıştır. ReLU'nun seçilmesinin nedeni gradyen kaybolma sorununa karşı daha dayanıklı olmasıdır. Eğitim süreci için ayarlanan parametreler arasında başlangıç öğrenme katsayısı 0.15, mini-batch büyüklüğü 32 ve optimizasyon algoritması olarak ADAM yer almıştır. En yüksek devir sayısı çok yüksek bir değere atanarak eğitim sürecinin takip edilmesine ve aşırı uyum olmadan elle sonlandırılmasına olanak tanınmıştır.

Her bir hisse için farklı zaman çerçevelerinde eğitim verileri kullanılarak, iki model ayrı ayrı eğitilmiştir. Eğitilen modeller, ardından o hisseye ait hem eğitim hem de test verileriyle beslenmiş ve tahminler elde edilmiştir. Bu tahminler, gerçek değerlerle önceki örnekte de kullanılan MAPE ve EDOC yöntemleriyle karşılaştırılmıştır. EDOC'un doğruluk değeri ise, (5.4)'te verilen denklemle hesaplanmış ve sonuçlar kısmında sunulmuştur.

$$\text{Doğruluk} = \frac{Toplam\ doğru\ tahmin\ edilen\ EDOC\ sayısı}{Toplam\ tahmin\ sayısı} \tag{5.4}$$

6. BULGULAR

Bu kısımda yapılan çalışmadaki iki örnek ile elde edilen sonuçlar sunulmaktadır.

6.1 MLP'nin Farklı Düğüm Sayıları ve Aktivasyon Fonksiyonları ile Denenmesinin Sonuçları

Borsada işlem gören hisse değerlerinin değişiminin yön tahmininin büyük bir öneme sahip olması nedeniyle, farklı aktivasyon fonksiyonları ve düğüm sayıları karşılaştırılırken EDOC doğruluğu dikkate alınmıştır. Eğitilen modellere hem eğitim verisi hem de test verisi verilerek tahminler yapılmış ve bu tahminler gerçek değerlerle karşılaştırılmıştır. BIST üzerinde yapılan denemelerde, genel olarak ReLU'nun sigmoid fonksiyonundan daha iyi sonuçlar verdiği gözlemlenmiştir. Bunun nedeninin, ReLU'nun türevinin sigmoidinkinden daha stabil olması ve bu özelliğin YSA'nın eğitimi sırasında gradiyent sönümünü engellemesi olduğu söylenebilir. Test verilerinin EDOC sonuçları, eğitim verilerine göre bazen daha yüksek bazen ise daha düşük olmuştur. Bunun sebebi, test verilerinin sayısının yalnızca 10, eğitim verilerinin ise 490 olmasıdır. Bu durum, eğitim verilerinin daha stabil sonuçlar verirken, test verilerinin sonuçlarının dalgalanmasına yol açmıştır.

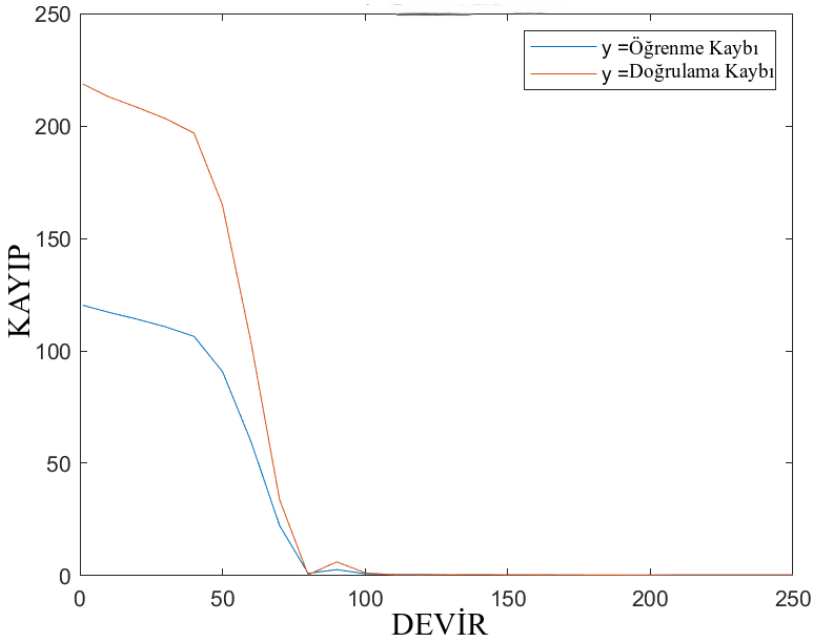
Bu örnekte, en iyi sonuç VAKBN hissesine ait test verisi ile elde edilmiş ve %90 EDOC doğruluğu sağlanmıştır. Diğer denemelerde ise %90 EDOC doğruluğu elde edilememiştir. Bu denemede kullanılan MLP modelinin katmanlarında sırasıyla ReLU aktivasyon fonksiyonu ile 8-5-3-1 sayıda algılayıcı bulunuyordu. Çizelge (6.1)'de, en iyi deneme ile elde edilen %80 doğruluğa sahip diğer bazı denemelerin katmanlarındaki algılayıcı sayıları ve MAPE ile EDOC doğruluk yüzdeleri verilmiştir. Diğer denemelerin EDOC doğrulukları %80'in altında olduğu için başarılı kabul edilmemiş ve tabloda yer almamıştır. MAPE değeri en düşük olan denemenin, aynı zamanda en yüksek EDOC doğruluğuna sahip olduğu gözlemlenmiş ve bu deneme çizelgede koyu renk ile işaretlenmiştir.

Çizelge 6.1 : Farklı aktivasyon fonksiyonları ve düğüm sayıları içeren MLP ile yapılan denemelerden alınan en iyi sonuçlar ve algılayıcı sayıları

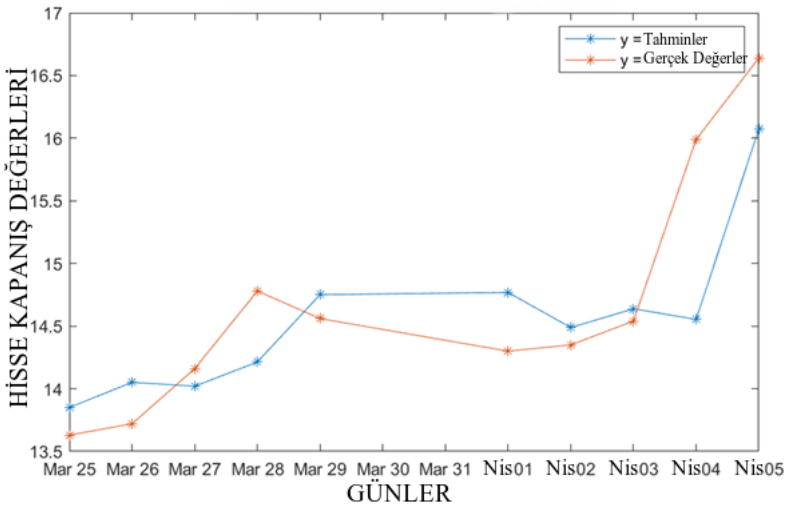
Katmanlardaki Algılayıcı Sayıları	TEST Verisinin MAPE'si	Yüzdesel EDOC Doğruluğu
8-5-3-1	**2.75**	**%90**
8-4-3-1	4.46	%80
8-6-3-1	3.47	%80
8-5-2-1	3.79	%80
8-5-4-1	4.95	%80
…	…	…

En iyi sonuçların alındığı denemede, düğüm sayılarının yaklaşık hesap kuralına uyduğu gözlemlenmiştir.

Şekil 6.1'de, en iyi sonucun öğrenme süreci gösterilmektedir. Grafiğin y ekseninde kayıp fonksiyonunun çıktısı yer alırken, x ekseninde ise öğrenme sürecindeki tur sayısı gösterilmektedir. Bu grafikte doğrulama kaybı, modelin aşırı öğrenme ya da yetersiz öğrenme gibi sorunlar yaşamadığını ortaya koymaktadır. Şekil 6.2'de ise test verilerindeki her bir iş günü için gerçek sonuçlarla tahminler karşılaştırılmaktadır. Grafikte, y ekseninde tahmin edilen ve gerçek değerler, x ekseninde ise günler yer almaktadır. Mavi çizgi üzerindeki yıldızlarla işaretlenen noktalar tahmin edilen değerleri, kırmızı çizgi ise gerçek değerleri göstermektedir. Buradan, tahminlerin gerçek değerlerden önemli bir sapma göstermediği net bir şekilde görülmektedir.



Şekil 6.1 : Farklı aktivasyon fonksiyonları ve düğüm sayıları içeren MLP ile yapılan denemelerin en iyisinin öğrenme süreci



Şekil 6.2 : Farklı aktivasyon fonksiyonları ve düğüm sayıları içeren MLP ile yapılan denemelerin en iyisinin tahminleri ve gerçek değerleri

Sonuçların yalnızca belirli bir aralıkta tutarlı olup olmadığını anlamak için, en iyi sonuçları veren parametrelerin diğer hisselerdeki test ve eğitim verileriyle elde edilen sonuçları EDOC çerçevesinde incelenmiştir. Bu inceleme sırasında her bir durumun sonuçları karışıklık matrisleri ile sunulmuştur.

Aşağıdaki çizelgeler, MLP modelinin en iyi sonuçları elde ettiği parametrelerle dört farklı hisse üzerindeki performansını göstermektedir. Çizelge 6.4'te, her hisse için sonuçlar karışıklık matrisleri şeklinde verilmiştir. Bu matrislerdeki P, A, I, N, D harfleri sırasıyla tahmin, gerçek, artış, değişmemiş ve azalış parametrelerini simgelemektedir. Bu karışıklık matrislerine ait kesinlik, duyarlılık, doğruluk ve F1 skoru verileri ise Çizelge 6.3'te yer almaktadır. Ayrıca, her hissenin test ve eğitim verileri üzerinden hesaplanan ortalama MAPE sonuçları Çizelge 6.2'de sunulmuştur.

Çizelge 6.2 : Farklı aktivasyon fonksiyonları ve düğüm sayıları içeren MLP ile yapılan denemelerin en iyisinin parametreleri ile farklı hisselerin test ve eğitim verilerinin denenmesinin MAPE sonuçları

HİSSE KODU	TEST VERİSİ MAPE	EĞİTİM VERİSİ MAPE
SOKM	4.29	3.78
OYAKC	9.01	5.84
ECILC	6.01	5.01
VAKBN	**2.75**	3.53

Çizelge 6.3 : Farklı aktivasyon fonksiyonları ve düğüm sayıları içeren MLP ile yapılan denemelerin en iyisinin parametreleri ile farklı hisselerin test ve eğitim verilerinin denenmesinin EDOC istatistikleri

HİSSE KODU		TEST VERİSİ			EĞİTİM VERİSİ		
		I	D	N	I	D	N
SOKM	Doğruluk	0.5	0.5	1	0.36	0.36	0.98
	Kesinlik	0.5	NaN	NaN	0.41	0.25	0
	Duyarlılık	1	0.5	NaN	0.47	0.22	0
	F1 Skoru	0.67	0	NaN	0.44	0.23	NaN
OYAKC	Doğruluk	0.5	0.5	1	0.36	0.36	0.99
	Kesinlik	0.5	NaN	NaN	0.39	0.31	NaN
	Duyarlılık	1	0.5	NaN	0.45	0.26	0
	F1 Skoru	0.67	0	NaN	0.42	0.28	NaN
ECILC	Doğruluk	0.3	0.3	1	0.34	0.34	0.99
	Kesinlik	0.3	NaN	NaN	0.40	0.25	NaN
	Duyarlılık	1	0	NaN	0.43	0.23	0
	F1 Skoru	0.46	NaN	NaN	0.42	0.24	NaN
VAKBN	Doğruluk	0.9	0.9	1	0.28	0.28	0.96
	Kesinlik	0.89	1	NaN	0.26	0.28	0
	Duyarlılık	1	0.5	NaN	0.26	0.28	0
	F1 Skoru	0.94	0.67	NaN	0.26	0.28	NaN
BÜTÜN HİSSELERİN ORTALAMASI	Doğruluk	0.55	0.55	1	0.33	0.34	0.98
	Kesinlik	0.54	1	NaN	0.37	0.27	0
	Duyarlılık	1	0.05	NaN	0.41	0.25	0
	F1 Skoru	0.7	0.1	NaN	0.39	0.26	NaN

Çizelge 6.4 : Farklı aktivasyon fonksiyonları ve düğüm sayıları içeren MLP ile yapılan denemelerin en iyisinin parametreleri ile farklı hisselerin test ve eğitim verilerinin denenmesinin karmaşıklık matrisleri

HİSSE KODU		EĞİTİM VERİSİ			TEST VERİSİ		
	P / A	I	N	D	I	N	D
SOKM	I	5	0	0	124	1	138
	N	0	0	0	6	0	3
	D	5	0	0	169	2	47
OYAKC	I	5	0	0	113	0	137
	N	0	0	0	3	0	0
	D	5	0	0	175	0	62
ECILC	I	3	0	0	115	0	150
	N	0	0	0	3	0	1
	D	7	0	0	171	0	50
VAKBN	I	8	0	0	61	3	172
	N	0	0	0	6	0	6
	D	1	0	1	170	4	68
BÜTÜN HİSSELERİN TOPLAMI	I	21	0	0	413	4	597
	N	0	0	0	18	0	10
	D	18	0	1	685	6	227

Eğitim verileriyle test edilen MLP'nin, hisselerin artıp artmayacağına dair yön tahminindeki %EDOC doğruluğu, farklı hisselerde benzer bir seviyeye sahip olup, genellikle %50'nin altında kalmıştır. Ancak test verileriyle yapılan denemelerde, doğruluk oranları %30 ile %90 arasında değişmiştir. Bu durum, daha önce açıklandığı gibi, kullanılan test veri miktarının az olmasından kaynaklanmaktadır. Bu nedenle, yön tahminlerinin rastgele yürüyüş hipoteziyle uyumlu olarak başarısız olduğu söylenebilir.

Hisselerin önceki değerlerine göre değişmeyeceklerine dair tahminlerin doğruluk oranı ise hem test hem de eğitim verileriyle neredeyse %100'dür. Ancak, değişim olmama durumu hiç doğru tahmin edilememiştir. Bunun nedeni, değişim olmama durumunun çok nadir olması ve bu durumun tahmin edilmesinin diğer durumlara kıyasla çok daha az sayıda gerçekleşmesidir. Bu yüzden doğruluk oranı yüksek görünse de, değişim olmama durumu aslında doğru tahmin edilememiştir.

Sonuç olarak, ilk örnekte farklı aktivasyon fonksiyonları ve düğüm sayılarıyla elde edilen sonuçlar, parametrelerin değişiminin farklı sonuçlara yol açtığını göstermektedir. Yön tahmini bazı durumlarda %90 doğruluğa ulaşsa da, genel olarak %50'nin altında kalmıştır, bu da yön tahmininde başarılı olunmadığını göstermektedir. MAPE sonuçları ise EDOC doğruluğuna göre daha tutarlı olmuştur. MAPE açısından, VAKBN test verileri %2,75 doğrulukla en iyi performansı gösterirken, OYAKC test verileri %9,01 doğrulukla en düşük performansı sergilemiştir. Ayrıca, ReLU aktivasyon fonksiyonunun sigmoid'e göre daha başarılı olduğu gözlemlenmiştir. Düğüm sayıları ise yaygın olarak kullanılan yaklaşık hesaplama kuralıyla uyumlu sonuçlar vermiştir.

6.2 MLP ve LSTM'in Değişen Zaman Çerçevelerinde Farklı Frekanstaki Değişkenlerden Oluşan Veriler Kullanılarak Denenmesin Sonuçları

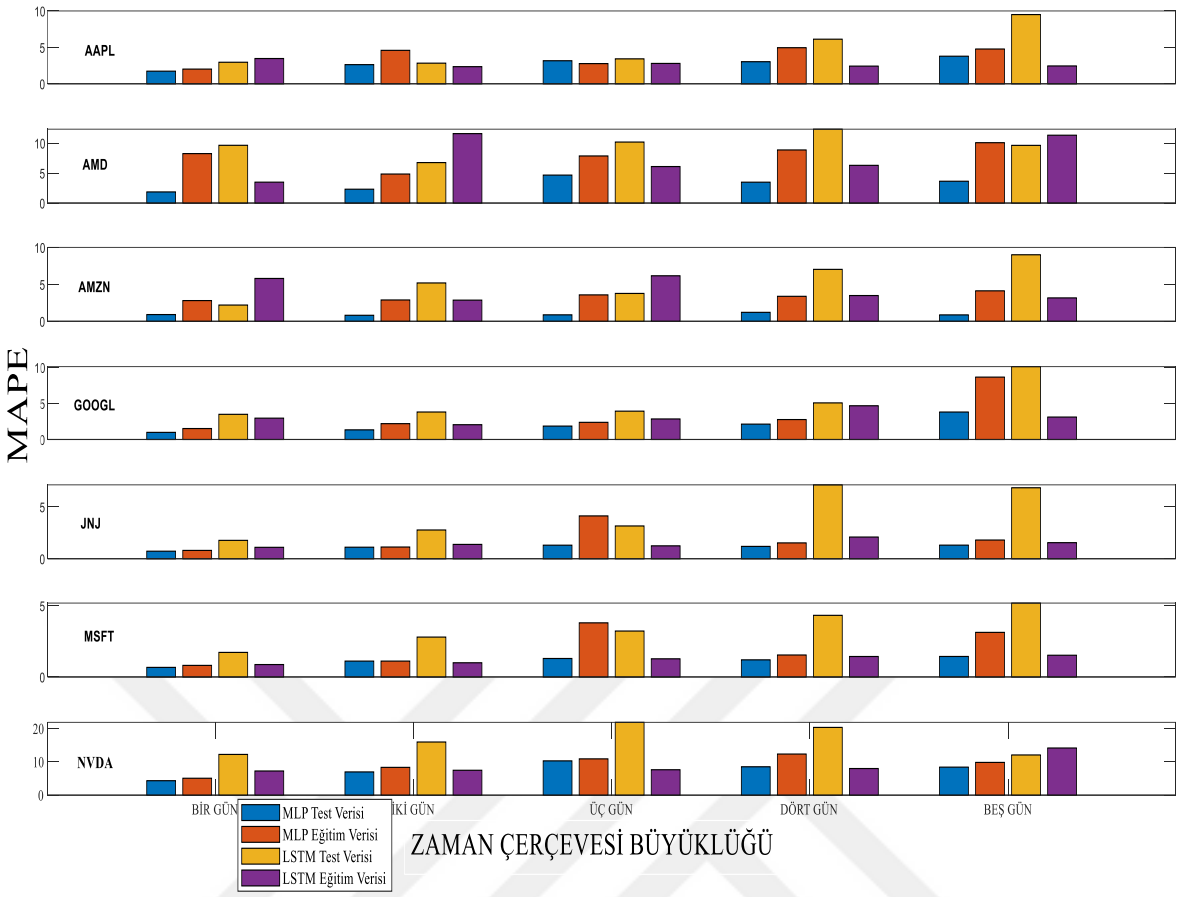
İlk örnekte, farklı aktivasyon fonksiyonları ve düğüm sayıları içeren MLP ile yapılan denemelerde hisselerin yönü başarılı bir şekilde tahmin edilememiştir. Bu başarısızlığın nedeni, rastgele yürüyüş teorisinde de belirtildiği gibi, hisselerin kısa vadede yüksek gürültülü hareketler göstermesi olarak değerlendirilmiştir. Bu durumu çözmek amacıyla, kısa dönem tahmini yerine uzun dönem tahminleri tercih edilerek farklı zaman çerçevelerinde denemeler yapılmıştır. Bu kapsamda, ertesi günün tahminine ek olarak, hisselerin iki, üç, dört ve beş gün sonraki değerleri de tahmin edilip karşılaştırılmıştır. Ayrıca, önceki örnekten farklı olarak yalnızca MLP kullanılmamış, LSTM de kullanılarak

tahminler üretilmiş ve karşılaştırılmıştır. Bu örnekte, önceki örnekte kullanılan günlük frekansa sahip 8 değişken yerine farklı frekanslar içeren 23 değişken kullanılmıştır.

Bu örnekte kullanılan hisseler, önceki örnekte hisselerin alındığı BIST yerine NASDAQ borsasından seçilmiştir. Böylece, önerilen yöntemin farklı borsalar ve hisseler üzerinde uygulanabilirliği de test edilmiştir. Tahminler yine MAPE ve EDOC ile karşılaştırılmış, ancak EDOC doğruluğuna öncelik verilmiştir.

İlk örnekten alınan sonuçlara dayanarak, ReLU aktivasyon fonksiyonu tercih edilmiştir ve düğüm sayıları yaygın kullanılan yaklaşık hesap kuralına göre belirlenmiştir. Bu örnekte de iki gizli katmana sahip MLP kullanılmıştır. LSTM ise tek gizli katmana sahip olacak şekilde ayarlanmıştır. Bunun seçimin sebebi, LSTM hücrelerinin karmaşıklığının MLP düğümlerinden daha fazla olması ve bu durumun gizli katman sayısı arttıkça modelin aşırı öğrenmesine sebep olabilmesidir. LSTM modelindeki gizli katmanda kullanılan hücre sayısı, giriş ve çıkış katmanlarındaki düğüm sayılarının ortalaması olan 12 olarak belirlenmiştir. Bu sayı, fazla hücre kullanıldığında aşırı öğrenme, az hücre kullanıldığında ise yetersiz öğrenme yaşanmaması için dikkatlice seçilmiştir. Bu nedenle, ortalama bir değer tercih edilmiştir [57].

İkinci örnekte kullanılan NASDAQ'tan seçilen 7 hisse senedine ait MAPE değerleri, farklı hisse senetleri ve zaman çerçeveleri için Şekil 6.3'te gösterilmiştir. Bu grafiklerde x ekseni bir, iki, üç, dört ve beş günlük zaman çerçevelerini, y ekseni ise MAPE değerlerini göstermektedir. Test ve eğitim verileriyle denenen MLP ve LSTM modellerinin sonuçları her zaman çerçevesi ve hisse için dört farklı sütun halinde sunulmuştur ve her bir durum farklı renklerle belirtilmiştir.



Şekil 6.3 : LSTM ve MLP'nin farklı hisseler için test ve eğitim verileri ile farklı zaman çerçevelerindeki denenmelerinin MAPE sonuçları

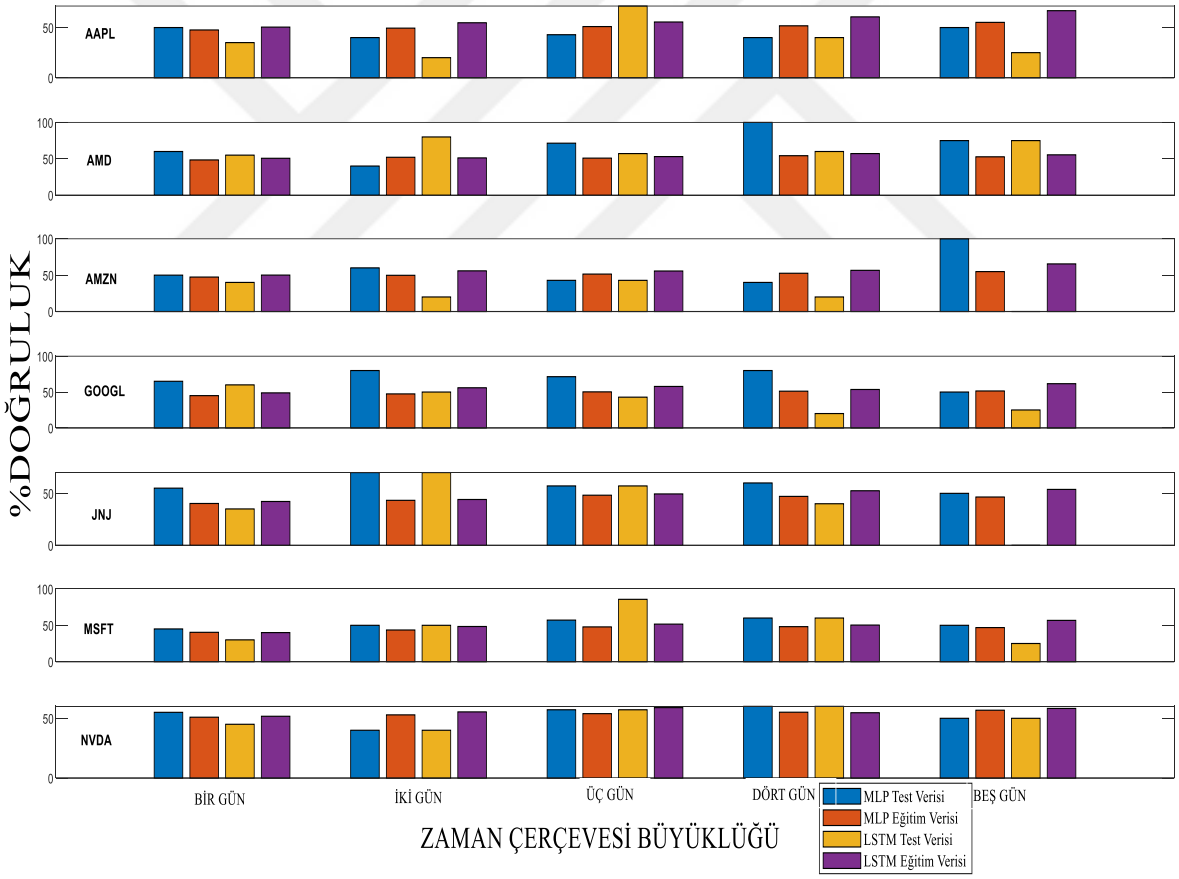
Genel olarak, zaman çerçeveleri arttıkça MAPE değerlerinin de arttığı gözlemlenmiştir, ancak bu artış çok fazla değildir. Farklı zaman çerçevelerinde MAPE değerlerinde en fazla değişim, GOOGL hissesinde görülmüştür. En yüksek MAPE değerlerinin %20'nin biraz üzerinde olduğu ve NVDA hissesinin üç ve dört günlük zaman çerçevelerinde LSTM ile yapılan denemelerde elde edildiği tespit edilmiştir. Hisseler arasında ortalama MAPE değerlerine bakıldığında, en yüksek değerlere sahip olanlar AMD ve NVDA, en düşük değerlere sahip olanlar ise JNJ ve MSFT'dir. Bütün hisse senetleri için yapılan tüm denemelerin ortalama MAPE değerlerine bakıldığında, yalnızca NVDA'nın %10,61 ile %10'un üzerinde bir değeri olduğu görülmüştür. LSTM ve MLP'nin test ve eğitim verileriyle yapılan denemeler karşılaştırıldığında, en yüksek MAPE değerlerinin test verileriyle denenen LSTM modelinde olduğu gözlemlenmiştir. Eğitim verileriyle yapılan LSTM ve MLP denemelerinin MAPE sonuçları ise genellikle birbirine yakın ve çok değişken olmamıştır.

Tüm hisselere ait ortalama MAPE değerleri Çizelge 6.5'te verilmiştir.

Çizelge 6.5 : LSTM ve MLP'nin bütün hisselerin test ve eğitim verileri ile farklı zaman çerçevelerindeki denenmelerinin ortalama MAPE sonuçları

	MLP		LSTM	
Zaman Çerçevesi	**Test Verisi**	**Eğitim Verisi**	**Test Verisi**	**Eğitim Verisi**
Bir Günlük	1.60	3.04	4.86	3.56
İki Günlük	2.33	3.58	5.73	4.10
Üç Günlük	3.35	5.06	7.08	4.01
Dört Günlük	2.97	5.05	8.91	4.06
Beş Günlük	3.32	6.05	8.90	5.32

Her hisse senedinin EDOC doğruluk sonuçları Şekil 6.4'te sunulmuştur. Bu grafikte, MAPE sonuçlarındaki gibi x ekseni zaman çerçevelerini artan gün sayısıyla göstermektedir. Ancak bu kez, y ekseni EDOC doğruluğunun yüzde oranını göstermektedir. Eğitim ve test verileri yine ayrı ayrı verilmiş olup, her bir veri türü farklı renklerle belirtilmiştir.



Şekil 6.4 : LSTM ve MLP'nin farklı hisseler için test ve eğitim verileri ile farklı zaman çerçevelerindeki denenmelerinin %EDOC doğruluğu sonuçları

Hem LSTM hem de MLP modelleri, her hisse senedinin eğitim verileri ile denendiğinde, artan zaman çerçevelerinde EDOC doğruluğunun düzenli bir şekilde arttığı gözlemlenmiştir. Ancak, test verilerinin EDOC doğruluk değerleri aynı tutarlılığı göstermemektedir. Bunun nedeni, farklı hisseler için test verileriyle yapılan denemelerde MLP ve LSTM'nin en yüksek EDOC doğruluğuna ulaşan zaman çerçevelerinin birbirinden farklı olması ve zaman çerçeveleri arttıkça doğruluğun sürekli artmayıp, bazen azalabilmesidir.

Tüm hisse senetlerine ait test ve eğitim verilerindeki MLP ve LSTM modellerinin EDOC doğruluğunun genel ortalamaları Çizelge 6.6'da sunulmuştur.

Çizelge 6.6 : LSTM ve MLP'nin bütün hisselerin test ve eğitim verileri ile farklı zaman çerçevelerindeki denenmelerinin ortalama ortalama %EDOC doğruluğu

	MLP		LSTM	
Zaman Çerçevesi	**Test Verisi**	**Eğitim Verisi**	**Test Verisi**	**Eğitim Verisi**
Bir Günlük	54.29	45.70	42.86	47.70
İki Günlük	54.29	48.34	47.14	52.24
Üç Günlük	57.14	50.50	59.18	54.58
Dört Günlük	62.86	51.44	42.86	55.07
Beş Günlük	60.71	52.05	28.57	59.74

LSTM ile elde edilen en yüksek % EDOC doğruluğu, beş günlük zaman çerçevesinde eğitim verileriyle yapılan denemelerde elde edilmiştir. MLP ile ise en yüksek EDOC doğruluğu, dört günlük zaman çerçevesinde test verileriyle yapılan denemelerde elde edilmiştir. Her iki modelin de eğitim verileriyle yapılan denemelerinde doğruluğun, zaman çerçevesi arttıkça yükseldiği gözlemlenmiştir. LSTM'in test verileriyle yapılan denemeleri ise, tüm sonuçlar arasında en büyük varyasyonları göstermektedir. Bunun nedeni olarak, daha önce de bahsedilen test verilerinin sayısının az olması gösterilebilir.

7. TARTIŞMA

Bu çalışmada farklı modeller, parametreler, farklı frekanslardaki değişkenlerden oluşan verilerin ve farklı zaman çerçevesi kullanımının YSA kullanılarak yapılan hisse senedi tahminlerinin performansı üzerindeki etkisi iki ayrı örnekte test edilmiştir. İlk örnekte MLP farklı düğüm sayıları ve aktivasyon fonksiyonlarıyla denenmiştir. Bu örnekten elde edilen sonuçlara göre, en iyi performans gösteren düğüm sayılarının yaklaşık hesap kuralına uygun olduğu görülmüştür. Ayrıca, ReLU aktivasyon fonksiyonunun, sigmoid fonksiyonuna göre daha iyi sonuçlar verdiği gözlemlenmiştir. İkinci örnekte ise, farklı frekansta değişkenlerden oluşan veriler kullanılarak LSTM ve MLP ile farklı zaman çerçevelerinde tahminler yapılmıştır. Bu örnekte LSTM'in MLP'ye göre daha başarılı olduğu gözlemlenmiştir. Artan zaman çerçevelerinde yön tahmini doğruluğunun arttığı da fark edilmiştir.

İlk örnekten alınan sonuçlar değerlendirirken, bir yön tahmininin başarılı sayılabilmesi için doğruluğunun %50'nin üzerinde olması gerektiği belirtilmelidir. Çünkü %50'nin altındaki doğruluk oranları, tahminlerin rastgele olduğunu gösterir. Elde edilen sonuçlara göre, hisselerin bir günlük zaman çerçevesinde gideceği yönün doğru bir şekilde tahmin edilemediği söylenebilir. Bu durum, hisselerin kısa vadede tahmin edilemeyeceğini savunan rastgele yürüyüş teorisi ve EMH ile tutarlıdır. Özellikle test ve eğitim verileri incelendiğinde, test verilerinin ortalama EDOC doğruluğu %55 iken, eğitim verilerinin ortalama doğruluğu sadece %35'tir. Bu, farklı zaman çerçevelerinde yapılan tahminlerin de farklı sonuçlar verdiğini göstermektedir. Aynı hisse senedinin öngörülebilirliğindeki değişimler de dikkat çekicidir; örneğin, VAKBN hissesinin test verilerindeki EDOC doğruluğu %90 iken, eğitim verilerindeki doğruluk %30'dur. Bu sonuçlar, hisse senedi fiyatlarının ne kadar öngörülemez olabileceğini ortaya koymaktadır. Öte yandan, MAPE sonuçları EDOC sonuçlarına göre daha tutarlıydı. MAPE değeri, en düşük %2,75 ve en yüksek %9,01 arasında değişmekte olup, farklı hisse senetleri arasında büyük dalgalanmalar gözlemlenmemiştir. Bu da, yön tahmini tam olarak yapılamasa da, gelecekteki değerler hakkında daha yakın tahminler yapılabileceğini göstermektedir. Ancak, hisse senedi

yatırımlarında yön tahmininin önemli bir rol oynaması nedeniyle, MAPE sonuçları başarılı olsa da bu başarı, genel başarı üzerinde büyük bir etki yapmamaktadır.

İlk örnekte, ReLU aktivasyon fonksiyonu, sigmoid fonksiyonuna kıyasla daha iyi sonuçlar elde etmiştir. Bu farkın temel nedeni, ReLU'nun önceki bölümlerde de belirtildiği gibi, gradyen sorununu daha etkili bir şekilde çözmesidir. Ayrıca, ReLU'nun doğrusal olmayan ilişkileri daha iyi modelleyebilmesi, karmaşık veri yapılarında üstün performans sergilemesini sağlar. Sigmoid fonksiyonu, çıktıyı 0 ile 1 arasında sıkıştırarak doğrusal olmayan veri ilişkilerini modellemede yetersiz kalabilirken, ReLU'nun doğrusal olmayan verilere uyum sağlaması daha verimli olmaktadır. İki fonksiyonun türevleri karşılaştırıldığında, bu fark daha net anlaşılabilir. Sinir ağı eğitimi sırasında, aktivasyon fonksiyonlarının türevleri alındığında, sigmoid fonksiyonu zamanla sönümlenerek verimsiz hale gelir ve gradyen kaybı sorununu ortaya çıkarır. Oysa ReLU, negatif değerler için sıfır, pozitif değerler için doğrusal bir ilişki kurarak bu sorunu aşar ve daha hızlı öğrenmeyi mümkün kılar. Bu durum, elde edilen sonuçlara doğrudan yansımaktadır. Ayrıca, düğüm sayılarının incelenmesi sonucunda, elde edilen değerlerin yaygın olarak kabul edilen yaklaşık hesap kuralıyla uyumlu olduğu gözlemlenmiştir. Bu da, hisse senedi tahminleri gibi problemlerin çözümünde, sinir ağı modelinin seçimi açısından ReLU aktivasyon fonksiyonu ve yaklaşık hesap kuralının faydalı olabileceğini göstermektedir.

Yapılan örneklerde kullanılan verilerin karşılaştırması, her ikisinde de ortak olan MLP'nin bir günlük zaman çerçevesinde denenmesinden alınan sonuçların incelemesi ile yapılmıştır. İlk örnekte yapılan denemelerden elde edilen en iyi MAPE değeri 2.75 iken, ikinci örnekte bu değer 1.60'dır. Bu önemli bir iyileşmedir. Her iki örnekte de ReLU aktivasyon fonksiyonu taşıyan düğüm sayıları yaklaşık hesap kuralı ile belirlenen iki gizli katmanlı MLP kullanılarak bir günlük zaman çerçevesinde tahmin yapılmasına rağmen, ikinci örnekte MAPE hatası neredeyse yarı yarıya azalmıştır. EDOC açısından ise, ilk örnekte eğitim ve test verileriyle sırasıyla %55 ve %34 doğruluk oranı elde edilirken, ikinci örnekte sırasıyla %54 ve %45 doğruluk oranlarına ulaşılmıştır. Bu da EDOC açısından da bir ilerleme olduğunu göstermektedir. MAPE ve EDOC değerlerindeki bu gelişmeler, iki örnekte kullanılan verilerin farklı olması ile açıklanabilir. İlk örnekteki sadece günlük frekansa sahip 8 değişkenin kullanılması, ikinci örnekte ise farklı frekanslardan 23 değişkenin kullanılmasına kıyasla YSA'nın ilgili hisse hakkında daha az bilgi edinmesine ve bu sebeple tahminlerin daha başarısız olmasına yol açmış olabilir. Yani, daha önce de bahsedildiği üzere hisselerin neredeyse her şeyden etkilenmeleri dolayısıyla mümkün

olduğunca farklı bilgiler içeren çok sayıda değişken kullanılmasının YSA'nın hisseler hakkında daha iyi eğitilmesine yardımcı olabileceği ve bu sayede tahminlerin güçlendirilebileceği gözlemlenmiştir. Bu da kullanılan değişkenlerin önemini ortaya koymaktadır.

İkinci örneğin sonuçları incelendiğinde, MLP ile test verilerinde LSTM'ye göre daha düşük MAPE değerleri elde edildiği görülmüştür. Ancak eğitim verilerinde, özellikle üç ve daha fazla günlük zaman çerçevelerinde, LSTM'in MAPE değeri daha başarılı olmuştur. Verilerin ayrılmasında, test verilerindeki zaman damgası sayısının eğitim verilerine kıyasla çok daha az olduğu daha önce belirtilmiştir. Bu durum, test verileriyle elde edilen sonuçların daha fazla dalgalanmasına yol açmıştır. Dolayısıyla, eğitim verileriyle elde edilen sonuçların daha güvenilir olduğu söylenebilir. Sonuç olarak, daha büyük zaman çerçevelerinde LSTM'in MLP'den MAPE açısından daha başarılı olduğu sonucuna varılabilir.

İkinci örneğin sonuçlarını EDOC açısından incelediğimizde, LSTM ile eğitim verilerinin doğruluğunun beş günlük zaman çerçevesinde zirveye ulaştığı görülmüştür. MLP modeli ise test verilerinde doğruluğu dört günlük zaman çerçevesinde en yüksek seviyeye çıkarmıştır. Eğitim verileri açısından hem MLP hem de LSTM modellerinde doğruluk, zaman çerçevesi arttıkça yükselmiştir. Bu da daha uzun zaman çerçevelerinde hisse senetlerinin yönlerinin daha doğru tahmin edilebileceğini göstermektedir. LSTM'in test verilerindeki sonuçlar ise en büyük varyasyonları göstermiştir. Bu varyasyon, test verilerinin sayısının az olmasından kaynaklanmaktadır. Bu nedenle, eğitim verileriyle elde edilen sonuçların daha tutarlı olduğu söylenebilir. Sonuç olarak, MAPE sonuçlarında olduğu gibi, burada da LSTM'in artan zaman çerçevelerinde MLP'ye göre daha iyi sonuçlar verdiği görülmektedir.

İkinci örnekten elde edilen genel bulgulara göre, zaman çerçevesi uzadıkça tahminlerin doğruluğunun arttığı görülmüştür. Bu, hisse senetlerinin yönlerinin uzun vadede daha doğru bir şekilde tahmin edilebileceğini gösteriyor. Ancak, artan zaman çerçeveleriyle birlikte MAPE değeri bir miktar yükselmiştir. Bu artış büyük olmasa da, zaman çerçevesi daha da uzatıldığında, yön tahmini doğru yapılsa bile, gerçek değere yakınsanma oranının azalabileceği düşünülmektedir. Yani, en fazla beş günlük tahminlere kıyasla, aylık ya da yıllık tahminler yapıldığında, yön tahmininde bir artış olsa da, gerçek değere yakınlık azalacaktır. Bu durum, EDOC ile MAPE arasında bir denge olduğunu göstermektedir. Bu da, zaman çerçevesini artırmanın bir noktadan sonra verimli olmayacağını düşündürmektedir. Ayrıca, LSTM modelinin, zaman çerçevesi uzadıkça MLP modeline göre

daha iyi performans gösterdiği gözlemlenmiştir. Bu, LSTM'nin uzun vadeli hisse senedi tahmininde MLP'ye göre üstün olduğunu ortaya koymaktadır. LSTM'nin bu başarısının, geçmiş verileri de dikkate alarak tahmin yapmasından kaynaklandığı söylenebilir. Bu durum, Dow teorisiyle uyumlu olarak, hisse senetlerinin geçmiş değerleriyle bir ilişki olduğunu ortaya koymaktadır.

Yapılan çalışmanın literatürdeki diğer çalışmalarla kıyaslaması Çizelge 7.1'de yapılmıştır. Bu çizelgede, önceki çalışmaların ve mevcut çalışmanın kullandığı modellerle elde edilen sonuçlar, tahmin yapılan zaman çerçeveleri ve değerlendirme yöntemleriyle birlikte sunulmaktadır. Buradan, mevcut çalışmanın MAPE açısından bazı çalışmalardan geride kalsa da, EDOC açısından oldukça iyi bir konumda olduğu görülebilir. Ayrıca, bu çalışmanın önceki çalışmalardan farklı olarak çeşitli zaman çerçevelerini denediği de söylenebilir. Hem regresyon analizi hem de yön tahmini yapması bakımından da diğer çalışmalardan ayrılmaktadır. Literatürdeki çalışmalarla bu çalışmanın sonuçları arasındaki farklar; veri farklılıkları, ekonomik faktörler, model ve parametre farklılıkları, kullanılan yöntemler ve öngörülen gün sayısı gibi etmenlerle açıklanabilir. Bunun nedeni, hisse senedi piyasalarının hızla değişmesi ve pek çok farklı faktörden etkilenmesidir.

Çizelge 7.1 : YSA ile hisse senedi tahmini konusunda yapılmış bazı önceki çalışmaların ve bu çalışmanın sonuçları

Yapılan Çalışma	Zaman Aralığı	Metot	Karşılaştırma Yöntemi	Sonuç
[17]	Sonraki gün sonu	RNN LSTM	MAE	27.469 32.283
[18]	Sonraki ay sonu yönü	MLP LSTM	%DOĞRULUK	~%50
[19]	Sonraki gün sonu	MLP LSTM	MAPE	4.365 3.191
[20]	Sonraki gün sonu	MLP LSTM	MAPE	~0.3 ~0.5
[21]	Sonraki gün sonu	ARIMA LSTM	MAE	3.2698 3.5043
[22]	Sonraki gün sonu yönü	LSTM	%DOĞRULUK	~%52
[23]	Sonraki gün sonu yönü	LSTM	%DOĞRULUK	%53.53
[25]	Sonraki gün sonu	MLP	MAPE	~%1
Bu çalışma	Sonraki bir, iki, üç, dört, beş gün sonu	MLP LSTM	MAPE %DOĞRULUK	1.6 %59.74

Yapılan çalışmadan elde edilen sonuçlar, kısa vadeli tahminlerde hisselerin yönlerinin rastgele olduğunu, ancak zaman çerçevesi arttıkça bu yönlerin daha belirgin hale

geldiğini göstermektedir. Bu bulgular, hem hisselerin tahmin edilemeyeceğini savunan rastgele yürüyüş hipotezi ve EMH'yi hem de hisselerin tahmin edilebileceğini öne süren Dow teorisi ve rastgele olmayan yürüyüş hipotezini destekler niteliktedir. Ayrıca, hisse senetlerinin yön tahmininde çok uzun vadeler dışında, artan zaman çerçevelerinin daha iyi sonuçlar verdiği belirlenmiştir. LSTM gibi geçmiş verileri de dikkate alan yapıların, bu tür tahminler için daha etkili olduğu görülmüştür. Farklı frekanslardan gelen, mümkün olduğunca fazla değişkenin kullanılmasıyla daha iyi sonuçlar elde edilebileceği gözlemlenmiştir. Ayrıca, YSA parametreleri seçilirken yaklaşık hesap kuralının kullanılabileceği gösterilmiştir. Aktivasyon fonksiyonu olarak ise, ReLU'nun sigmoid fonksiyonuna kıyasla daha iyi performans sergilediği tespit edilmiştir.

Sonuçların genelleştirilebilirliği bazı sınırlamalara sahiptir. İlk örnek için, dünya çapında her hisse senedi için tüm etkili faktörlere ve veri noktalarına ulaşmak ve bunları hesaba katmak neredeyse imkansızdır. Ayrıca, tüm düğüm sayıları ve aktivasyon fonksiyonlarıyla MLP modeli üzerinde denemeler yapmak oldukça zordur. İlk örnekte, MLP denemeleri, belirli bir zaman diliminde BIST'ten dar bir değişken setiyle alınan verilerle sınırlı sayıda düğüm ve aktivasyon fonksiyonu kullanılarak yapılmıştır. İkinci örnekte de benzer kısıtlamalar mevcuttur. Yatırımcı duygusu, siyaset ve popüler kültür gibi faktörler hisse senedi fiyatları üzerinde etkili olabilir. Ayrıca, hisse senedi fiyatları, örneğin bir ülke arasındaki çatışmalar gibi durumlarla birlikte zaman içinde öngörülebilirlik açısından değişebilir. Farklı ülkelerdeki hisse senedi piyasaları, farklı ekonomik, sosyal ve politik faktörlerden etkilenebilir, bu da sonuçların her ülkede farklılık göstermesine yol açar. Aynı piyasa içindeki farklı hisse senetleri arasında da etkiler farklılık gösterebilir. Ayrıca, YSA katmanları ve düğümleri için çok sayıda farklı kombinasyon olduğu için, tüm olasılıkları test etmek imkansızdır. Her potansiyel kombinasyonu tüm hisse senetlerine uygulamak gerçekçi değildir. Bu çalışmada yalnızca belirli bir zaman diliminde yedi NASDAQ ve dört BIST hisse senedi incelendiğinden, genelleştirilebilir sonuçlar çıkarmak için yeterli veri bulunmamaktadır. Bu nedenle, sonuçların genelleştirilebilirliği sınırlıdır. Ancak, bulgular, hem mevcut literatürle hem de kendi içindeki verilerle tutarlı olduğu için geçerliliğe sahiptir.

8. SONUÇ VE ÖNERİLER

Bu çalışmada, YSA kullanarak hisse senetlerinin gelecekteki fiyatlarını tahmin etmek için farklı algılayıcı sayıları, aktivasyon fonksiyonları, zaman çerçeveleri, modeller ve farklı frekanslara sahip değişkenlerden oluşan veriler denenmiştir. Elde edilen sonuçların, hisse senedi tahmini gibi karmaşık sorunları çözmeye yönelik karar destek sistemlerinin geliştirilmesinde yardımcı olması hedeflenmiştir. Bu amaçla, alandaki önceki çalışmalar incelenmiş ve bu çalışmalardan faydalanarak ilerleme kaydedilmiştir. Karşılaştırmalar ise iki farklı örnekte yapılmıştır.

İlk örnekte, BIST'ten rastgele seçilen dört hisse senedinin son gün kapanış fiyatlarını tahmin etmek için MLP'nin performansı, farklı aktivasyon fonksiyonları ve düğüm sayıları kullanılarak MAPE ve EDOC açısından incelenmiştir. MLP, iki gizli katmanda değişen algılayıcı sayıları ve iki farklı aktivasyon fonksiyonu olan sigmoid ile ReLU ile test edilmiştir. En iyi sonuç, 8-5-3-1 düğüm sayılarıyla ve ReLU aktivasyon fonksiyonu ile VAKBN hissesinin test verisi kullanılırken elde edilmiştir. Bu düğüm sayılarının yaklaşık hesap kuralına uyması, gelecekteki çalışmalarda bu kuralın kullanılabileceğini göstermektedir. Ayrıca, ReLU'nun bu iş için daha uygun olduğu ve gelecekteki çalışmalar için tercih edilebileceği ortaya konmuştur. Sonuçların sadece belirli bir aralıkta ve tek bir hisse üzerinde etkili olup olmadığını görmek amacıyla, en iyi sonuçları veren parametrelerle tüm hisselerin test ve eğitim verileri kullanılarak elde edilen tahminler, EDOC açısından incelenmiştir. Tahminlerin gerçek değerlerle karşılaştırılması sonucunda, MAPE'nin düşük ve istikrarlı olduğu, ancak EDOC'un yüksek oynaklık gösterdiği ve ortalama doğruluğunun düşük olduğu görülmüştür. Bu durum, seçilen hisse senetlerinin belirli bir zaman dilimi boyunca oynak ve tahmin edilemez olduğunu göstermektedir. Sonuç olarak, bulgular rastgele yürüyüş hipotezi ve EMH gibi hisselerin tahmin edilemeyeceğini savunan görüşlerle uyumlu bulunmuştur.

İkinci örnekte, iki farklı YSA modeli olan MLP ve LSTM, farklı frekanslara sahip değişkenlerden oluşan verilerle çeşitli zaman çerçevelerinde test edilmiştir. İlk çalışmadan elde edilen sonuçlara dayanarak, kısa vadede hisse senetlerinin tahmin edilemeyeceği sonucuna varılmış ve bu nedenle daha uzun zaman çerçevelerinde de tahminler yapılmıştır. Ayrıca, ilk örnekte başarılı olduğu gözlemlenen yaklaşık hesap kuralı ve ReLU aktivasyon

fonksiyonu, ikinci örnekte kullanılan parametrelerin seçiminde temel alınmıştır. Yedi NASDAQ hissesi üzerinde farklı zaman çerçevelerinde MLP ve LSTM kullanılarak tahminler yapılmıştır. Elde edilen tahminler, gerçek sonuçlarla MAPE ve EDOC açısından karşılaştırılmıştır. Bu örnekteki, önceki örnektekinden farklı olarak, makroekonomik ve mikroekonomik ölçeklerde toplamda 23 farklı frekansta değişkenden oluşmaktaydı. Veriler, bir, iki, üç, dört ve beş günlük zaman çerçevelerine ayrılarak farklı zaman çerçeveleri oluşturulmuştur. Her zaman çerçevesi ve hisse için MLP ve LSTM ayrı ayrı eğitilmiştir. Eğitilen modellere hem eğitim hem de test verileri verilerek tahminler oluşturulmuştur. Elde edilen bulgulara göre, zaman çerçevesi büyüdükçe MAPE'nin arttığı, ancak EDOC doğruluğunun da arttığı gözlemlenmiştir. İlk örnek ile kıyaslandığında, farklı frekanstan değişkenlerden oluşan verilerin kullanılmasının daha iyi sonuçlar elde edilmesine yardımcı olduğu görülmüştür. Ayrıca, eğitim verilerinin sonuçlarının daha stabil olduğu ve uzun vadeli tahminlerde LSTM'nin MLP'ye göre daha iyi performans gösterdiği tespit edilmiştir. Sonuç olarak, zaman çerçevesi uzadıkça hisse senetlerinin yön tahminlerinin daha doğru yapılabildiği gözlemlenmiştir. Bu bulgular, hisse senetlerinin geçmiş değerleriyle olan ilişkisini gösterdiği için Dow teorisiyle uyumludur.

Çalışmadan elde edilen sonuçlar, hisse senedi fiyatlarının yönünün kısa vadede, örneğin ertesi günün sonunda rastgele olduğunu, ancak uzun vadede bu yönün belirlenebileceğini, yani belirli bir yönde hareket ettiğini göstermektedir. Bu sonuçlar, hisse senedi fiyatlarının tahmin edilemeyeceğini savunan rastgele yürüyüş hipotezi ve EMH gibi görüşlerle, fiyatların tahmin edilebileceğini ileri süren row teorisi ve rastgele olmayan yürüyüş hipotezi gibi görüşleri belirli yönlerden desteklemektedir. Bir hisse senedinin yönünün, artan zaman çerçevesi boyutuyla beraber daha iyi tahminlenebildiği sonucuna varılmıştır. Ayrıca, LSTM gibi tahmin üretirken geçmiş zaman damgalarını da kullanan yapıların bu tür görevler için daha uygun olduğu gözlemlenmiştir. Farklı frekanslardan gelen çok sayıda değişken içeren veriler kullanıldığında daha iyi sonuçlar alındığı da belirlenmiştir. YSA parametrelerini seçerken yaklaşık hesap kuralının uygulanabileceği gözlemlenmiş, ayrıca ReLU'nun sigmoid aktivasyon fonksiyonuna kıyasla daha uygun bir tercih olduğu sonucuna varılmıştır.

Gelecekteki çalışmalar, daha fazla değişken ve daha eski verilerle desteklenmiş, farklı hisse senetlerinin test edilmesini içerebilirler. Bu veriler, sosyal medya ve internet sitelerinden alınan bilgilerle zenginleştirilebilir. Tahminler, saatlik, haftalık ve aylık gibi kısa ve uzun vadeli farklı zaman çerçevelerinde yapılabilir. Ayrıca, farklı modeller üzerinde

de denemeler gerçekleştirilebilir. Bunun yanı sıra, dünya çapındaki çeşitli borsalarda ve daha uzun zaman dilimlerinde birçok hisse senedi üzerinde çalışmalar yapılabilir. MLP ve LSTM modelleri için farklı katman ve düğüm sayıları test edilebilir.

Yapılan çalışmanın literatüre katkıları şu şekilde özetlenebilir:

• MLP modelinde, iki gizli katmandaki farklı düğüm sayılarının performansları, ReLU ve Sigmoid aktivasyon fonksiyonlarıyla karşılaştırılmıştır.

• BIST'ten dört, NASDAQ'tan yedi farklı hisse senedi kullanılmıştır.

• Performans analizi için hem MAPE hem de EDOC değerlendirme kriterleri kullanılmıştır.

• Farklı zaman çerçevelerinde tahminler yapılmış ve karşılaştırılmıştır.

• Hem yüksek hem de düşük frekanslı değişkenler içeren veriler kullanılmıştır.

• MLP ve LSTM karşılaştırılmıştır.

Bu çalışmada, hedeflendiği gibi YSA kullanarak hisse senetlerinin tahmininde farklı algılayıcı sayıları, aktivasyon fonksiyonları, zaman çerçeveleri, modeller ve farklı frekanslardan değişkenler içeren veriler test edilmiştir. Elde edilen sonuçlar sunulmuş ve yorumlanmıştır. Gelecekte, hisse senedi tahmini gibi karmaşık problemlerin çözümünde bu çalışmadan elde edilen sonuçların kullanılabileceği söylenebilir.

KAYNAKLAR

[1] **Lu, W., Li, J., Li, Y., Sun, A., & Wang, J**. (2020). A CNN-LSTM-based model to forecast stock prices. *Complexity*, *2020*, 1–10.

[2] **Borsa Istanbul.** (n.d.). https://borsaistanbul.com/en

[3] ***Stock market, data updates, reports & news*. Nasdaq.** (n.d.). https://www.nasdaq.com/

[4] **Samuelson, P. A., & Nordhaus, W. D.** (2010). *Economics*. McGraw-Hill.

[5] **Bishop, G. W.** (1960). *Charles H. Dow and the dow theory*. Appleton-Century-Crofts.

[6] **Lo, A. W.-C., & McKinlay, A. C.** (2011). *A non-random walk down wall street*. Princeton University Press.

[7] **Read, C.** (2013). *The efficient market hypothesists: Bachelier, samuelson, fama, Ross, Tobin, and Shiller*. Palgrave Macmillan.

[8] **Malkiel, B. G.** (2003). A random walk down wall street. W.W. Norton & Company

[9] **Aggarwal, C. C.** (2018). *Neural Networks and Deep Learning A Textbook* (1st ed.). Springer Cham.

[10] **Kolisnyk, T., Ruban, O., Fil, N., & Kutsenko, S.** (2018). Application of an Artificial Neural Network for Design of Sustained-Release Matrix Tablets Containing Vaccinium Myrtillus Leaf Powder Extract. *Asian Journal of Pharmaceutics*, *12*(2), 137–145.

[11] **Deng, T.** (2023). Effect of the number of hidden layer neurons on the accuracy of the back Propagation Neural Network. *Highlights in Science, Engineering and Technology*, *74*, 462–468.

[12] **Bengio, Y., & LeCun, Y.** (2007). Scaling learning algorithms toward ai. *Large-Scale Kernel Machines*, 321–360.

[13] **Uzair, M., & Jamil, N.** (2020). Effects of hidden layers on the efficiency of Neural Networks. *2020 IEEE 23rd International Multitopic Conference (INMIC)*.

[14] **Titiya, M., Bala, A., & Degadwala, S.** (2024). Analyzing the Effect of Different Activation Functions in Deep Learning on Accuracy and Execution time. *International Journal of Intelligent Systems and Applications in Engineering*, *12*(4), 742–749.

[15] **Burzawa, L., Li, L., Wang, X., Buganza-Tepole, A., & Umulis, D. M.** (2020). Acceleration of PDE-based biological simulation through the development of neural network metamodels. *Current Pathobiology Reports*, *8*(4), 121–131.

[16] **Deep learning toolbox. MATLAB.** (n.d.). https://www.mathworks.com/products/deep-learning.html

[17] **Zaheer, S., Anjum, N., Hussain, S., Algarni, A. D., Iqbal, J., Bourouis, S., & Ullah, S. S.** (2023). A multi parameter forecasting for Stock Time Series data using LSTM and deep learning model. *Mathematics*, *11*(3), 590.

[18] **Chudziak, A.** (2023). Predictability of stock returns using neural networks: Elusive in the long term. *Expert Systems with Applications*, *213*, 119203.

[19] **Ren, S., Wang, X., Zhou, X., & Zhou, Y.** (2023). A novel hybrid model for stock price forecasting integrating encoder forest and informer. *Expert Systems with Applications*, *234*, 121080.

[20] **Oukhouya, H., & El Himdi, K.** (2023). Comparing machine learning methods—SVR, XGBoost, LSTM, and MLP— for forecasting the Moroccan stock market. *IOCMA 2023*.

[21] **Yang, Z., & Wang, Z.** (2024). The research of Nvidia Stock Price Prediction based on LSTM and Arima model. *Highlights in Business, Economics and Management*, *24*, 896–902.

[22] **Yang, C., Zhai, J., & Tao, G.** (2020). Deep learning for price movement prediction using convolutional neural network and long short-term memory. *Mathematical Problems in Engineering*, *2020*, 1–13.

[23] **Choi, J., Yoo, S., Zhou, X., & Kim, Y.** (2023). Hybrid information mixing module for stock movement prediction. *IEEE Access*, *11*, 28781–28790.

[24] **B L, Shilpa & Shambhavi, Ravi.** (2021). Combined deep learning classifiers for stock market prediction: integrating stock price and news sentiments. Kybernetes. ahead-of-print. 10.1108/K-06-2021-0457.

[25] **Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A.** (2020). Stock closing price prediction using Machine Learning Techniques. *Procedia Computer Science*, *167*, 599–606.

[26] **Amit Kumar Tyagi, & Ajith Abraham.** (2022). *Recurrent neural networks: Concepts and applications*. CRC Press.

[27] **Kumar, N. B., & Mohapatra, S.** (2015). *The use of technical and fundamental analysis in the stock market in emerging and developed economies*. Emerald Group Publishing Limited.

[28] **Standfield, K.** (2005). *Intangible finance standards: Advances in fundamental analysis & technical analysis*. Elsevier/Academic Press.

[29] **Kwong, I. C. K.** (2001). *Financial Forecasting Using Neural Network or Machine Learning Techniques* (thesis).

[30] **Kap.** (n.d.). https://www.kap.org.tr/tr/

[31] ***The Statistics Portal*. Statista.** (n.d.). https://www.statista.com/

[32] **Dean, E., Elardo, J., Green, M., Wilson, B., & Berger, S.** (2016). *Principles of microeconomics: Scarcity and social provisioning* \ (2nd ed.). Open Textbook Library.

[33] **Kanat, M. Ş.** (2011). *Hisse senedi getirilerini etkileyen faktörlerin analizi: İMKB'de bir uygulama* (thesis).

[34] **DEMİR, .A.** (2001). Hisse Senedi Fiyatını Etkileyen İşletme Düzeyindeki Faktörler ve Mali Sektör Üzerine İMKB'de Bir Uygulama. Süleyman Demirel Üniversitesi İktisadi Ve İdari Bilimler Fakültesi Dergisi, 6(2).

[35] ***Gold Spot Live Price Chart: Trade Gold*. Gold Spot Live Price Chart | Trade Gold.** (n.d.). https://capital.com/gold-price

[36] ***The long term perspective on markets*. Macrotrends.** (n.d.). https://www.macrotrends.net/

[37] **Tarihsel Fiyat Bilgileri. Tarihsel Fiyat Bilgileri | İş Yatırım.** (n.d.). https://www.isyatirim.com.tr/tr-tr/analiz/hisse/Sayfalar/Tarihsel-Fiyat-Bilgileri.aspx

[38] **Goodfellow, I., Bengio, Y., & Courville, A.** (2016). Deep learning. MITP.

[39] **Nielsen, M.** (2019). *Neural Networks and Deep Learning*

[40] **Suzuki, Kenji.** (2013). Artificial Neural Networks – Architectures and Applications.

[41] **VURAL, B. B.** (2007). *Yapay Sinir Ağlari ile Finansal Tahmin* (thesis).

[42] **Vui, Chang & Soon, Gan & On, Chin & Alfred, Rayner & Anthony, Patricia.** (2013). A review of stock market prediction with Artificial neural network (ANN). Proceedings - 2013 IEEE International Conference on Control System, Computing and Engineering, ICCSCE 2013. 477-482. 10.1109/ICCSCE.2013.6720012.

[43] **OKUTKAN, C.** (2014). Borsa İstanbul Şirketlerinin Hisse Senedi Getirilerinin Yapay Sinir Ağlari ve Çoklu Regresyon Yöntemleri Kullanarak Analizi (thesis).

[44**] Ogasawara, E., Martinez, L. C., de Oliveira, D., Zimbrao, G., Pap, G. L., & Mattoso, M.** (2010). Adaptive normalization: A novel data normalization approach for non-stationary Time Series. *The 2010 International Joint Conference on Neural Networks (IJCNN)*.

[45] **YUSSIF, A.-R. B.** (2020). *Yapay Sinir Ağlari ile Borsa Endeksi Tahmini: Gana Borsasi Üzerine Bir Uygulama* (thesis).

[46] **KALYONCU, S.** (2020). *Borsa Analizi ve Tahmini İçin Derin Öğrenme Ağlari* (thesis).

[47] **Nwankpa, Chigozie & Ijomah, W. & Gachagan, Anthony & Marshall, Stephen.** (2020). Activation Functions: Comparison of trends in Practice and Research for Deep Learning. 10.48550/arXiv.1811.03378.

[48] **Doğan, Ahmet & Okatan, Ali & Çetinkaya, Ali.** (2021). Vehicle Classification and Tracking Using Convolutional Neural Network Based on Darknet Yolo with Coco Dataset.

[49] **Vujicic, T., Matijevic, T., Ljucovic, J., Balota, A., & Sevarac, Z.** (2016). Comparative Analysis of Methods for Determining Number of Hidden Neurons in Artificial Neural Network.

[50] **Heaton, J.** (2005). Introduction to Neural Networks with Java, Heaton Research Inc

[51] **Salem, F. M.** (2022). *Recurrent neural networks*. Springer International Publishing.

[52] **Wang, L., Yang, Y., Min, R., & Chakradhar, S.** (2017). Accelerating deep neural network training with inconsistent stochastic gradient descent. *Neural Networks*, *93*, 219–229. https://doi.org/10.1016/j.neunet.2017.06.003

[53] **Gu, H., Wang, Y., Hong, S., & Gui, G.** (2019). Blind Channel Identification Aided Generalized Automatic modulation recognition based on Deep Learning. *IEEE Access*, *7*, 110722–110729.

[54] **Géron, A.** (2018). *Neural networks and deep learning*. O'Reilly.

[55] **Cerqueira, V., Torgo, L., & Mozetič, I.** (2020). Evaluating time series forecasting models: An empirical study on performance estimation methods. *Machine Learning*, *109*(11), 1997–2028.

[56] **Chandrika, P. V., & Srinivasan, K. S.** (2021). Predicting stock market movements using artificial neural networks. *Universal Journal of Accounting and Finance*, *9*(3), 405–410.

[57] **Ergen, T., Mirza, A. H., & Kozat, S. S.** (2020). Energy-efficient LSTM networks for online learning. *IEEE Transactions on Neural Networks and Learning Systems*, *31*(8), 3114–3126.

ÖZGEÇMİŞ

Ad-Soyad : Mustafa Ozan DUMAN

ÖĞRENİM DURUMU:

- **Lisans :** 2019, Orta Doğu Teknik Üniversitesi, Mühendislik Fakültesi, Elektrik Elektronik Mühendisliği

MESLEKİ DENEYİM:

- 2020-2021 İ2İ Bilişim'de Yerli Milli 5G projesinde Analyst Developer olarak çalıştım.
- 2021'de BVK Teknoloji'de Embedded Software Developer olarak çalıştım.
- 2022'de İSDEMİR'de elektrik otomasyon mühendisi olarak çalıştım.
- 2023- Tübitak 1001 projesinde bursiyer olarak yer almaktayım.
- 2024- Bursa Uludağ Üniversitesi'nde Araştırma Görevlisi olarak çalışmaktayım.

YÜKSEK LİSANS TEZİNDEN TÜRETİLEN ÇALIŞMALAR

- **Duman, M. O., & Tağluk, M. E. (2024).** Assessment of an optimal multilayer perceptron model in estimation of stock's price rates; a sample from the Turkish Stock Exchange bist. *2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP), 3,* 1–5. https://doi.org/10.1109/idap64064.2024.10710792
- **Duman, M. O., & Tağluk, M. E. (2024).** Performance assessment of temporal frame size in stock price prediction. *2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP),* 1–5. https://doi.org/10.1109/idap64064.2024.10711131