

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**YİNELEMELİ SİNİR AĞLARI İLE SERMAYE PİYASASI
YÖN TAHMİNİ ÜZERİNE BİR ÇALIŞMA**

YÜKSEK LİSANS TEZİ

Muhiddin Çağlar EREN

Endüstri Mühendisliği Anabilim Dalı

Endüstri Mühendisliği Lisansüstü Programı

Tez Danışmanı: Prof. Dr. Alp ÜSTÜNDAĞ

HAZİRAN, 2022

**YİNELEMELİ SİNİR AĞLARI İLE SERMAYE PİYASASI
YÖN TAHMİNİ ÜZERİNE BİR ÇALIŞMA**

YÜKSEK LİSANS TEZİ

**Muhiddin Çağlar EREN
(507181138)**

Endüstri Mühendisliği Anabilim Dalı

Endüstri Mühendisliği Lisansüstü Programı

Tez Danışmanı: Prof. Dr. Alp ÜSTÜNDAĞ

HAZİRAN, 2022

İTÜ, Lisansüstü Eğitim Enstitüsü'nün 507181138 numaralı Yüksek Lisans Öğrencisi Muhiddin Çağlar EREN, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı “YİNELEMELİ SİNİR AĞLARI İLE SERMAYE PİYASASI YÖN TAHMİNİ ÜZERİNE BİR ÇALIŞMA” başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : **Prof. Dr. Alp ÜSTÜNDAĞ**
İstanbul Teknik Üniversitesi

Jüri Üyeleri : **Dr. Öğr. Üyesi Ömer Faruk BEYCA**
İstanbul Teknik Üniversitesi

Prof. Dr. Selim ZAIM
İstanbul Sabahattin Zaim Üniversitesi

Teslim Tarihi : 1 Haziran 2022
Savunma Tarihi : 29 Haziran 2022





Aileme,



ÖNSÖZ

Tüm hayatım boyunca yanımda olan, en zorlu kararlarımda beni cesaretlendiren, hiçbir zaman içten ve sıcak desteğini esirgemeyen aileme sonsuz şükranlarımı sunarım.

Tez çalışmam boyunca yönlendirmeleriyle ve bilgisiyle bana yol gösteren danışmanım Prof. Dr. Alp ÜSTÜNDAĞ’a teşekkürlerimi sunarım.

Haziran 2022

Muhiddin Çağlar Eren



İÇİNDEKİLER

Sayfa

ÖNSÖZ	vii
İÇİNDEKİLER	ix
KISALTMALAR	xi
ÇİZELGE LİSTESİ.....	xiii
ŞEKİL LİSTESİ.....	xv
ÖZET	xvii
SUMMARY	xix
1. GİRİŞ	1
1.1 Finansal Piyasalarda Tahmin	1
1.2 Derin Öğrenme	2
1.3 Tezin Amacı	3
1.4 Tez Akışı	4
1.5 Literatür Araştırması	4
2. METODOLOJİ	11
2.1 Zaman Serileri	11
2.2 Yapay Sinir Ağları	13
2.2.1 İleri Beslemeli Sinir Ağları	16
2.2.2 Yinelemeli Sinir Ağları	16
2.2.2.1 LSTM	19
2.2.2.2 GRU	21
2.3 Boyut İndirgeme.....	22
2.3.1 Özellik Seçimi	24
3. UYGULAMA TASARIMI	29
3.1 Deneyisel Ortam Bilgileri	30
3.2 Veri Kümesi	30
3.3 Veri Etiketleme ve Ölçekleme	31
3.4 Özellik Seçim Stratejisi	32
3.5 Zaman Serisi Verileri Dizileme Yaklaşımları	32
3.6 Tasarlanan Model Parametreleri	36
3.7 Performans Metrikleri	37
3.8 Sonuçların Değerlendirilmesi	38
4. MODEL PERFORMANS ÇIKTILARI.....	39
4.1 Geri Bakma Değeri 100 Olan Dizilemede Tahmin Performansı	39
4.2 Geri Bakma Değeri 200 Olan Dizilemede Tahmin Performansı	41
4.3 Geri Bakma Değeri 300 Olan Dizilemede Tahmin Performansı	43
4.4 Geri Bakma Değeri 400 Olan Dizilemede Tahmin Performansı	45
4.5 Geri Bakma Değerleri İçin Performans Karşılaştırması	46
4.6 Özellik Seçim Yöntemleri İçin Performans Karşılaştırması	48
5. SONUÇLAR VE ÖNERİLER	51
KAYNAKLAR	55

ÖZGEÇMİŞ.....	61
---------------	----



KISALTMALAR

ANN	: Artificial Neural Network
AR	: Autoregressive Model
ARCH	: Autoregressive Conditional Heteroskedasticity Model
ARIMA	: Autoregressive Integrated Moving Average
ARMA	: Autoregressive Moving Average Model
AUC	: Area Under the Curve
BİST	: Borsa İstanbul
CNN	: Convolutional Neural Network
FNN	: Feedforward Neural Network
GARCH	: Generalized Autoregressive Conditional Heteroskedasticity Model
GRU	: Gated Recurrent Unit
HSIC	: Hilbert Schmidt Independence Criterion
LASSO	: Least Absolute Shrinkage and Selection Operator
LSTM	: Long Short Term Memory
MA	: Moving Average Model
MLP	: Multilayer Perceptron
MRMR	: Minimum Redundancy Maximum Relevance
NAR	: Nonlinear Autoregression Model
NSE	: National Stock Exchange of India
NYSE	: New York Stock Exchange
PCA	: Principal Component Analysis
RF	: Random Forest
RFE	: Recursive Feature Elimination
RNN	: Recurrent Neural Network
SVM	: Support Vector Machines
TAR	: Threshold Autoregression Model



ÇİZELGE LİSTESİ

Sayfa

Çizelge 3.1 : Python paket versiyonları.	30
Çizelge 3.2 : AKBNK, TCELL, FROTO veri kümesi detayları.	30
Çizelge 3.3 : LSTM ve GRU model parametreleri.	37
Çizelge 4.1 : AKBNK 100 adımlı dizileme performans sonuçları.	40
Çizelge 4.2 : TCELL 100 adımlı dizileme performans sonuçları.	40
Çizelge 4.3 : FROTO 100 adımlı dizileme performans sonuçları.	41
Çizelge 4.4 : AKBNK 200 adımlı dizileme performans sonuçları.	42
Çizelge 4.5 : TCELL 200 adımlı dizileme performans sonuçları.	42
Çizelge 4.6 : FROTO 200 adımlı dizileme performans sonuçları.	43
Çizelge 4.7 : AKBNK 300 adımlı dizileme performans sonuçları.	43
Çizelge 4.8 : TCELL 300 adımlı dizileme performans sonuçları.	44
Çizelge 4.9 : FROTO 300 adımlı dizileme performans sonuçları.	44
Çizelge 4.10: AKBNK 400 adımlı dizileme performans sonuçları.	45
Çizelge 4.11 : TCELL 400 adımlı dizileme performans sonuçları.	45
Çizelge 4.12 : FROTO 300 adımlı dizileme performans sonuçları.	46



ŞEKİL LİSTESİ

Sayfa

Şekil 1.1 : Zaman serisi segmentlere ayırma işlemi.	8
Şekil 2.1 : Yapay sinir ağlarında nöron modeli.	15
Şekil 2.2 : Tek gizli katmanlı ileri beslemeli ağ mimarisi.	16
Şekil 2.3 : RNN hücresi.	17
Şekil 2.4 : RNN dizi açılımı.	18
Şekil 2.5 : Tek bir katman içeren standart RNN ağı.	18
Şekil 2.6 : LSTM hücresi.	20
Şekil 2.7 : LSTM dizisi.	21
Şekil 2.8 : GRU mimarisi.	22
Şekil 3.1 : Hisse senedi yön değişimi tahminleme diagramı.	29
Şekil 3.2 : Zaman serisi üzerinde kayan pencereler yöntemi ile dizi oluşturma ve etiketleme.	33
Şekil 3.3 : Geri bakma değeri 100 için veri kümesinin dizilere ayrılması.	34
Şekil 3.4 : Geri bakma değeri 200 için veri kümesinin dizilere ayrılması.	34
Şekil 3.5 : Geri bakma değeri 300 için veri kümesinin dizilere ayrılması.	35
Şekil 3.6 : Geri bakma değeri 400 için veri kümesinin dizilere ayrılması.	36



YİNELEMELİ SİNİR AĞLARI İLE SERMAYE PİYASASI YÖN TAHMİNİ ÜZERİNE BİR ÇALIŞMA

ÖZET

Finansal piyasaları etkileyen faktörler ele alınarak yapılacak finansal piyasa geleceğinin tahminleme süreci, doğru portföy yönetimini sağlayarak kaybı azalıp kazancı arttırabilir. Fakat finansal veriler doğrusal olmayan dinamik ve kaotik karakteristiğe sahip olduğundan finansal piyasalarda karar vermek zorlu bir süreçtir. Bu probleme çözüm sağlayabilmek amacıyla kullanılabilecek farklı yöntemler mevcuttur.

Bunlardan biri olan derin öğrenme, girdi ve çıktı arasındaki ilişkiyi birden fazla sinir ağı katmanında modelleyerek yapabilen makine öğrenimi yöntemleri olarak tanımlanır. Veri biliminin birçok alanında karmaşık veri kümeleri üzerinde yüksek performans gösterdiğinden, son zamanların ilgi çekici yöntemlerinden biri olan derin öğrenmede, geleneksel modellerin aksine, yüksek boyutlu çok değişkenli problemler ve doğrusal olmayan ilişkiler modellenebilir. Bu nedenle derin öğrenme ile sağlanan algoritma başarıları tahminleme çalışmalarında büyük yer tutmaktadır.

Bu çalışmada, derin öğrenme kavramı altında değerlendirilen yinelemeli sinir ağları mimarilerinden LSTM ve GRU, finansal veri ilişkilerini modellemede kullanılmıştır. Zaman serisi olarak modellenen finansal veri kümeleri Reuters kaynak alınarak oluşturularak derin sinir ağı tahminleme modelleri ile Borsa İstanbul'da işlem gören AKBNK, TCELL ve FROTO hisse senetlerinin fiyat hareketlerinin yön tahmini için karşılaştırmalı yöntemler sunulmuştur.

Çalışmada derin öğrenme yöntemlerinin bir finansal veri kümesi kullanılarak başarılı bir tahmin performansı gösterip gösteremeyeceğinin ortaya konulması amaçlanmıştır. Bununla beraber, çalışma için hazırlanan finansal veri kümeleri üzerinde doğrusal olmayan özellik seçim yöntemleri ile yapılan boyut indirgeme yaklaşımları ve finansal zaman serisi olarak modellenen finansal veri kümelerinin farklı dizileme yaklaşımları ile segmentlere ayrılarak model girdisi olarak kullanılmasının baz model tahmin performansını iyileştirip iyileştirmeyeceğinin karşılıklı analizi yapılmıştır.

Boyut indirgeme sürecinde, finansal veri kümeleri etiketleme yapılarak denetimli öğrenme algoritmalarına elverişli hale getirildikten sonra HSIC Lasso, mRMR, RFE özellik seçim yöntemleri; segmentasyonda da 100, 200, 300 ve 400 adımlı dizileme birer adımlı kayan pencereler yöntemi tasarımı uygulanmıştır. Çalışmanın amacı doğrultusunda LSTM ve GRU mimarileri üzerinde hiper parametre ayarlamaları literatür baz alınarak yapılmıştır. Bununla beraber, tüm tahmin senaryolarında 70 olarak alınmış indirgenmiş özellik sayısına, model stres testlerinden sonra karar verilmiştir.



A STUDY ON DIRECTION PREDICTION OF CAPITAL MARKETS WITH RECURRENT NEURAL NETWORKS

SUMMARY

The forecasting process of the future of the financial market, which is made by considering the factors affecting the financial markets, can reduce the loss and increase the profit by providing the right portfolio management. However, since financial data has non-linear, dynamic and chaotic characteristics, it is a difficult process to make decisions in financial markets.

Investor decisions can be affected by many macroeconomic factors such as changing economic conditions, political relations, periodic sector based changes, news in the media, expectations and preferences of investors, and non-linear indicators. Decision-making in capital markets, which are not only dependent on economic indicators, but also seriously affected by non-financial factors, has become a very important research area for investors to manage their portfolios correctly. Providing reliable forecasts for capital markets is of great importance for effective market strategies. Therefore, it is necessary to forecast the stock market in order to contribute to the decision-making processes and reduce the investment risk of the investors. There are different methods that can be used to provide a solution to this forecasting problem.

Deep learning, defined as machine learning methods that can model the relationship between input and output in more than one neural network layer, is one of them. Contrary to traditional models, high-dimensional multivariate problems and nonlinear relationships can be modeled in deep learning. Traditional models can be inefficient in dealing with high-dimensional multivariate problems, nonlinear relationships, and incomplete datasets. These data features often occur in real-world data and can be modeled more efficiently using deep neural networks.

Recurrent neural network is one of the most interesting methods of recent times, since it performs well on complex data sets in many areas of data science. For this reason, algorithm achievements provided by deep learning have a great place in financial forecasting studies. Increasing computers' processing power and facilitating access to data in parallel with the development of hardware and software technologies are some of the important milestones that accelerate the scientific development in this field.

In this study, LSTM and GRU, recurrent neural network architectures evaluated under the concept of deep learning, are used to model financial data relationships. Financial datasets modeled as time series are created based on Reuters, and comparative methods are used for the direction prediction of the price movements of AKBNK, TCELL and FROTO, which are stocks traded in Borsa Istanbul, with deep neural network forecasting models.

Thus, it is aimed to reveal whether deep learning methods can show an acceptable forecasting performance by using financial datasets that include multi dimensional feature space containing many indicators such as company financials, technical indicators, periodic company reports and macroeconomic data. In addition, a comparative analysis has been made whether the use of dimension reduction approaches with nonlinear feature selection methods on the financial datasets prepared for the study and the use of financial datasets modeled as financial time series by segmenting with different sequencing approaches as model input will improve the base recurrent neural network models forecasting performance.

In the age of big data, where large amounts of high-dimensional data are widespread in a wide variety of fields, the rapid growth of data brings with it difficulties for effective and efficient data management. The necessity of applying data mining and machine learning techniques to obtain meaningful information from different types of data is better understood in the age of big data.

As a data preprocessing strategy, feature space size reduction strategies have proven to be effective and efficient, especially in preparing high-dimensional data for a variety of data mining and machine learning problems. Dimension reduction methodology aims to create simpler and more understandable models, improve data mining performance, and prepare clean, understandable data.

When algorithms are applied to high-dimensional data, phenomena known as the curse of dimensionality appear as a critical issue. This concept refers to the phenomenon of sparseness of data in high-dimensional space, which negatively affects algorithms designed for low-dimensional space. Also, with a large number of features, learning models tend to over-learn which can lead to poor performance on unseen data. High-dimensional data can significantly increase memory storage requirements and computational costs for data analytics. In addition, the simplification of the models will increase their applicability and explainability to real life. Models that produce faster results by keeping our model success performance criteria in similar value ranges will provide many benefits in real life. Dimension reduction in the data feature space is one of the most powerful methodology to solve these problems.

In the dimension reduction process applied in the study, HSIC Lasso, mRMR, RFE feature selection methods are used. Feature selection methods were applied to the labeled data before sending time sequences prepared by the sliding windows method to designed LSTM and GRU architecture. The total number of features of the stock datasets prepared in a broad perspective based on Reuters is over 170 and this space is reduced to 70 for all scenarios with the applied methods. The reduced feature number, which is taken as 70 in all forecast scenarios, is decided after the model stress tests.

In time series data segmentation phase, 100, 200, 300 and 400 steps sequencing is done with one day step sliding windows method. Two-dimensional arrays which are used as input to LSTM and GRU are created from the raw data in this time dependent sliding window approach. Models are trained with segmented data using 1240 days on the dataset covering 1270 trading days in total. Afterwards, the classification of 30-day stock directional movements is designed to give a buy and sell signal to investors. In line with the purpose of the study, hyper parameter adjustments on LSTM and GRU architectures are made depending on financial forecasting literature.

The models' performance outputs obtained in the study are presented for sequencing with 4 different look-back values for AKBNK, TCELL and FROTO stocks, and then performance comparisons are given by considering the look-back values and feature selection methods separately. Performance comparisons were made within the framework of 4 determined metrics. These metrics are long accuracy measuring "BUY" decision, short accuracy measuring "SELL" decision, weighted F1 score measuring overall model performances, and total accuracy metrics measuring total prediction accuracy.

During the decision-making phase, it is determined that the model could exhibit a balanced performance of more than 50% in long and short accuracy and increase over 55% in total forecast accuracy as an investment acceptance criterion. Performance metrics are evaluated in this respect, and the model that did not meet the investment acceptance criteria in at least one metric is evaluated as unsuccessful and was interpreted as open to improvement in terms of performance.





1. GİRİŞ

Finansal piyasalarda hisse senedi fiyat hareketlerinin derin öğrenme yöntemleriyle tahmin edilmesini, tahmin doğruluğunun farklı yöntemlerle iyileştirilebileceğini göstermeyi amaçladığımız çalışmamızın bu bölümünde finansal piyasalarda tahminleme ve derin öğrenme kavramlarını, literatürdeki farklı yaklaşımlar ile yapılmış uygulamaları ve çalışmamızın amaçlarını inceleyeceğiz.

1.1 Finansal Piyasalarda Tahmin

Ekonomide finans sektörünün ve reel sektörün uygun şekilde geliştirilmesi büyük önem taşımaktadır. İktisat teorilerine göre, söz konusu oluşumların finansal bozulmaya uğraması veya büyümemesi durumunda ekonomik büyüme ve kalkınma zorlaşacaktır. Ekonomik olarak gelişmiş ülkeler, para ve sermaye piyasalarını içeren güçlü bir finans sektörüne sahiptir. Finans sektörünün önemli bir alt kümesi olan sermaye piyasasının uygunsuz gelişimi, sadece ülkenin para sistemi üzerinde ek baskı oluşturmakla kalmaz, aynı zamanda üretim ve hizmet birimlerini aktif bir sermaye piyasasının faydalarından mahrum bırakır. Sermaye piyasasının en önemli amacı, şeffaf süreçlerin uygulanması yoluyla hissedarları tüzel kişiler ve ekonomik birimlerle buluşturmaktır.

Menkul kıymetler borsalarının sektördeki birimler arasında bir köprü konumunda olması, kaynakların etkin dağılımı ve kullanımı gibi üstlendiği işlevler düşünüldüğünde ekonomi üzerinde yadsınamaz bir etkisi vardır. Finansal ve sermaye piyasalarını önemli ölçüde etkilemesi nedeniyle küçük tasarrufları makroekonomik yatırımlara dönüştürecek bir mekanizma sağlamaktadır. Yatırım, pazarlanabilir ve pazarlanamaz finansal varlıklar da dahil olmak üzere çok çeşitli faaliyetleri kapsar. Finansal piyasalara yatırım yapan herkesin amaçları arasında getiri elde etmek vardır. Yatırım için teşvik edici bir faktör yoksa, insanlar tipik olarak mevcut tüketimi gelecekteki tüketime tercih ederler. Bu nedenle, bireysel yatırımcı, mevcut yatırımlarından pozitif getiri elde etmeyi ve gelecekteki tüketim fırsatlarından

yararlanmayı bekleyebilmelidir. Yatırım süreci, yatırım için menkul kıymet türü, yatırım miktarı ve yatırım zamanı konusunda karar vermeyi gerektirir [1].

Finansal piyasalarda karar vermek ise finansal piyasaların doğası gereği zor bir süreci ifade etmektedir. Zaman serileri olarak modellenen finansal verilerin doğrusal olmayan dinamik ve kaotik karakteristiği bu serilerin tahmin edilebilmesini oldukça zorlaştırmaktadır. Yatırımcı kararları değişen ekonomik koşullar, devletler arasındaki politik ilişkiler, dönemsel sektör bazlı değişimler, medyada yer alan haberler, yatırımcıların beklenti ve tercihleri, doğrusal olmayan göstergeler gibi birçok makroekonomik faktörden etkilenebilmektedir. Yalnızca ekonomik göstergelere bağlı olmayan finans dışı faktörlerden de ciddi şekilde etkilenen sermaye piyasalarında karar verme yatırımcıların doğru portföy yönetimleri için oldukça önemli bir araştırma alanı haline gelmiştir. Sermaye piyasaları için güvenilir tahminlerin ortaya konulması etkin pazar stratejileri için büyük önem taşımaktadır [2]. Bu nedenle, karar verme süreçlerine katkı sağlamak ve yatırımcıların yatırım riskini azaltmak için hisse senedi piyasasının tahmin edilmesi gerekmektedir.

1.2 Derin Öğrenme

Derin öğrenme kavramı son yıllarda veri biliminin birçok alanında karmaşık veri kümeleri üzerinde gösterdiği yüksek performanstan dolayı ilgi çekmektedir. Nesne algılama, doğal dil işleme, finansal zaman serileri tahmini gibi alt araştırma alanlarında derin öğrenme algoritmalarının gösterdiği başarımlar oranları bu alanda yapılacak çalışmaların da önünü açarak literatürde kendine geniş bir yer bulmaya başlamıştır. Geleneksel modeller, yüksek boyutlu çok değişkenli problemlerle, doğrusal olmayan ilişkilerle ve eksik veri kümeleriyle mücadele etmekte verimsiz kalabilir. Bu özellikler genellikle gerçek dünya verilerinde ortaya çıkar ve derin sinir ağları kullanılarak daha verimli modellenenirler. Temel olarak derin öğrenme, girdi ve çıktı arasındaki ilişkiyi birden fazla sinir ağı katmanında modelleyerek yapabilen makine öğrenimi yöntemlerine verilen genel bir isimdir. Donanım ve yazılım teknolojilerinin gelişmesine paralel olarak artan işlem gücü ve veriye erişimin kolaylaşması bu alandaki gelişimin önünü açan önemli kilometre taşlarından birkaçıdır.

Zaman serileri dizi modelleme probleminin alt araştırma alanlarından biridir. Derin öğrenme algoritmaları ile literatürde daha popüler olan makine çevirisi, doğal dil işleme gibi diğer dizi modelleme çalışmalarının yanında zaman serilerine olan ilginin

de arttığı söylenebilir. Bir zaman serisi, temel olarak zamana göre indekslenen bir dizi veri noktasıdır. Zaman serisi analizi, bir dizinin temel yapısını doğru bir şekilde karakterize edebilen geçmiş gözlemlerden faydalı istatistiksel özellikler ve kavrayışlar çıkarma amacına sahiptir [3].

Derin Öğrenme Algoritmaları, doğrusal olmayan ilişkileri modellemek için ara katmanları kullanarak girdileri ve çıktıları ilişkilendiren sinir ağlarını kullanır. Bir katmandaki her birim, verilerin belirli bir temsilini kullanır. Zaman serisi verileri için girdi katmanı sayısal değerlerin bir vektörüne veya yardımcı verileri içeren bir matrise karşılık gelebilir.

1.3 Tezin Amacı

Tez çalışmasının amacı, zaman serisi olarak modellenen finansal veri kümeleri kullanarak hisse senedinin fiyat hareketlerinin yön tahmini için derin sinir ağı tahmin yöntemlerini oluşturmak, test etmek ve uygulamak için gerekli olan teorik temelleri ve pratik adımları sunmaktır. Seçilen sinir ağı algoritmaları finansal zaman serilerinin tahmini için yapılan modellemelerde umut verici sonuçlar gösteren yinelemeli sinir ağı algoritmalarından LSTM ve GRU yöntemleridir. Bununla beraber finansal veri kümeleri üzerinde yapılacak boyut indirgeme yaklaşımları özelinde veri ön işleme süreçlerinin ve finansal verinin segmentlere ayrılması aşamasında kullanılan zaman serisi pencere aralıklarının yinelemeli sinir ağı modellerinin tahmin başarımlarının performansı üzerindeki karşılaştırmalı etki analizlerinin yapılmasıdır. Veri işleme sürecinde kullanılan boyut indirgeme yöntemleri ham veri ile yapılan çalışmalara göre daha iyi performans çıktıları üretebilen, hesaplama maliyetini kabul edilebilir ölçülerde azaltan ve modellemeyi daha anlaşılır hale getiren HSIC Lasso, mRMR ve RFE yöntemleridir. Çalışmada derin sinir ağlarının girdileri segmentlenirken kayan pencereler yönteminden yararlanılmış ve gelecek hareketlerin efektif olarak tahmin edilebilmesi için uygun geri bakma aralıkları tespit edilmeye çalışılmıştır. Tezin amacı doğrultusunda LSTM ve GRU modelleri üzerinde hiper parametre ayarlamaları literatür temel alınarak yapılmış ve model testleri sırasında sabit tutulmuştur. Çalışmanın sonunda 1240 sıralı piyasa işlem günü verisiyle eğitilen modeller kullanılarak 30 gün olarak tanımlanan test tahmin uzayı üzerinde AKBNK, TCELL ve FROTO BİST kodlu hisse senetlerinin yön hareketlerini inceleyen toplamda 96 senaryonun performans karşılaştırılması yapılmıştır.

1.4 Tez Akışı

Çalışmanın ikinci bölümünde kullanılan yöntemler özelinde genel metodoloji ile ilgili bilgi verilmiştir. Zaman serileri bağlamında kullanılan kavramlar, karşılaştırmalı performans analizinde kullanılmış derin sinir ağı model mimarileri, finansal veri kümesine uygulanmış boyut indirgeme yöntemleri ele alınmıştır. Çalışmanın üçüncü bölümünde ise kullanılan veri kümesi açıklanmış, etiketleme stratejileri ortaya koyulmuş ve verinin hangi yöntemlerle diziler olarak segmentlere ayrıldığı aşamalarından bahsedilmiştir. Buna ek olarak çalışmalarda yararlanılan yazılımlar, tahmin performanslarının yorumlanmasında kullanılan metrikler hakkında bilgi verilmiştir. Dördüncü bölümde ise derin öğrenme modellerinin AKBNK, TCELL ve FROTO hisse senedi yön hareketlerinin tahminlemesi için kullanılan teknikler ile elde edilen tahmin çıktıları, performans metriklerine göre karşılaştırılmıştır. Karşılaştırma kapsamında dizileme ve özellik seçim yöntemlerinin model başarımlarına etkisi de ayrıntılı olarak ortaya koyulmuştur.

1.5 Literatür Araştırması

Finansal zaman serisi tahmin problemi literatürde hisse senedi fiyat tahmini, endeks tahmini, emtia fiyat tahmini, kripto paraların fiyat tahmini, tahvil hareketleri gibi çeşitli alt başlıklarda incelenmiştir. Çalışmamızda Borsa İstanbul'da işlem gören farklı sektörlerde faaliyet gösteren Akbank, Turkcell ve Ford Otosan'ın sırasıyla AKBNK, TCELL ve FROTO kodlu hisse senetlerinin derin öğrenme yöntemleri kullanılarak hareketlerinin tahmin edilmesi amaçlanmıştır.

Finansal piyasalarda işlem gören varlıkların gelecekteki fiyat hareketlerinin tahmini için tasarlanan sınıflandırma algoritmalarında modelin girdisi olarak çoğunlukla makroekonomik ve mikroekonomik faktörler ele alınmaktadır. Mikroekonomik faktörler kurum seviyesinde kendini gösterip değerlendirilirken, makroekonomik faktörler ise tüm ekonomiyi ilgilendiren petrol fiyatı, altın fiyatı, döviz kuru, faiz oranı, para arzı, tüketici fiyat endeksi, cari açık ve ödemeler dengesi gibi göstergeler üzerinden ele alınmaktadır [4].

Geleneksel hisse senedi fiyat tahmin modelleri, otoregresif hareketli ortalama modelleri [5, 6, 7], çoklu doğrusal regresyon modelleri ve üstel yumuşatma modeli [8, 9, 10] dahil olmak üzere temel olarak doğrusal modellerdir. Bununla birlikte, bu

doğrusal modeller finansal tahmin çalışmalarının ilerlemesini ve gelişimini teşvik etmede önemli bir rol oynamaktadır. Derin öğrenme yöntemlerinin gelişmesiyle birlikte, daha yüksek doğruluk ile hisse senedi fiyatını tahmin etmek için doğrusal olmayan sinir ağları giderek daha fazla kullanılmaktadır. Bununla birlikte hem veri işleme süreçlerinde hem de model katmanında doğrusal ve doğrusal olmayan yöntemleri birlikte kullanarak hibrit çalışmalar sunan araştırmacıların sayısı da azımsanmayacak kadar fazladır.

Yineleyen sinir ağları dizi tahmin problemlerinde sıra bağımlılığını öğrenebilen tekrarlayan sinir ağlarıdır. RNN temelli LSTM ve GRU modelleri bu tür karmaşık zaman serisi problemlerini çözmek için uygun özellikleri nedeniyle finansal zaman serilerinde yüksek performans göstermektedir. Geçmiş örüntüleri saklayabilen bu ağ mimarileri hisse senedi fiyat tahmininde işlevseldir; çünkü geçmiş örüntüler hisse senetlerinin gelecekteki fiyatlarını tahmin etmede önemli bir rol oynamaktadır [11].

Fama tarafından 1970 yılında önerilen Etkin Piyasa Hipotezi'ne göre mevcut fiyatlar borsadaki tüm bilgileri yansıtmaktadır [12]. Ayrıca Fama, hisse senedi fiyatlarının rastgele yürüyüşler izlediğini savunmuştur [13]. Bu iki teori, hisse senedi fiyat değişikliklerinin bağımsız olduğunu, piyasa fiyat dalgalanmalarının rastgele bir yürüyüş karakteristiğine sahip olduğunu, geçmiş bilgilerin piyasa davranışlarını tahmin etmede kullanılamayacağını ve piyasadan düzenli bir gelir elde edilemeyeceğini savunmaktadır [14]. Ancak literatürde makine öğrenmesini temel almış birçok çalışma bu doğrultudan farklı görüşler önermektedir.

Henrique ve arkadaşlarının makine öğrenmesi yöntemleriyle finansal marketlerin tahmini konusunda yaptığı literatür çalışmasında [15] incelenen 57 akademik çalışmada tahmin yöntemi olarak %70'e yakın sinir ağları kullanıldığı belirtilmiştir. Tahmin çeşidine göre yön tahmini yapan sınıflandırma yöntemlerinde performans göstergelerinden en çok doğruluk ve kümülatif kazanç toplamı, regresyon yöntemlerinde ise ortalama mutlak hata ve ortalama karekök sapması kullanıldığı görülmektedir.

Namini ve Namin yaptıkları çalışmada [16] hisse senedi yön tahmini problemi için LSTM ve ARIMA modellerini karşılaştırmışlardır. LSTM daha düşük tahmin hatası ve daha yüksek doğruluk ile ARIMA'ya performans olarak üstünlük sağlamıştır.

Zaman serisi tahmininde derin öğrenme yöntemleri ile geleneksel yöntemlerin gösterdiği performansın üzerinde sonuç almaları önemlidir.

Ghosh ve arkadaşları yaptıkları çalışmada [17] 1993-2018 yılları arasındaki veri kümesini kullanarak RF ve LSTM ağlarını S&P 500 hisselerinin yön tahmini için kullanmışlardır. Kurdukları finansal işlem stratejisi ile her işlem gününde model çıktılarına göre gün içi getiri medyanını temel alarak bu değeri geçmesi en düşük olasılıklı 10 hisseyi satıp en yüksek olasılıklı 10 hisseyi alarak endeks getirisinden daha karlı bir sonuca ulaşmışlardır. Günlük kazanç ortalamasında LSTM ağı RF'e göre daha iyi performans göstermiştir.

Literatürde birçok teknik gösterge, hisse senedi tahmin modellerine girdi özellikleri olarak kullanılmıştır, ancak teknik göstergelerin kalitesi ihmal edilebilen bir konu olabilmektedir, bu nedenle hisse senedi tahmin problemlerinde özellik mühendisliği uygulamalarının daha da genişletilmesi faydalı olacaktır. Bu doğrultuda Ji ve arkadaşları [18] yaptıkları çalışmada özellik kümesine uyguladıkları özellik seçim stratejileri ile SSEC, HSI, DJI ve S&P 500 endekslerindeki tahmin performans skorlarını artırarak daha az özellik ile daha yüksek tahmin doğrulukları elde edilebileceğini göstermişlerdir. Buna ek olarak ise zaman serilerini oluştururken kullandıkları geriye bakma değerlerini uygun bir şekilde artırmanın tahmin performansına pozitif etki verebileceğini de göstermişlerdir.

Enke ve arkadaşları ise temel bileşenler analizi (principal component analysis) ile boyut indirgeme yaptıkları veri kümesinin derin sinir ağlarında ham verinin kullanıldığı temel modele göre daha iyi doğruluk ile tahmin performansı gösterdiğini ortaya koymuştur [19].

Hamiche ve Moghar [20] RNN özellikle de LSTM mimarisi ile elde edilecek performans sonuçlarının hiper parametreler ile ne ölçüde iyileştirilebileceğini üzerine çalışmışlardır. Farklı veri kümelerinde yaptıkları gözlemlerde daha az veri ve daha yüksek sayıda seçilen iterasyon değerlerinin test performansını iyileştirdiğini belirtmişlerdir.

Budiharto hisse senetlerinin fiyat tahmini problemi için LSTM mimarisini kullandığı çalışmasında [21] veri kümesi üzerinde kayan pencereler yönteminin kullanıldığı tahmin metodolojisi önermiş ve farklı büyüklükteki eğitim kümelerinin tahmin performansına etkisini araştırmıştır.

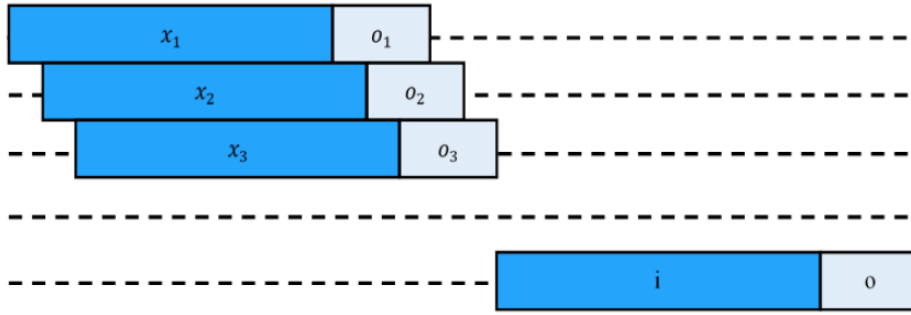
Kwon ve Sun yapay sinir ağılarından yararlandıkları çalışmalarında [22] hisse senedi işlem strateji etkinliğini artırmak için orijinal zaman serilerini parçalı doğrusal segmentlere dönüştürmüşler ve sinir ağına girdi olarak bu verileri vermişlerdir. Çalışma sonucunda önerdikleri model stratejisinin al-tut stratejisine göre daha karlı sonuç verdiğini göstermişlerdir. Model stratejisi ANN çıktı katmanındaki olasılık çıktısında 0.5 ve üstünü al sinyali olarak altını ise sat sinyali olarak değerlendirmektedir. Al-Tut stratejisi ise ilk test işlem günü için hissenin alındığı ve tüm test dönemi boyunca hissenin elde tutulmasıdır.

Tsiamas ve Borovkova'nın çeşitli teknik indikatörün ağ girdisi olarak kullanıldığı çalışmalarında [23] topluluk yöntemiyle uyguladıkları LSTM yapısının performansı LASSO ve Ridge lojistik regresyon modelleriyle AUC performans ölçütü alınarak karşılaştırılmıştır. Yıllık işlem verisini 5'er dakikalık aralıklara bölerek her hisse senedi için yaklaşık 19000 gözlem noktası oluşturmuşlar ve hem sektörel hem de zaman serisi bazlı teknik indikatörler yaratmışlardır. Hisse senedinin yön tahmini için özellik uzayına ilgili şirketin rekabet içinde olduğu ana şirketin özellik uzayını da ekleyerek performansı iyileştirmeye odaklanmışlardır. Her zaman adımında hisseleri fiyat yönüne göre al ve sat olarak etiketlenmişlerdir. Veri segmentlemesi kayan pencereler yöntemi ile yapılmış ve çalışma sonucunda önerilen yapının al sat stratejileri için kullanılabileceği ve lojistik regresyon modellerinden daha iyi sonuç verdiğini göstermişlerdir.

Veri kümesi üzerinde boyut indirgemeye odaklanmış bir diğer çalışmada [24] ise Gao ve arkadaşları LSTM ve GRU modellerini kullanarak piyasa tahmini yapmışlar ve çeşitli parametreler altında LSTM ve GRU modellerinin performanslarını karşılaştırmışlardır. Çalışma çıktılarında LSTM ve GRU'nun tahmin performansı olarak birbirine yakın sonuçlar verdiği, her iki model için de LASSO özellik seçimi ile entegre edilen yaklaşımın PCA ile özellik çıkarımı yapılan senaryolardan daha efektif olduğunu belirtmişlerdir. Sonuçları farklı geri bakma değerleri ve model hiper parametreleri altında değerlendirmişler ve temel olarak boyut indirgeme yöntemleri ile veri kümesi boyutu azaltılmasına rağmen daha yüksek performans çıktılarının alınabileceğini göstermişlerdir.

Faurina ve arkadaşları da [25] bu doğrultuda kurdukları hisse senedi fiyatlarının yön tahmin modelini farklı boyut indirgeme yöntemleriyle kurgulamışlardır. Model kurulurken ilk aşamada Lasso, Elastic Net ve PCA boyut indirgeme yöntemleri veri

kümesine uygulanmıştır. İkinci aşamada indirgenmiş veri kümesinde girdi ve çıktılar Şekil 1.1’deki gibi kayan pencereler yöntemi ile zaman dizilerine ayrılmıştır. x_1 olarak segmentlenen dilim t adımımdan $t+9$ adımına kadar 10 günlük verileri içeren zaman adımlarının hepsidir. Her bir satır tek bir eğitim verisi olarak düşünüldüğünde ağa beslenecek ikinci veri de $t+1$ anından $t+10$ anına kadar olan verileri tutacaktır. Son aşamada da oluşturulmuş zaman dizileri LSTM ağına beslenerek yön tahmini için bir sınıflandırma yapılmıştır. PCA-LSTM yöntemi diğer tasarımlara göre daha yüksek doğrulukla sonuç vermiştir.



Şekil 1.1 : Zaman serisi segmentlere ayırma işlemi.

Hiransha ve arkadaşları da [26] NSE ve NYSE’de işlem gören çeşitli şirketlerin hisse senedi fiyat verileri üzerinden farklı geri bakma değerleri için LSTM modelinde performans analizlerini ortaya koymuşlardır. 1996-2015 arasında 4861 işlem gününü eğitim kümesi olarak kullanarak 50, 100, 150, 200, 250 geri bakma değerleri ile segmentlere ayırdıkları zaman serisini gelecekteki belirli bir aralıkta hisse senedi kapanış fiyatlarını tahmin etmek için kullanmışlardır. Bu aralığı 10, 20, 30 ve 40 gün olarak ayrı ayrı test etmişler ve model performans çıktılarında en düşük ortalama mutlak yüzde hatasının (MAPE) çerçeve aralığı olarak 200 kullandıklarında ve gelecek 10 günü tahmin ettiklerinde olduğunu belirtmişlerdir.

Liu ve arkadaşlarının yaptığı çalışmada [27] hisse senedi endeks yönünü tahminlemek için veri ön işleme süreçlerine odaklanılmış ve trend bazlı teknik indikatörlerin kullanıldığı bir veri hazırlama süreci oluşturulmuştur. Sonrasında orijinal özellik kümesi ve oluşturulan diğer girdi kümeleri farklı derinliğe ve nöron sayısına sahip RNN, LSTM ve GRU modellerine girdi olarak verilerek performans karşılaştırılması yapılmıştır. Çalışmanın sonucunda 3 katmanlı ve ön işleme ile oluşturulmuş veri

kümesi ile GRU modelinin diğer kombinasyonlara göre daha iyi doğruluk ile çıktı ürettiği belirtilmiştir.

LSTM ve GRU modellerinin karşılaştırmasını da ele alan Shah ve arkadaşlarının yaptığı çalışmada [28] gerçek zamanlı piyasa verisi birer dakikalık aralıklarla işlenmiş ve yatırımcıya gönderilen al ve sat sinyallerinin performansı doğruluk ve toplam parasal getiri ile ölçülmüştür.

Finansal piyasa tahmininde veri kümesi üzerinde çok aşamalı boyut indirgeme yaklaşımlarının incelendiği çalışmalar [29,30,31] tahmin başarımının özellik seçim stratejilerinin geliştirilmesi ve model tasarımında kullanıldığı fazların önemine dikkat çekmektedir.

mRMR, HSIC Lasso ve RFE özellik seçim yöntemleri ile makine öğrenmesi modellerinin performans çıktılarının iyileştirilmesini konu alan incelenen diğer çalışmalar [32,33,34,35] ilgili yöntemlerin tahmin çalışmalarında veri boyut karmaşıklığını azaltıp hedef değişken ile doğru bağlantıların seçilerek model performansını direkt olarak olumlu yönde etkilediğini göstermektedir.



2. METODOLOJİ

Bu bölümde zaman serileri bağlamında kullanılan kavramlar, sinir ağı model mimarileri, tez çalışmasında veri setine uygulanmış boyut indirgeme yöntemleri ve performans metrikleri tanıtılacaktır.

2.1 Zaman Serileri

Bir zaman serisi, temel olarak zamana göre indekslenen bir dizi veri noktasıdır. Bir değişken, örnekleme aralığı olarak bilinen sabit bir aralıkta veya zaman içinde sırayla ölçüldüğünde, elde edilen veriler bir zaman serisi oluşturur [36]. Herhangi bir zaman noktasındaki tek bir veri noktasına zaman adımı denir. Zaman serilerine örnek olarak, bir evdeki nem ve sıcaklık değerlerini belirli zaman adımlarında kaydeden bir sensörden alınan veriler, her an resim yakalayan bir kameradan oluşan bir video kaydı veya bir şirketin her dakika kayıt altına alınan hisse senedi fiyatları verilebilir. Farklı zaman serisi verilerini ayırt etmenin önemli bir yolu, her bir veri noktasının boyutsallığıdır.

Tek değişkenli zaman serileri yalnızca bir değişken içeren veri serisi olarak tanımlanırken; çok değişkenli zaman serileri ise iki veya daha fazla değişken içeren veri serisi olarak tanımlanır. Yapılan çalışmada kullanılan veri kümesi çok değişkenli bir zaman serisidir.

Zaman serisi tahmini, gelecekteki zaman adımlarının verilerini tahmin etme sürecidir. Tarihsel verileri işlemek veya analiz etmek ve bu öğrenilen bilgileri gelecek hakkında tahminlerde bulunmak için kullanmak için bir mekanizma olarak düşünülebilir. Bir şirketin hisse senedinin ertesi gün yönünün tahmin edilmesi bu anlamda örnek olarak verilebilir. Zaman serisi tahmini temel olarak iki tipte sınıflandırılır. Zaman serisinin bir sonraki zaman adımını tahmin etmeyi amaçlayan yaklaşımlar ve zaman serisinin bir dizi zaman adımının öngörülmesini hedefleyen yaklaşımlar bu sınıflandırmanın unsurlarıdır.

Zaman serilerinin modellenmesi dinamik bir araştırma alanıdır. Zaman serisi modellemenin temel amacı, seri bağlamını uygun bir şekilde tanımlayan modellerin geliştirilmesi için zaman serisinin geçmiş gözlemlerini dikkatli bir şekilde toplamak ve titizlikle incelemektir. Geliştirilen model daha sonra seri için gelecek değerleri oluşturmak, yani tahminler yapmak için kullanılır. Bu nedenle zaman serisi tahmini, geçmişini anlayarak geleceği tahmin etme işlemi olarak adlandırılabilir [37]. Gerçek dünyadaki birçok çalışma alanında zaman serilerinin tahmininin yadsınamaz önemi sebebiyle tahmin doğruluğunu iyileştiren verimli modellerin geliştirilmesi araştırmacılar arasında uzun yıllar boyunca popüler konu olarak yer edinmiş ve literatürde birçok zaman serisi tahmin modeli geliştirilmiştir [38].

Literatürde sıklıkla kendine yer etmiş zaman serilerinin modellemesinde kullanılan istatistiksel yöntemlere ARIMA, AR, MA ve ARMA modelleri örnek olarak verilebilir. Altta yatan varsayım zaman serisinin lineer olduğu ve belirli bir istatistiksel dağılımı takip ettiğidir; ancak bu varsayımlar gerçek veri dünyasında karşımıza çıkan birçok zaman serisi probleminin çözümünde veriler üzerindeki lineer olmayan ilişkisel bağları ortaya çıkarabilme özelinde yetersiz kalmıştır. Bu bağlamda doğrusal olmayan veri modellemesi için ARCH, GARCH, TAR, NAR vb. stokastik yaklaşımlar önerilerek çalışmalar yapılmıştır. Zaman serileri karakteristiğinin kavranabilmesine çözüm arayışının sonucunda yapay sinir ağları zaman serilerinin modellemesinde artan ilgiyi üzerine çekmiştir. Yapay sinir ağlarının yeteneklerinden biri gözlem verilerinin izlediği dağılım ve trendler hakkında varsayım olmaksızın doğrusal olmayan modelleme yeteneklerine sahip olmasıdır. Son yıllardaki araştırmalarının çoğunun da temelinde yapay sinir ağı mimarilerini kullanan modeller üzerinde yapıldığı görülmektedir. Aralarında en yaygın olarak kullanılan yöntemlerden ilki FNN ile karakterize edilen MLP'lerdir. Zaman serilerinin modellenmesinde yaygın olarak üzerinde çalışılmış diğer bir model ise Vapnik ve arkadaşlarının ortaya koyduğu SVM yöntemidir. Ortaya konma amacı başlarda örüntü sınıflandırma problemlerine çözüm arayışı olmasına rağmen sonrasında regresyon, sinyal işleme, zaman serisi tahmini gibi birçok alanda uygulanmıştır. Eğitim verilerinin daha iyi genelleştirilmesi için ortaya koyduğu yaklaşımla literatürde iyi bilinen tekniklerden biri haline gelmiştir. SVM uygulamalarında girdi uzayı çekirdek fonksiyonları yardımıyla yüksek boyutlu bir özellik uzayına eşlenir ve bu yaklaşımla yüksek boyutlarda bile iyi bir genelleme sağlanarak lineer olmayan problemlerde de iyi sonuçlar verebilir [38].

2.2 Yapay Sinir Ağları

Yapay sinir ağlarının tarihine bakıldığında çok sayıda kapsamlı faaliyet dönemi görülmektedir. YSA araştırmalarındaki ilk tepe noktasının 1940'larda McCulloch ve Pitts'in nöron sinirsel aktivitelerin mantıksal hesaplamalarının yapılması konusundaki öncü çalışması olduğu söylenebilir [36]. İkinci tepe noktası ise Rosenblatt'ın yaptığı çalışmalarıdır. Rosenblatt 1958'de perceptron algoritmasını [39] tanıttı ve 1960'larda perceptron yakınsama teoremini ortaya koydu. Perceptron, bilinen ilk yapay sinir ağıdır ve girdi verilerini iki sınıfa bölebilir. Algoritma, girdilerin ağırlıklı toplamını hesaplar ve ardından birim basamak eşik fonksiyonu kullanır. Girdilerin ağırlıklarının toplamı belirli bir eşikten büyükse girdi “a” sınıfına, daha düşükse “b” sınıfına aittir.

Sonraki dönemde Minsky ve Papert'in yaptığı çalışma [40] tek katmanlı basit bir perceptronun limitlerini ve mantıksal XOR işlemlerini gerçekleştiremeyeceğini gösterdi. Bu yayın sonrasında literatürde yapay sinir ağlarına olan ilginin azaldığı ve yapılan çalışmalarda bir süre duraklama dönemine geçildiği söylenebilir. Werbos [41] tarafından önerilen ve Rumelhart ve arkadaşları [42] tarafından popüler hale getirilen çok katmanlı algılayıcılar için geri yayımlı öğrenme algoritması 1980'lerde YSA'lara olan ilginin yeniden canlanmasını sağladı. Çok katmanlı algılayıcıların XOR problemini çözebilmesinin yanında, ağırlık hızlı bir şekilde eğitilmesinin önünü açan geri yayılım algoritması sinir ağları için büyük bir ilerleme olmuştur. Zamanla gelişen bilgisayar depolama ve hesaplama birimleri ile de edinilen büyük veriyi işleyebilecek ve üzerinde karmaşık hesaplamaları yapabilecek derin sinir ağlarının önü açılmıştır. Ağların dağıtık mimarisine daha çok uyan grafik işlemcilerin yaygınlaşmasının da 2000'lerde bu alanda ivmeyi arttırdığı söylenebilir. Hesaplama ve depolama gücüne ek olarak geliştirilen sinir ağlarındaki aşırı öğrenme problemlerine çözüm arayan veri regülarizasyon teknikleri, ağı doğrusal olmayan bir karakteristik veren farklı aktivasyon fonksiyonları gibi performans iyileştirici tekniklerin de bilimsel ilerlemedeki etkisi göz ardı edilemez [43]. 1991 yılında Hochreiter tarafından dizilerin modellenmesi için sunulan yinelemeli sinir ağları öncü çalışmalar arasında yer almıştır. Bu mimaride uzun girdi dizileri için yaşanan bilgi kaybı problemine 1997'de sunulan [44] uzun kısa dönem bellekli sinir ağları (LSTM) ile çözüm bulunmuştur. [45]'de LeCun ve arkadaşları tarafından ortaya konan çalışma ile de konvolüsyonel sinir ağları

(CNN) görüntü ve ses işlemede literatürde sıklıkla kullanılan mimariler arasında kendine yer bulmuştur.

Yapay sinir ağları, insan beyninin çalışma mekanizmasından esinlenerek, makinelere öğrenme, hatırlama, genelleme, yeni bilgiler türetebilme gibi fonksiyonları kazandırmaya odaklanan sentetik modellerdir. Beyin temel olarak deneyimlerinden öğrenir. Yapay sinir ağları da geleneksel matematiksel modellerin aksine girdi ve çıktılar arasındaki ilişkileri deneyimlerinden öğrenmeye çalışır. Bir sinir ağı, deneysel bilgileri depolamak ve kullanıma hazır hale getirmek için doğal bir eğilime sahip basit işlem birimlerinden oluşan büyük ölçüde paralel dağıtılmış bir işlemcidir. Bilgiyi bellekte belirli bir yerde depolayan geleneksel bilgisayar sistemlerinin aksine bilgiyi tüm sinir ağı boyunca dağıtır. Bilgi ağı tarafından bir öğrenme süreci ile elde edilir ve sinaptik ağırlıklar olarak bilinen nöronlar arasındaki bağlantılar edinilen bilgileri depolamak için kullanılır. Ağı öğrenme sürecini gerçekleştirmek için kullanılan öğrenme algoritması, istenen çıktı hedefine ulaşmak amacıyla ağı sinaptik ağırlıklarını düzenli olarak günceller [46]. İterasyon başlangıcında rastgele seçilen bağlantı ağırlıkları ağı çıktısındaki hata değerini istenilen seviyelere getirmek için güncellenir. Bağlantılardaki ağırlıkların güncellenmesi işleminde geriye doğru hatanın her bir ağırlık değerine göre türevi alınarak bağlantıların ağı çıktısında elde edilen hata değerine etkisi hesaplanarak sinaptik ağırlıklar artırılır ya da azaltılır. Sonraki iterasyonda da aynı süreç tekrarlanarak ağı kayıp fonksiyonundaki hata değerinin azaltılması amaçlanır.

Yapay sinir ağlarında nöronlar üç temel unsur tarafından inşa edilmiştir.

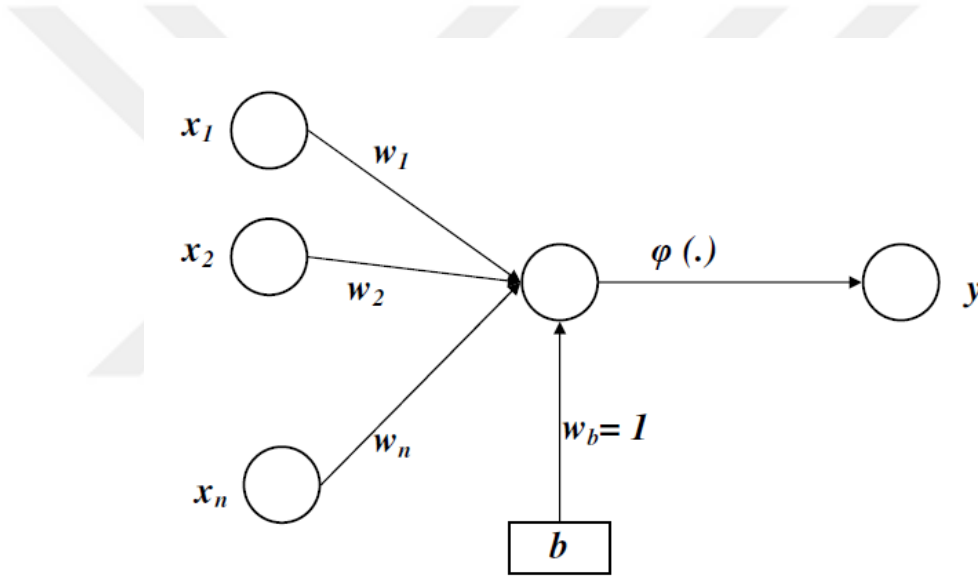
1. Her biri kendi ağırlığı ile karakterize edilen bir dizi bağlantı noktası
2. Nöronların ilgili bağlantı noktaları ile ağırlıklandırılan giriş sinyallerini toplamak için lineer birleştirici
3. Nöronun çıktı genliğini sınırlamak için bir aktivasyon fonksiyonu

Bir nöron matematiksel olarak aşağıdaki şekilde gösterilebilir.

$$u_k = \sum_{j=1}^n w_j x_j \quad (2.1)$$

$$y_k = \phi(u_k + b_k) \quad (2.2)$$

$x_1, x_2, \dots, x_n$ giriş sinyallerini, $w_1, w_2, \dots, w_n$ nöronu için sinaptik ağırlıkları, u_k giriş sinyalleri için lineer birleştiriciyi, u_k bias değerini, $\phi(.)$ aktivasyon fonksiyonunu ve y_k nöronun çıkış sinyalini göstermektedir. Yapay sinir ağlarındaki temel amaç model çıktısının en iyi skoru üreteceği w ve b değerlerini hesaplamaktır. Fonksiyonel diyagram Şekil 2.1'deki gibi gösterilebilir.



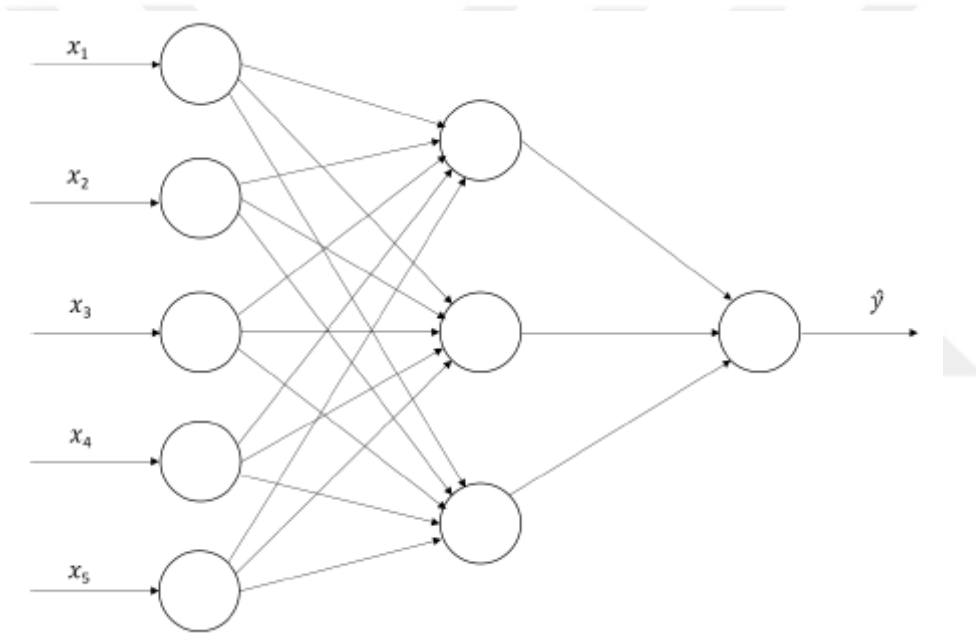
Şekil 2.1 : Yapay sinir ağlarında nöron modeli.

$\phi(.)$ aktivasyon fonksiyonları lineer birleştiriciden gelen verilerin nasıl bir değişimden geçmesi gerektiğini belirleyen, doğrusal olmayan fonksiyonlardır. Model, aktivasyon fonksiyonları ile basit bir doğrusal fonksiyon olmaktan çıkar ve karmaşık gerçek dünya verilerini öğrenebilecek kapasiteye ulaşabilir. Literatürde genellikle sigmoid kökenli aktivasyon fonksiyonları kullanılmaktadır.

YSA'lar yapılarına göre genel olarak ileri beslemeli ve geri beslemeli ağlar olarak sınıflandırılabilirler.

2.2.1 İleri Beslemeli Sinir Ağları

Tek katmanlı algılayıcının uzantısı olarak sadece doğrusal olarak ayrılabilir desenleri sınıflandırabilen tek katmanlı algılayıcının limitlerinin üstesinden gelmek için tasarlanmış çok katmanlı algılayıcıdır. Aynı zamanda ileri beslemeli sinir ağı olarak da adlandırılırlar, çünkü bilgi giriş katmanından çıkışa akar ve aynı katmanlar arasında bağlantılar bulunmaz. Nöron çıktılarından geri besleme bağlantısı yoktur. Şekil 2.2’de görüleceği üzere ileri beslemeli mimari en az bir gizli nöron katmanına sahiptir. Zincirlenmiş katmanların toplam uzunluğu, derin öğrenmeye de adını veren ağ modelinin derinliğini temsil eder, ancak bir modelin “derin” olarak sınıflandırılması için literatürde herhangi bir düşünce birliğinin olmadığını da belirtmek gerekir.



Şekil 2.2 : Tek gizli katmanlı ileri beslemeli ağ mimarisi.

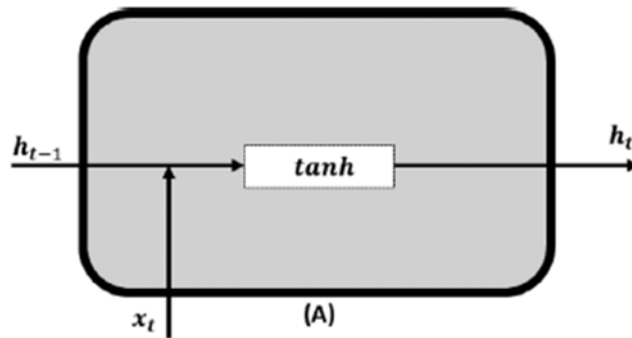
2.2.2 Yinelemeli Sinir Ağları

RNN, metin, konuşma, borsa, DNA dizisi, vb. gibi makine öğrenimi kullanım alanlarında genellikle uygulanabilir olan, sabit ve değişken uzunluktaki zaman dizilerini işleme konusunda yeteneklidir. Parametreleri ağ boyunca paylaşmak veri dizilerinin genelleştirilmesine izin verir. RNN’nin bu tasarım yönü, veri dizisini sağlayan zaman serileri için uygulanabilir. Ayrıca, bir değişkenin mevcut değerinin

gelecekteki bir zaman adımında değişken değerini etkileyebileceği finansal zaman serileri için, RNN, ağ boyunca parametreleri paylaşarak diziler arasında bu tür örüntülerin bulunmasına izin verir. Geleneksel ileri beslemeli ağ, her bağımsız giriş özelliği için ayrı parametrelere sahipken, yinelemeli sinir ağı aynı parametreleri birbirinden farklı zaman adımlarında paylaşır.

$t-1$ anındaki girdi için verilen karar, t anında verilecek olan kararı da etkilemektedir. Yinelemeli sinir ağlarında girdiler şimdiki ve önceki bilgilerin birleştirilmesi ile çıktı üretirler. İleri beslemeli sinir ağlarında her girdi vektörünün birbirinden bağımsız olduğu varsayımı yapılır. Her yeni girdi, yalnızca girdi vektöründen gelen bilgilere dayanan bir çıktı üretir, daha önce verilen girdi vektörleri hakkında hiçbir bilgi veya nöronların herhangi bir kalıntı aktivasyonu ağ üzerinde kaydedilmez.

Yinelemeli sinir ağları, 1980'lerde önerilen vektör dizileri üzerinde çalışmamıza izin veren yapay sinir ağı modelleridir. Metin oluşturma, makine çevirisi, konuşma tanıma, video sınıflandırma, görüntü alt yazılama veya finansal tahmin problemlerinde uygulanabilirler. İleri beslemeli sinir ağları ve yinelemeli sinir ağları arasındaki en önemli fark, yinelemeli sinir ağlarının önceki girdilerden hesaplanan bilgileri depoladıkları bir belleğe sahip olmalarıdır, yani son çıktı yalnızca önceki girdiden değil, ağa beslenmiş tüm girdilerden etkilenir [47]. Çıkış aktivasyonu gelecek zaman adımında katmana girdi olarak verilerek veri kümesi üzerindeki zaman bağımlılıkları modellenmeye çalışılır. Temel RNN hücresi Şekil 2.3'te gösterilmiştir.

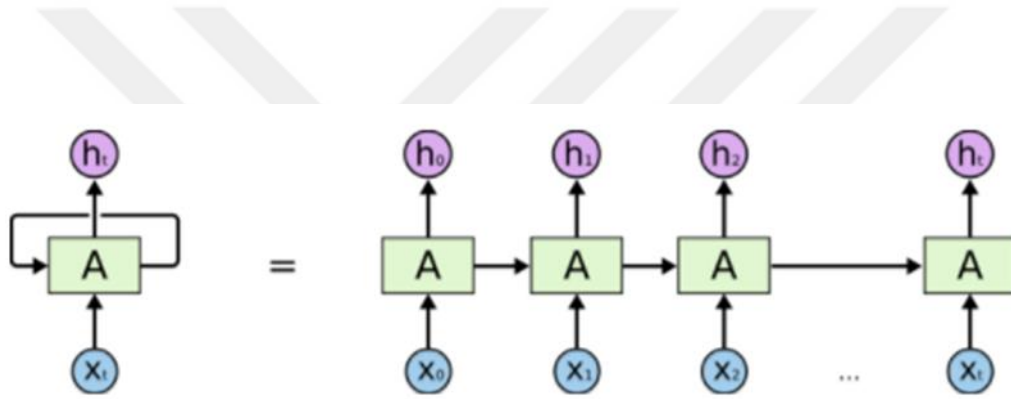


Şekil 2.3 : RNN hücresi.

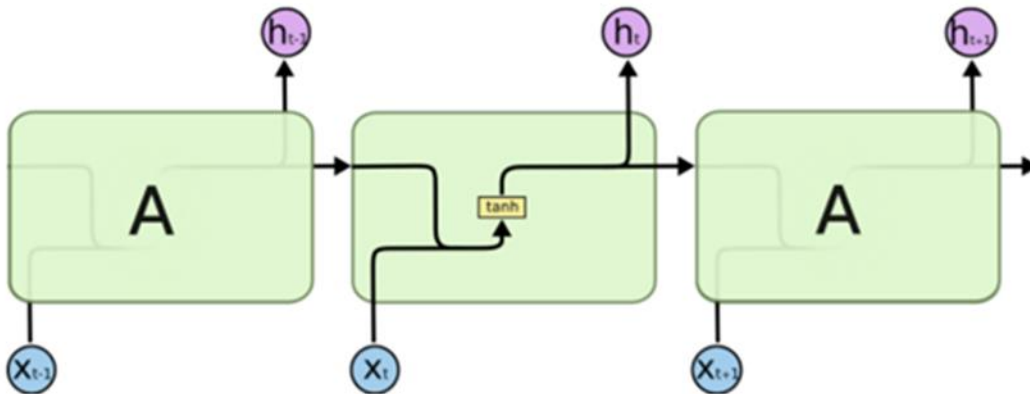
Bir RNN hücresinin matematiksel modellenmesi aşağıdaki gibi yapılabilir.

$$h_t = \tanh(W[h_{t-1}, x_t] + b) \quad (2.3)$$

Yukarıdaki eşitlikte W ağırlık matrisini, b bias vektörlerini göstermektedir. h_{t-1} ve h_t sırasıyla önceki zaman adımı ve şimdiki zaman adımıdaki gizli katman çıktılarıdır. Matematiksel model olarak ileri beslemeli sinir ağlarına çok benzer hesaplamalar gerçekleştirilir. Temelde RNN'deki her bir nöronda h_{t-1} ve x_t birleştirilip bir ağırlık matrisi ile çarpılarak bir bias vektörü eklenir. Çıktı hiperbolik tanjant fonksiyonundan geçirilerek h_t 'ye atanır. Nörondaki tanh fonksiyonu gerçek değerleri -1 ile 1 arasında kalacak şekilde ölçeklendirmek için kullanılan aktivasyon fonksiyonudur. RNN hücresinin t zaman serisi için açılımı Şekil 2.4'te, standart RNN ağı ise şekil 2.5'te gösterilmiştir.



Şekil 2.4 : RNN dizi açılımı.



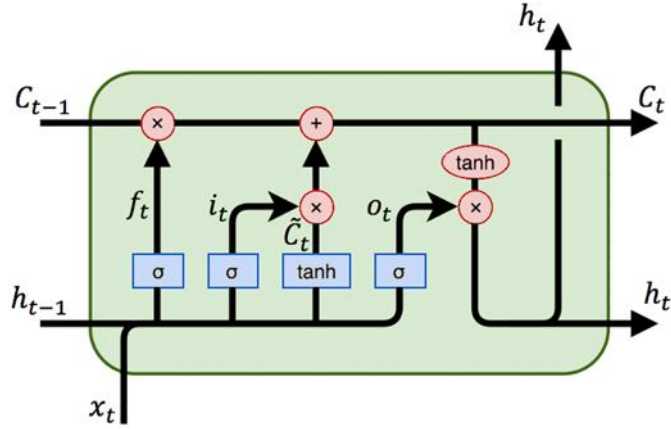
Şekil 2.5 : Tek bir katman içeren standart RNN ağı.

Ağın eğitilmesinde kullanılan geri yayılım algoritması geleneksel BP (back-propagation) algoritmasından biraz farklı olan zaman içinde geri yayılım (BPTT) olarak bilinen zamana bağlı bir dizi işlemin tümüne geri yayılım uygulanmasıdır. BPTT her zaman adımında hataları toplar ileri beslemeli ağlar ise her katmanda ağırlık parametrelerini paylaşmadıklarından dolayı buna ihtiyaç duymazlar. RNN mimarisinin ana dezavantajı uzun girdi dizilerini eğitirken gradyanlarda karşılaşılan zorluklarıdır. Bu yönüyle zaman serileri tahminlerinde belirli bir noktadan sonra eksik kalırlar ve öğrenemezler [48]. Ağ eğitilirken türevlerin geriye doğru hesaplanması aşamasında türevin yok olması (vanishing gradient) veya kontrolsüz olarak yüksek değerlere çıkması (exploding gradient) araştırmacıları yeni mimari arayışına itmiştir [49]. Hata eğrisi boyunca gradyan vektörü çok küçük olduğunda ağırlık parametreleri önemsiz hale gelene kadar küçülmeye devam eder. Bu gerçekleştiğinde model artık eğitilemez duruma gelmektedir. Gradyanlar çok büyüdüğünde de kararsız bir model oluşur ve ağırlıklar çok büyür. Buna bağlı olarak ağ yine öğrenememeye başlar. RNN mimarilerindeki bu sorunlara aktivasyon katmanında ReLu fonksiyonu kullanmak ya da gizli katmanların sayısını azaltarak karmaşıklığı indirmek gibi çözüm arayışları olmuştur. İlerleyen aşamada bu dezavantajların üstesinden gelebilmek için LSTM ve GRU mimarileri ortaya atılmıştır.

2.2.2.1 LSTM:

Sepp Hochreiter ve Juergen Schmidhuber tarafından [44] RNN'deki gradyan problemlerine bir çözüm olarak tanıtılan yinelemeli sinir ağı mimarisidir. İnsan beyninin sıralı verileri işleme şeklini taklit eder. Bir metin okuduğumuzu düşünürsek insan beyni aradığı bilgiyi bulduğunda metnin geri kalan kısımlarını genelde unutar. LSTM mimarileri de bunu içerdiği ek birimlerle yapabilir.

Şekil 2.6'da gösterilen LSTM yapılarında temel birime hafıza hücresi denir ve durumu c_t ile tutulur. c_t önceki dizi girdilerinin bilgilerini tutar ve hücre boyunca taşır. Giriş kapısı, bilgilerin ne olduğunu öğrenerek girdilerin hafıza hücresine akışını düzenler. Çıkış kapısı, hafıza hücresinin çıktılarının ağa akışını kontrol eder. Çıkış kapısı değeri tahmin için kullanılır ve zaman adımındaki durum değeri olarak bir sonraki katmana ilerler. Unutma kapısı ise bellekte hangi bilgilerin tutulup hangi bilgilerin tutulmayacağını denetler. RNN'lerde yaşanan gradyanların bir süre sonra çok azalarak ağ ağırlıklarını sıfırlaması durumunu giriş ve çıkış kapıları engeller.



Şekil 2.6 : LSTM hücresi.

LSTM katmanındaki kapılarda yapılan işlemler aşağıdaki şekilde matematiksel olarak formülize edilebilir.

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (2.4)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \quad (2.5)$$

$$g_t = \tanh(W_{xg}x_t + W_{hg}h_{t-1} + b_g) \quad (2.6)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (2.7)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (2.8)$$

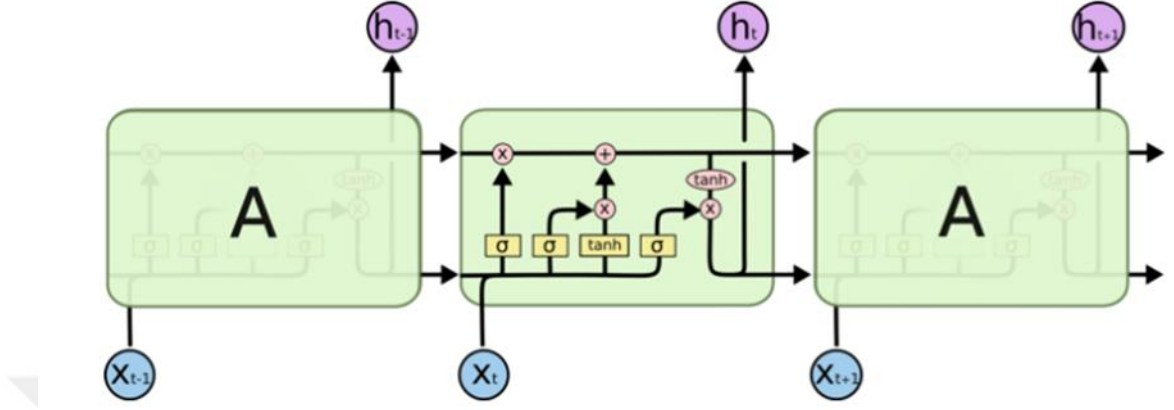
$$h_t = o_t \odot \tanh(c_t) \quad (2.9)$$

$$y_t = \phi(h_t) \quad (2.10)$$

Eşitliklerde W ağırlık matrislerini, b bias vektörlerini, x_t ve h_t giriş - çıkış dizilerini, i_t giriş kapısını, f_t unutma kapısını, o_t çıkış kapısını temsil etmektedir. g_t ile giriş ve bir önceki adımdaki durum $\tanh$ aktivasyonundan geçirilir. c_t ise LSTM hücresinin durum güncellemesini yapar. t zaman serisindeki adımları, σ sigmoid fonksiyonu, $\odot$ ise nokta tabanlı eleman çarpımını gösterir. $\sigma(\cdot)$ ve $\tanh(\cdot)$ aktivasyon fonksiyonları aşağıdaki gibi gösterilebilir.

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (2.11)$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.12)$$

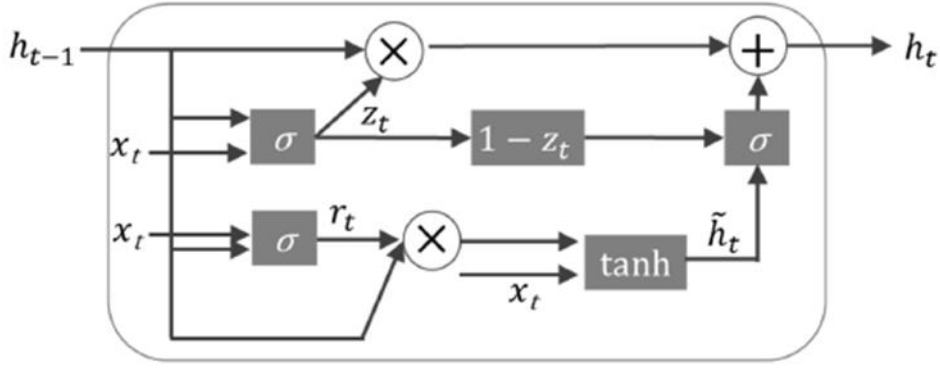


Şekil 2.7 : LSTM dizisi.

2.2.2.2 GRU:

Cho ve arkadaşları tarafından [50] 2014'te ortaya konan GRU'lar LSTM mimarisi üzerinde bazı değişiklikler yapmışlardır. Giriş kapısı ve unutma kapısı tek bir güncelleme kapısı altında birleştirilmiştir. Hücre durumunu ve gizli durumu birlikte modellemişlerdir. GRU'larda temel olarak sıfırlama kapısı ve güncellenme kapısı bulunur. Güncelleme kapısı hangi bilgilerin ağda saklanacağına veya ağdan çıkarılacağına buna ek olarak hangi yeni bilgilerin ağa iletileceğine karar vermekle görevlidir. Sıfırlama kapısı ise zaman serisinin önceki bilgilerinin ne kadarının ağda tutulacağını denetler. Kısa dönem bağlarını sıfırlama kapısı yakalarken, uzun dönem bağlarını güncelleme kapısı tutar. Kapılar sigmoid katmanlardır, girdileri sıfır ve bir arasına dönüştürürler.

Model LSTM mimarisinden daha basittir ve literatürde giderek daha popüler hale gelmektedir. Şekil 2.8'de gösterilen GRU mimari katmanlarında yapılan işlemler matematiksel olarak aşağıdaki gibi ifade edilebilir.



Şekil 2.8 : GRU mimarisi.

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (2.13)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (2.14)$$

$$\hat{h}_t = \tanh(W_h x_t + U_z (r_t \circ h_{t-1}) + b_z) \quad (2.15)$$

$$h_t = (1 - z_t) \circ h_{t-1} + z_t \circ \hat{h}_t \quad (2.16)$$

Eşitliklerde W , U ağırlık matrislerini, b bias vektörlerini, x_t ve h_t giriş çıkış dizilerini, z_t güncelleme kapısını, r_t sıfırlama kapısını temsil ederken “ $\circ$ ” işlemi Hadamard çarpımıdır [51].

2.3 Boyut İndirgeme

Büyük miktarda yüksek boyutlu verinin çok çeşitli alanlarda yaygınlaştığı büyük veri çağındaki verilerin hızlı büyümesi, etkili ve verimli veri yönetimi yapılabilmesi için zorlukları da beraberinde getirir. Farklı türlerdeki verilerden anlamlı bilgilerin edinilmesi için veri madenciliği ve makine öğrenimi tekniklerinin uygulanması gerekliliği büyük veri çağındaki daha iyi anlaşılmaktadır [52].

Bir veri ön işleme stratejisi olarak boyut indirgeme stratejilerinin, çeşitli veri madenciliği ve makine öğrenmesi problemleri için özellikle yüksek boyutlu verilerin hazırlanmasında etkili ve verimli olduğu kanıtlanmıştır. Boyut indirgeme metodolojisi

daha basit ve daha anlaşılır modeller oluşturmayı, veri madenciliği performansını iyileştirmeyi ve temiz, anlaşılır veriler hazırlamayı amaçlar.

Yüksek boyutlu verilere ilgili alanda algoritmalar uygulandığında, kritik bir konu olarak boyutsallığın laneti (curse of dimensionality) olarak bilinen fenomenler karşımıza çıkmaktadır. Bu kavram düşük boyutlu uzay için tasarlanmış algoritmaları olumsuz yönde etkileyen, yüksek boyutlu uzayda verilerin seyrekleşmesi olgusunu ifade eder [53]. Ayrıca, çok sayıda özellik ile öğrenme modelleri, görünmeyen verilerde performansın düşmesine neden olabilecek şekilde aşırı öğrenme eğilimindedir. Yüksek boyutlu veriler, veri analitiği için bellek depolama gereksinimlerini ve hesaplama maliyetlerini önemli ölçüde artırabilir. Bunlara ek olarak modellerin sadeleşmesi gerçek hayata uygulanabilirliğini ve açıklanabilirliğini artıracaktır. Model başarı performans ölçütlerimizi benzer değer aralıklarında tutarak daha hızlı sonuç üreten modellemeler gerçek hayatta birçok fayda sağlıyor olacaktır. Veri özellik uzayında boyut indirgeme bu sorunları çözmek için en güçlü araçlardan biridir.

Boyut indirgeme temel olarak özellik çıkarma (feature extraction) ve özellik seçimi (feature selection) olarak ikiye ayrılabilir. Özellik çıkarma, orijinal yüksek boyutlu özellik uzayını düşük boyutlu yeni bir özellik uzayına yansıtır. Yeni oluşturulan özellik uzayı, genellikle orijinal özelliklerin doğrusal veya doğrusal olmayan bir birleşimidir. Özellik seçimi ise, özellik uzayındaki ilgili özelliklerin bir alt kümesini seçmeyi amaçlayan boyut indirgeme yaklaşımıdır [54]. Hem özellik çıkarma hem de özellik seçimi, öğrenme performansını iyileştirme, hesaplama verimliliğini artırma, bellek depolamayı azaltma ve daha iyi genelleme yapabilen analitik modelleri oluşturma avantajlarına sahiptir. Bu nedenle, her ikisi de etkili boyutsallık azaltma teknikleri olarak kabul edilir. Bir yandan, ham girdi verilerinin belirli bir öğrenme algoritması tarafından anlaşılabilir herhangi bir özellik içermediği birçok uygulama için özellik çıkarımı tercih edilir. Öte yandan, özellik çıkarma bir dizi yeni özellik oluşturduğundan, bu özelliklerin fiziksel anlamlarını koruyamadığımız için daha fazla analiz sorunludur. Buna karşılık, orijinal özellikleri koruyarak yapılan özellik seçimi, orijinal özelliklerin fiziksel anlamlarını korur ve modellere daha iyi okunabilirlik ve yorumlanabilirlik sağlar. Bu nedenle, metin madenciliği ve genetik analiz gibi birçok uygulamada özellik seçimi sıklıkla tercih edilmektedir. Bazı durumlarda, özellik boyutluluğu o kadar yüksek olmasa da özellik çıkarma/seçiminin öğrenme

performansını iyileştirme, aşırı öğrenmeyi önleme ve hesaplama maliyetlerini düşürme gibi önemli bir rol oynadığına dikkat edilmelidir.

2.3.1 Özellik Seçimi

Forman'a göre [55] özellik seçiminin temel amacı, model performansını makul bir değerde tutarak ham veri kümesini temsil edebilecek en iyi alt kümeyi seçme işlemidir. Özellik seçimi, ham verileri ön işleme ve hedef değişkenleri tahmin etmede bize yardımcı olabilecek anlamlı tahmincilere dönüştürme sürecini ifade eder [56].

Veri kümelerinde kritik özelliklere kıyasla modellemeye katkısı neredeyse hiç olmayan özellikler olabilir. Bu özelliklerin varlığı tahminin doğruluğunu kısıtlayabilir. Ayrıca bu özellikler için gereksiz yere kaynak ayrılması gerekir. Sonuç olarak, tasarlanan modelinin eğitilmesi daha fazla zaman alır ya da aşırı öğrenme problemine neden olabilir. Bu nedenle bu tür senaryoların önüne geçebilmek amacıyla özellik seçimi kullanılmalıdır.

Özellik seçimi genel olarak denetimli, denetimsiz ve yarı denetimli yöntemler olarak sınıflandırılabilir. Denetimli özellik seçimi genellikle sınıflandırma veya regresyon problemleri için tasarlanmıştır. Farklı sınıflardan örnekleri ayırt edebilen veya regresyon hedeflerine yaklaşabilen en iyi özellik alt kümelerini bulmayı amaçlar.

Özellik seçim aşamaları öğrenme algoritmalarından bağımsız olabilir veya şimdiye kadar seçilen özelliklerin kalitesini değerlendirmek için bir sınıflandırıcının veya bir regresyon modelinin öğrenme performansından yinelemeli olarak yararlanabilir veya özellik seçimini temel alınan modele uyarlamak için bir öğrenme algoritmasının içsel yapısını kullanabilir.

Denetimsiz özellik seçimi genellikle kümeleme sorunları için tasarlanmıştır. Etiketlenmiş verilerin elde edilmesi süreçlerinin hem zaman hem de çaba açısından maliyetli olduğu problemlerde denetimsiz özellik seçimi yöntemleri ilgi görmüştür. Bu yöntemler özelliklerin önemini etiket bilgisi olmadan değerlendirmek için özellik alaka düzeyini tanımlamak amacıyla alternatif kriterler arar. Denetimli özellik seçiminden farklı, denetimsiz özellik seçimi, genellikle özellik seçimi aşamasında mevcut olan tüm örnekleri kullanır.

Denetimli özellik seçimi, yeterli etiket bilgisi mevcut olduğunda çalışır, denetimsiz özellik seçim algoritmaları ise herhangi bir sınıf etiketi gerektirmez. Bununla birlikte,

birçok gerçek dünya uygulamasında, genellikle sınırlı sayıda etiketlenmiş veriye sahibiz. Bu nedenle, hem etiketli hem de etiketsiz veri örneklerinden yararlanarak yarı denetimli yöntemler geliştirmek arzu edilir.

Özellik seçim stratejileri ile ilgili olarak, özellik seçim yöntemleri genel olarak filtre, sarmalayıcı ve gömülü yöntemler olarak kategorize edilebilir.

Filtre yöntemleri, herhangi bir öğrenme algoritmasından bağımsızdır. Özelliğin önemini değerlendirmek için veri özellikleri arasındaki farklı istatistik skorlarını temel alırlar. Filtre yöntemleri, hesaplama maliyeti olarak verimlidir [57]. Ancak özellik seçim aşamasını yönlendiren belirli bir öğrenme algoritmasının olmaması nedeniyle seçilen veri kümesi hedef öğrenme algoritmaları için optimal olmayabilir. Tipik bir filtre yöntemi iki adımdan oluşur. İlk adımda, bazı özellik değerlendirme kriterlerine göre özelliğin önemi skorlanır. İkinci adımda ise düşük skorlanmış özellikler filtrelenir. Filtre yöntemlerine örnek olarak Pearson korelasyon, spearman korelasyon, anova, ki-kare yöntemleri verilebilir.

Sarmalayıcı yöntemler, seçilen özellik alt kümelerini değerlendirmek için önceden tanımlanmış bir öğrenme algoritmasının tahmin performansına dayanır. Özellik seçimi ve öğrenme süreçlerini birleştirirler. Özellik kümesi arama bileşeni, önce özelliklerin bir alt kümesini oluşturur; daha sonra öğrenme algoritması, öğrenme performansına dayalı olarak bu özellik alt kümelerinin belirlenmiş performans kriterine göre değerlendirilmesi için işlev görür. En yüksek öğrenme performansı elde edilene veya istenen sayıda seçilen özellik elde edilene kadar tüm süreç yinelemeli olarak çalışır. Bu yaklaşımın ana dezavantajı yüksek karmaşıklık ve zayıf genelleme performansdır. Küçük veri kümelerinde aşırı öğrenme sorunu sarmalayıcı yöntemler için bir sorunken filtre yöntemleri büyük veri kümelerinde de kullanılabilirler [57]. Genetik algoritmalar, RFE, SFE örnek olarak verilebilir. Seçim aşaması alt kümelerin hangi sistematikle seçileceğine göre ileri doğru seçim (step forward selection), geriye doğru seçim (step backwards selection) ve adımsal seçim (stepwise selection) olarak kategorize edilebilir.

Gömülü yöntemler, özellik seçimi için filtre ve sarmalayıcı yöntemler arasında bir denge kurar. Seçim işlemini kullanılan algoritmanın kendisi yapar. En yaygın olarak kullanılan gömülü yöntemler, uydurma hatalarını en aza indirerek bir öğrenme modeline uymayı hedefleyen ve özellik katsayılarını aynı anda çok küçük veya sıfır

olmaya zorlayan modellerdir. L1 regülarizasyon, L2 regülarizasyon, Elastic Net yöntemleri örnek olarak verilebilir.

Literatürde bazı kaynaklar seçim stratejisi perspektifinden özellik seçim yöntemlerini hibrit yöntemleri de dahil ederek dört kategoride sınıflandırmışlardır [58]. Hibrit yöntemler, çoklu seçim algoritmalarının bir kombinasyonu olarak kabul edilebilir. Ana hedef, mevcut seçim algoritmasının kararsızlık ve bozulma sorunlarını çözmektir. Farklı yöntemlerden seçilen birden çok özellik alt kümesini bir araya getirerek daha güvenilir sonuçların üretilmesi hedeflenir.

Özellik seçiminin doğrusal ve doğrusal olmayan yaklaşımlar odağında üzerinde çalışılan problemin doğasına ve kullanılan algoritmalara bağlı olarak değerlendirilmesi daha güvenilir ve performanslı sonuçlar alınmasını sağlayacaktır.

Gömülü seçim yöntemlerinden olan L1 regülarizasyonu yaklaşımı LASSO yönteminin düşük gözlem sayısına ve yüksek örnek uzayına sahip kümeler üzerindeki uygulamaları oldukça popülerdir [59]. Maliyet fonksiyonuna katsayıların mutlak değerinin büyüklüğüne eşit bir ceza terimi ekleyen doğrusal regresyonun düzenlenmiş bir versiyonudur. Amaç belirli özelliklerin katsayılarını sıfıra indirerek maliyet fonksiyonunun enazlanmasıdır. Lasso'nun temel sınırlarından ilki özellik uzayındaki bileşenler arasındaki doğrusal olmayan ilişkileri yakalayamaması olarak söylenebilir. Bu sebeple Yamada ve arkadaşları tarafından önerilen [60] HSIC Lasso (Hilbert-Schmidt Independence Criterion Lasso) özellik seçim modeli doğrusal olmayan problemlerde özellik ilişkilerini değerlendirmek için kullanılmaktadır. HSIC Lasso bağımlı değişkene bağlı, seçilen diğer kestiricilere ise bağlı olmayan bir dizi değişkeni seçip çıkartır. Değişkenler arasındaki bağımlılıklar parametrik olmayan HSIC istatistikleri ile değerlendirilir. Özellik seçimini, öğrenme ve tahmin sürecini birleştiren Lasso'nun aksine HSIC Lasso yalnızca özellik seçimi için kullanılır. Tahmin yönteminden ziyade çıktı değişkenini bir Gram matrisine dönüştürdüğünden ve direkt olarak çıktıyı tahmin edemediğinden özellik eleme yöntemi olarak düşünülebilir. HSIC Lasso'nun özellik seçim yaklaşımı mRMR yöntemine benzerdir. Yöntem iki özellik ya da özellik ve çıktı değişken arasındaki doğrusal olmayan bağımlılıkları değerlendirmek için klasik Lasso'nun aksine kernel tabanlı ayrıklık ölçüsü olan HSIC istatistiklerini kullanır. HSIC istatistik değerleri negatif olmayan değerler alır ve yalnızca iki değişken istatistiksel olarak bağımsız ise sıfırlanır [61].

Doğrusal olmayan özellik seçimi problemlerinde kullanılan bir diğer yöntem ise 2019 yılında Uber makine öğrenmesi platform mühendisleri tarafından sunulan [62] çalışmasından sonra popülerliği artan 2003 yılında [63] ile ortaya konan mRMR (minimum redundancy maximum relevancy) yöntemidir. Uber çalışmasında mrmr ile ortaya konan çıktıda çok çeşitli pazarlama faaliyetlerinde otomatik özellik seçim yöntemi ile kullanıcı edinme, çapraz satış ve kullanıcıyı platforma yeniden katma gibi çeşitli problem çözümlerinde efektif şekilde kullanıldığı ortaya konmuştur.

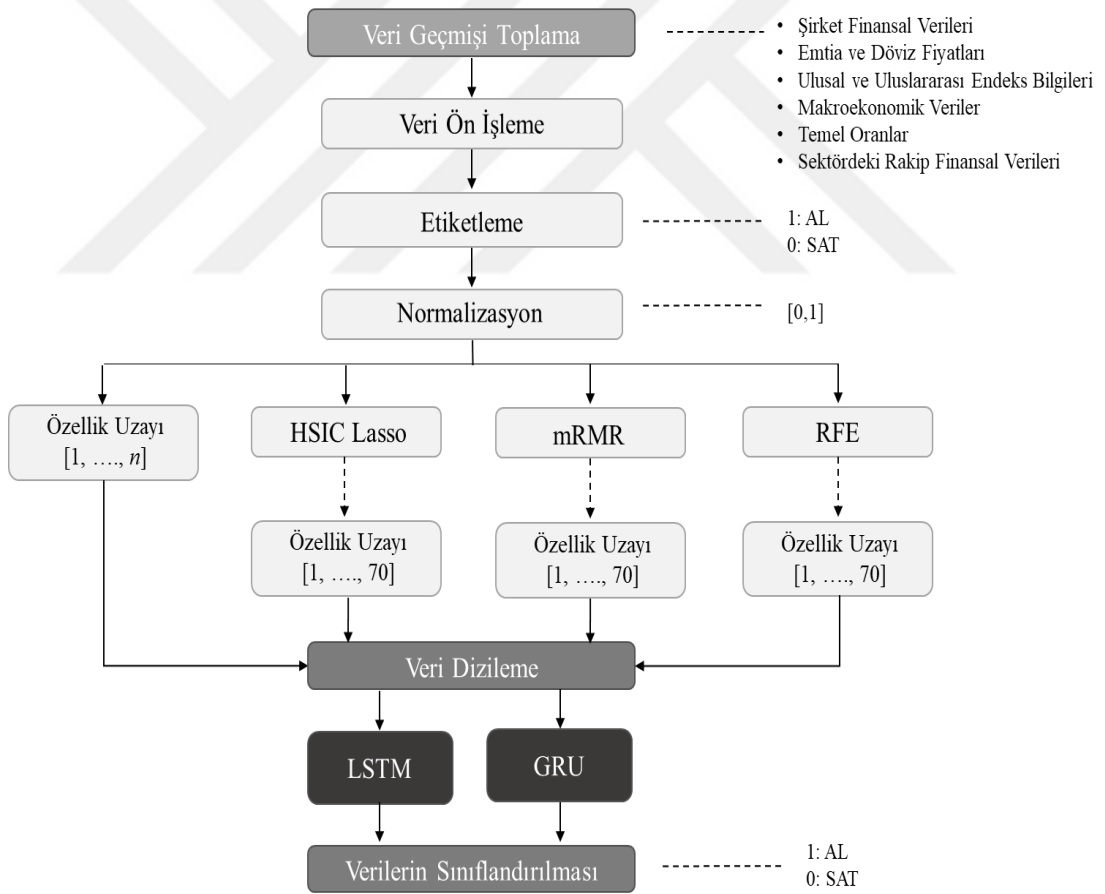
Özellikleri mrmr kriterlerine göre yinelemeli olarak çalışarak sıralayan yaklaşım, alaka düzey seviyelerini F istatistiklerinden ve ortak bilgi (mutual information) değerlerini hesaplayarak sıralar. Her yinelemede en iyi özelliği tanımlar ve onu seçilen özellik listesine ekler. Özellik seçilmiş listeye eklendiğinde bir daha çıkmaz. İstenen sayıda özellik listeye eklenene kadar iterasyon adımları devam eder. mRMR yöntemi sınıf çıktısıyla yüksek korelasyona, kendi aralarında ise düşük korelasyona sahip özellikleri seçme eğiliminde olan bir özellik seçim yaklaşımıdır. Performanslı varyantlar için özellikler için sınıfla korelasyonu hesaplamak için F istatistikleri, kendi aralarındaki korelasyonu hesaplamak içinse Pearson korelasyon katsayısı kullanılabilir. Sonrasında uygunluk ve fazlalık fonksiyonu olan amaç fonksiyonunu enbüyüklemek için arama yöntemleri kullanılarak özellikler tek tek seçilir. Amaç fonksiyonunun yaygın olarak kullanılan iki türü, uygunluk ve fazlalık farkını veya bölümünü temsil eden MID ve MIQ kriterleridir [64].

Çalışmada uygulanan bir diğer yaklaşım ise sarmal yöntemlerden biri olan RFE yaklaşımıdır. Uygulamada ilk olarak veri kümesindeki mevcut özellik uzayı ile seçilen model eğitilir ve ardından özellikler seçilen ölçüt metriklerine göre önem skorlarına göre sıralanır. Ardından, en az önemli özellik kümeden çıkartılarak model yeni özellik uzayı alt kümeleri ile yeniden eğitilir ve daha önce seçilen metriklere dayalı olarak model performansı yeniden ölçülerek ilk adımdaki eleme işlemleri tekrarlanır. Özellik eleme süreci en iyi performans çıktısı sağlayan istenen “k” adet özellik kalana kadar devam eder. Dolayısıyla özellik seçimi sonunda elde edilmesi istenen özellik sayısı yöntemde kullanılan bir ayar parametresidir. Buna ek olarak özellik alt kümelerinin skorlarının belirlendiği kullanılan algoritmanın seçimi de belirlenmesi gereken diğer bir etkidir. Doğrusal olmayan ilişkileri yakalayabilmek adına RFE yöntemi RF algoritmasıyla beraber kullanılmıştır.



3. UYGULAMA TASARIMI

Bu bölümde uygulamanın tasarlandığı deneysel ortam, kullanılan hisse senedi veri kümeleri, etiketleme ve boyut indirgeme stratejileri, zaman serisi verilerinin kayan pencereler yöntemiyle farklı geri bakma değerleri ile dizilere ayrılması, kullanılan derin öğrenme modellerinin parametreleri ve model çıktılarının yorumlanmasında kullanılacak performans metrikleri detaylandırılmıştır. Çalışmada izlenen tahmin akışı detaylı olarak Şekil 3.1’de gözlemlenebilir.



Şekil 3.1 : Hisse senedi yön değişimi tahminleme diagramı.

3.1 Deneyisel Ortam Bilgileri

Programlama Spyder 5.1 versiyonlu IDE üzerinde Python 3.9 kullanılarak yapıldı. Veri işleme süreçlerinde numpy ve pandas paketleri kullanıldı. Sinir ağı modellemesi ve eğitimi TensorFlow üzerinde Keras kütüphanesi ile geliştirilmiştir. Boyut indirgeme ve performans metriklerinin hesaplanmasında kullanılan diğer paketlerin versiyon bilgileri Çizelge 3.1’de verilmiştir.

Çizelge 3.1 : Python paket versiyonları.

Paket	Versiyon
TensorFlow	2.8.0
Pandas	1.3.4
Numpy	1.20.3
Scikit-learn	0.24.2
Mrmr-selection	0.2.5
pyHSICLasso	1.4.2

3.2 Veri Kümesi

Borsa İstanbul’da işlem gören Akbank, Turkcell ve Ford Otosan hisse senetlerinin tez çalışması kapsamında oluşturulmuş veri kümeleri 02/01/2017 - 21/01/2022 aralığındaki 1270 sıralı işlem gününü kapsamaktadır.

Çalışmada kullanılan veri kümelerine ait tamamı sayısal özellik uzayı varlık fiyatları, teknik göstergeler, emtia fiyatları, döviz kurları, yerel ve uluslararası endeks bilgileri, aracı kurum tahminleri, faaliyet alanı eş şirket varlık verileri ve şirkete ait finansal bilgileri gibi değişkenleri içeren geniş bir perspektifte Reuters’ten alınarak oluşturulmuştur. Modellemede kullanılan hisse senetlerine ait veri kümesi bilgileri Çizelge 3.2’de gösterilmiştir.

Çizelge 3.2 :AKBNK, TCELL, FROTO veri kümesi detayları.

Hisse Senedi Kodu	Veri Kümesi Özellik Sayısı	Gözlem Tarih Aralığı
AKBNK	176	02/01/2017 – 21/01/2022
TCELL	185	02/01/2017 – 21/01/2022
FROTO	171	02/01/2017 – 21/01/2022

3.3 Veri Etiketleme ve Ölçekleme

Reuters kaynak alınarak oluşturulan, Borsa İstanbul'daki 1270 işlem gününü kapsayan AKBNK, TCELL ve FROTO hisse senedi veri kümelerinde aykırı ve kayıp veri bulunmamaktadır. Bu nedenle veriler üzerinde bu yönde bir ön işleme yapılmamıştır. İlk aşamada zaman serisi verilerinin denetimli bir sınıflandırma problemine dönüştürülmesi kapsamında her bir zaman adımı için etiketleme yapılmıştır. Üzerinde çalışılan problem zaman serisi olarak modellenmiş verileri kullanarak hisse senetlerinin bir sonraki işlem günündeki yön hareketlerini tahminlemeyi amaçlayan bir sınıflandırma problemidir. Bu doğrultuda her işlem günü için oluşturulmuş verilere bir sonraki işlem günündeki hisse senedi fiyat değişim yüzdesi eklenmiştir.

AKBNK, TCELL ve FROTO veri kümelerinde t günündeki veri t+1 gününde, t+1 günündeki veri de t+2 gününde hisse senedi fiyatının hangi oranda değiştiğine dair bilgi taşımaktadır. Bu noktada 3 farklı durum için zaman serisi verileri yön tahminine bağlı olarak al ve sat kararı verebilmek amacıyla etiketlenmiştir. t işlem günü verisi, t+1 günündeki hisse senedi yönü negatifse 0, pozitifse 1 sınıfına aittir. Eğer t+1 gününde yönde herhangi bir değişme olmamışsa ilgili etiket için t+2 gününe bakılmıştır. t+2 işlem günündeki değişim negatifse t verisi 0, pozitifse 1 sınıfına aittir.

Belirtilen veri etiketleme işleminden sonra özellik uzayı ölçeklendirilmiştir. Ölçekleme işlemi farklı değer aralıklarına veya birimlere sahip özelliklerin karşılaştırılabilir değerlere dönüştürülmesine yardımcı olur. Literatürde sıklıkla standardizasyon ve normalizasyon olmak üzere iki ölçekleme yöntemi kullanılır [65]. Sadece sayısal verilerden oluşan hisse senetlerine ait özellik uzayı scikit-learn kütüphanesi kullanılarak Min-Max ölçeklendirme yöntemiyle 0 - 1 arasına ölçeklendirilmiştir. Veri normalizasyonunda kullanılan matematiksel yaklaşım denklem 3.1'de gösterilmiştir.

$$\tilde{X} = \frac{x - \text{Min}}{\text{Max} - \text{Min}} \quad (3.1)$$

Max, zaman dizisi boyunca x'in en büyük değeri, Min ise en küçük değeridir.

3.4 Özellik Seçim Stratejisi

Çalışmanın boyut indirgeme kapsamında ele alınması gereken amaçlarından biri doğrusal olmayan hisse senedi verileri üzerinde gerçekleştirilen özellik seçim yaklaşımlarının farklı geri bakma değerleriyle oluşturulmuş zaman dizileri ile eğitilen derin ağlardaki tahmin performansına olan etkisinin incelenmesidir.

Bu amaç doğrultusunda da HSIC Lasso, mRMR ve RFE özellik seçim yöntemleri hisse senedi verileri üzerinde uygulanmış, LSTM ve GRU mimarilerinde gösterdikleri performanslar incelenmiştir. HSIC Lasso seçimi için Python'da pyHSICLasso, mRMR yöntemi için ise mrmr_selection paketinden sınıflandırma problemi kapsamında yararlanılmıştır. Scikit-learn paketleri kullanılarak gerçekleştirilmiş RFE özellik seçim yönteminde ise karar verici algoritma olarak random forest kullanılmıştır.

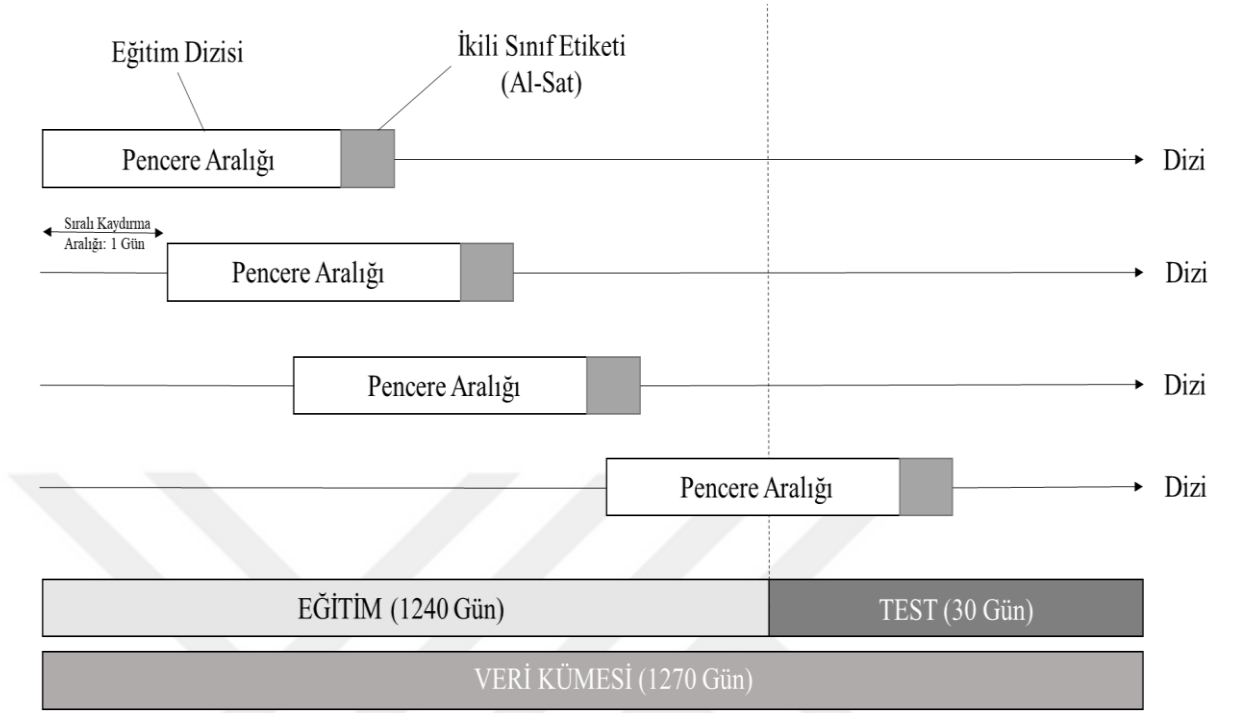
Modeller üzerinde farklı özellik sayıları için yapılan stres testlerinde optimum değer olarak 70 bulunmuş ve üç hisse senedi verisi için de çalışmada kullanılan tüm özellik seçim yöntemleriyle seçilecek özellik sayısı 70 olarak belirlenmiştir. Özellik seçimi yapılmayan senaryolarda modele verilecek girdi verisi Çizelge 3.2'de belirtilmiş tüm özellik uzayı olacaktır.

3.5 Zaman Serisi Verileri Dizileme Yaklaşımları

Ele alınan hisse senedi yön tahmini probleminde zamansal hisse senedi verileri üzerinde kayan pencereler yöntemi ile ham verilerden iki boyutlu diziler oluşturulmuştur. İzlenen ana yöntem Şekil 3.2'de gösterilmiştir. 1270 işlem gününü kapsayan hisse senedi verileri üzerinde belirli pencere aralıkları 1'er gün kaydırılarak derin öğrenme modellerinin girdi verileri oluşturulmuştur.

Hisse senedinin bir sonraki gün yapacağı hareketin geriye dönük kaç gün kullanılarak eğitilmiş model ile en başarılı şekilde tahminlenebileceğinin incelenmesi de çalışmanın amaçları arasında yer almaktadır. Model tahmin aralığı tüm deneylerde 30 gün olarak belirlenmiştir. 1240 günlük veri kullanılarak oluşturulan zaman dizileri

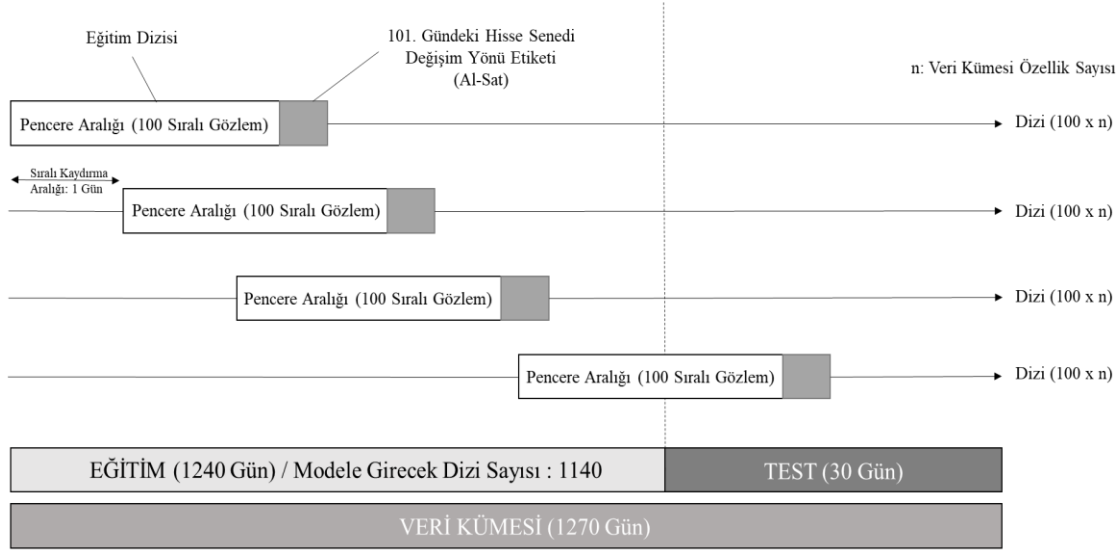
kullanılarak eğitilen modeller ile sonraki 30 günün hisse senedi yönlerinin tahminlenmesi amaçlanmıştır.



Şekil 3.2 : Zaman serisi üzerinde kayan pencereler yöntemi ile dizi oluşturma ve etiketleme.

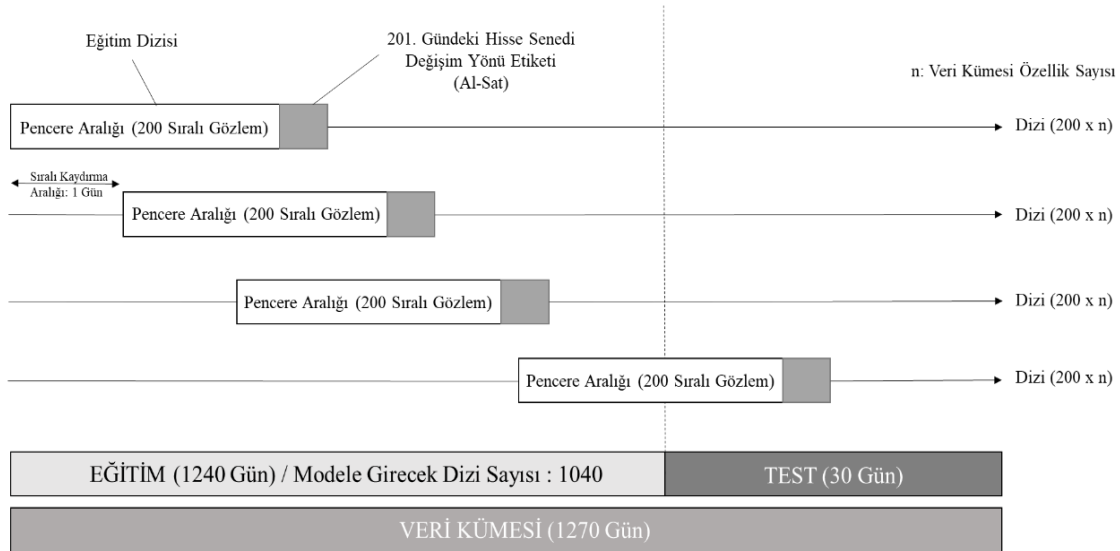
Çalışmada LSTM ve GRU modellerine girdi olarak sağlanan veri 4 farklı dizileme tekniğiyle oluşturulmuştur. Bu yaklaşımlardan ilki geri bakma değerinin 100 olarak belirlendiği Şekil 3.3’de gösterilen yöntemdir. Eğitim ve test olarak ayrılmış tüm veri kümesi üzerinde sıralı olarak 1’er gün için pencereler kaydırılarak tüm veri zamansal dizilere ayrılacaktır.

Eğitim aşamasında modele girecek olan 1140 zaman dizisinin her biri boyutsal indirgeme yapılmış senaryolarda 100x70’lik 2 boyutlu bir dizidir. Boyut indirgeme yapılmamış senaryolarda ise AKBNK için 100x176, TCELL için 100x185, FROTO için ise 100x171’dir. İlk dizi baz alındığında ilk 100 işlem günü ile oluşturulmuş eğitim dizisi 101. Gündeki hisse senedi değişim etiketine sahip olacaktır.



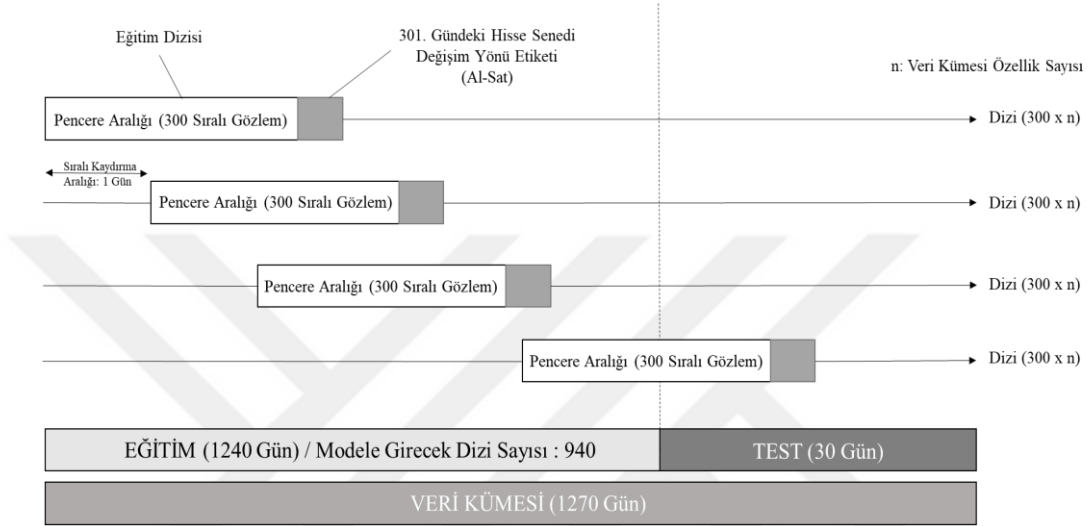
Şekil 3.3 : Geri bakma değeri 100 için veri kümesinin dizilere ayrılması.

Geri bakma değerinin 200 olarak belirlendiği diğer yöntem Şekil 3.4'te gösterilmiştir. Eğitim aşamasında modele girecek olan 1040 zaman dizisinin her biri boyutsal indirgeme yapılmış senaryolarda 200x70'lik 2 boyutlu bir dizidir. Boyut indirgeme yapılmamış senaryolarda ise AKBNK için 200x176, TCELL için 200x185, FROTO için ise 200x171'dir. İlk dizi baz alındığında ilk 200 işlem günü ile oluşturulmuş eğitim dizisi 201. Gündeki hisse senedi değişim etiketine sahip olacaktır.



Şekil 3.4 : Geri bakma değeri 200 için veri kümesinin dizilere ayrılması.

Geri bakma değerinin 300 olarak belirlendiği diğer yöntem ise Şekil 3.5'te gösterilmiştir. Eğitim aşamasında modele girecek olan 940 zaman dizisinin her biri boyutsal indirgeme yapılmış senaryolarda 300x70'lik 2 boyutlu bir dizidir. Boyut indirgeme yapılmamış senaryolarda ise AKBNK için 300x176, TCELL için 300x185, FROTO için ise 300x171'dir. İlk dizi baz alındığında ilk 300 işlem günü ile oluşturulmuş eğitim dizisi 301. Gündeki hisse senedi değişim etiketine sahip olacaktır.

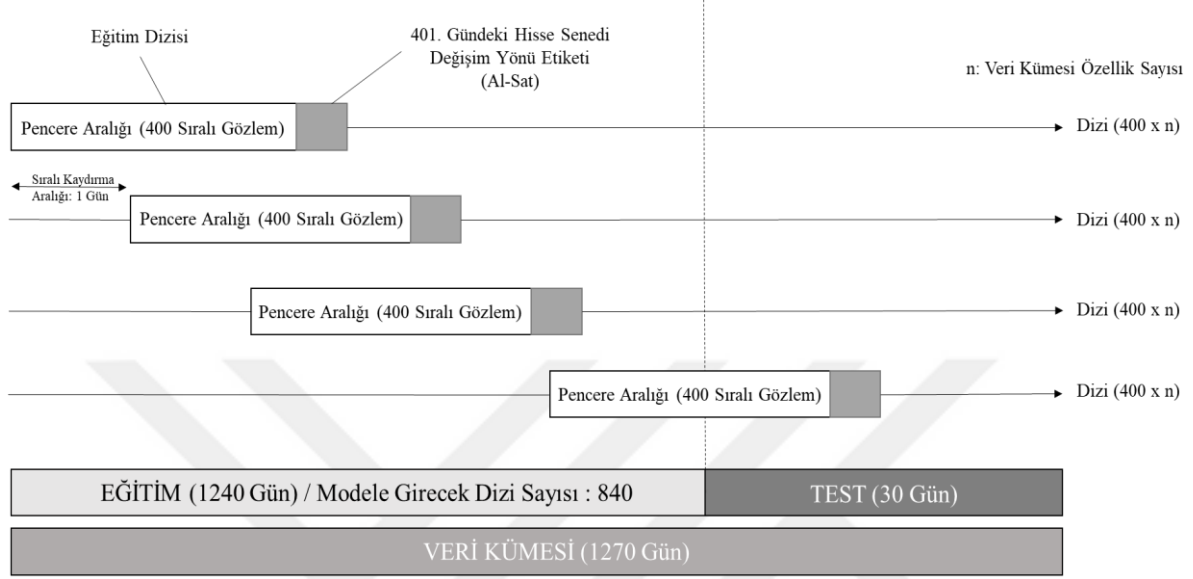


Şekil 3.5 : Geri bakma değeri 300 için veri kümesinin dizilere ayrılması

Geri bakma değerinin 400 olarak belirlendiği son yöntem ise Şekil 3.6'te gösterilmiştir. Eğitim aşamasında modele girecek olan 840 zaman dizisinin her biri boyutsal indirgeme yapılmış senaryolarda 400x70'lik 2 boyutlu bir dizidir. Boyut indirgeme yapılmamış senaryolarda ise AKBNK için 400x176, TCELL için 400x185, FROTO için ise 400x171'dir. İlk dizi baz alındığında ilk 400 işlem günü ile oluşturulmuş eğitim dizisi 401. Gündeki hisse senedi değişim etiketine sahip olacaktır.

Yöntemler sonunda 30 günlük test verisi için tahmin modellerine 30 dizi girecektir ve model her bir dizi için bir sınıf etiketi tahmini yapacaktır. Bu sınıf etiketleri önceki bölümlerde detaylandırılan etiketleme stratejisine göre belirlenmiş 1 ve 0 sınıflarıdır. 1 sınıfına ait bir işlem yatırımcı için AL sinyali verirken 0 sınıfına ait bir işlem yatırımcıya SAT sinyalini göndermektedir.

Özetle sonraki gün hisse senedi yön değişiminin tahmin süreci her bir zaman adımında geçmiş 100, 200, 300 ve 400 işlem günü kullanılarak eğitilmiş derin sinir ağı modelleri ile yapılmaktadır.



Şekil 3.6 : Geri bakma değeri 400 için veri kümesinin dizilere ayrılması

3.6 Tasarlanan Model Parametreleri

LSTM ve GRU modellerinde model tasarımcısının seçimine bırakılan, probleme ve veriye göre performansı değişiklik gösteren parametreler bulunmaktadır. En verimli model mimari parametre gruplarının bulunması ve veri üzerinde uygulanması da başlı başına ayrı bir araştırma alanıdır.

Tez çalışmasının amacı farklı geri bakma değerlerinin ve doğrusal olmayan özellik seçim yöntemlerinin bileşiminin eğitim senaryolarında 3 farklı sektörden hisse senedi üzerindeki tahmin performansına etkilerinin incelenmesidir. Bu doğrultuda tez çalışmasında kullanılan modeller için literatürdeki çalışmalar baz alınarak ve farklı parametrelerle yapılan validasyon testleri değerlendirilerek Çizelge 3.3'teki

parametreler belirlenmiştir. 4. Bölümde sonuçları sunulan modellerin çalışma süresince parametreler belirtilen değerlerde sabit tutulmuştur.

Çizelge 3.3 : LSTM ve GRU model parametreleri.

	LSTM	GRU
Kayıp Fonksiyonu	Çapraz Entropi	Çapraz Entropi
Optimizasyon Algoritması	Adam	Adam
Batch Boyutu	64	64
Aktivasyon Fonksiyonları	Tanh, sigmoid	Tanh, sigmoid
Öğrenme Hızı	0.001	0.001
Gizli Katman Sayısı	3	3
Gizli Katman Nöron Sayıları	40-20-20	20-10-10

3.7 Performans Metrikleri

Üzerinde çalışılan hisse senedi yön tahmini problemi denetimli bir sınıflandırma problemidir. Kayan pencereler yöntemiyle etiketlenmiş veriler üzerinden oluşturulmuş alt diziler model eğitim girdisi olarak kullanılmış ve aynı yöntemle alt dizilere ayrılmış test verileri üzerinde model performansları ölçülmüştür. Çalışmada yararlanılan ana metrikler aşağıdaki gibi açıklanabilir.

- Doğruluk, test veri kümesi üzerinde doğru olarak tahmin edilmiş örneklerin yüzdesidir.
- Duyarlılık, sınıfa ait örneklerin ne kadarının doğru olarak tahmin edildiğinin göstergesidir.
- Kesinlik, sınıf tahminlerinin hangi oranda gerçekten ilgili sınıfa ait olduğunun göstergesidir.
- F1 Skor, kesinlik ve duyarlılığın harmonik ortalamasıdır.

Model performansları için long doğruluk, short doğruluk, ağırlıklı ortalama F1 skoru ve toplam doğruluk göstergelerinden yararlanılmıştır. Long işlemler bir yatırım varlığının fiyatının artacağı, short işlemler ise azalacağı beklentisiyle yatırımcı taraf

tarafından açılmış işlemlerdir. Çalışmada 1 sınıfına ait “AL” tahminleri long pozisyonu, 0 sınıfına ait “SAT” tahminleri ise short pozisyonu ifade etmektedir. Bu doğrultuda genel karışıklık matrisi düşünüldüğünde long doğruluk 1 sınıfı için duyarlılık, short doğruluk da 0 sınıfı için duyarlılık verisini ifade etmektedir. İncelenen diğer göstergeler ise “AL” ve “SAT” sınıfları için elde edilen F1 değerlerinin ağırlıklı ortalaması ve tüm tahminlerdeki toplam doğruluk oranıdır.

3.8 Sonuçların Değerlendirilmesi

Çalışma kapsamında yatırım stratejilerinde kullanılması planlanan modellerde long ve short doğruluk değerlerinin ikisinin birden dengeli olarak %50 üzerinde olması beklenmektedir. Buna ek olarak toplam doğruluk ve ağırlıklı ortalama F1 skorun en az %55 olması modelin kabul edilebilir başarı kriteridir.

4. MODEL PERFORMANS ÇIKTILARI

Bu bölümdeki çıktılar tasarlanan derin öğrenme modellerinin Borsa İstanbul'da işlem gören farklı sektörlerden 3 şirketin hisse senedi yön tahminlerinin deneysel sonuçlarını içermektedir. Finansal tahmin problemlerindeki popüleritesi son zamanlarda artan ve literatürdeki benzer çalışmalarda iyi çıktılar üreten LSTM ve GRU mimarilerinin öncelikle gerçek finansal veriler üzerinde uygulanabilirliğinin ortaya konması, ikinci olarak ise model tahmin performansının farklı özellik seçim yöntemleri ve boyut olarak değişen zaman pencerelerine sahip zamansal veri dizileri ile iyileştirilebilir olup olmadığının gösterilmesi amaçlanmaktadır.

Tasarlanan modeller farklı sektörden 3 farklı hisse senedi üzerinde gerçekleştirilmiştir. Böylece ağ tasarımlarının genel olarak farklı şirketlerin veri kalıplarına uygulanabilirliğinin de incelenmesi çalışmanın diğer araştırma konuları arasındadır. Ele alınan hisse senedi işlem kodları sırasıyla Akbank için AKBNK, Turkcell için TCELL ve Ford Otosan için FROTO'dur.

Sonuçlar farklı geri bakma değerleri için belirtilen performans metrikleri ve başarı kriterleriyle ayrı bölümlerde incelenecek sonrasında sınıflandırma performansının geri bakma değerleri ve özellik seçim yöntemleri ile ilişkisi ortaya konacak son olarak ise en iyi tahmin performanslarını sağlayan modeller arasındaki ortak yöntemlerin değerlendirilmesi yapılacaktır.

4.1 Geri Bakma Değeri 100 Olan Dizilemede Tahmin Performansı

Çizelge 4.1'de AKBNK için tahmin performansı metrikleri verilmiştir. Short doğruluk, F1 skor ve toplam doğruluk göz önüne alındığında özellik seçimi ile veri üzerinde boyut indirgeme yapılmamış GRU modeli en iyi performansı göstermiştir. HSIC-Lasso-LSTM ve HSIC-Lasso-GRU modelleri toplam doğruluk ve ağırlıklı F1 skor performansı yönünden benzer çıktılar üretmiştir. Ancak long ve short tahminlerindeki dengesizlik belirlenen başarı kriterlerinin dışında kalmıştır. mRMR-

LSTM ile LSTM'e göre short doğrulukta iyileşme sağlandığı görülmektedir. RFE kullanılan modellerde LSTM ve GRU modellerine göre bir iyileşme sağlanmamıştır.

Çizelge 4.1 : AKBNK 100 adımlı dizileme performans sonuçları.

MODEL	Long Doğruluk	Short Doğruluk	Ağırlıklı F1 Skor	Toplam Doğruluk
LSTM	0.5000	0.5625	0.5333	0.5333
GRU	0.4286	0.8750	0.6475	0.6667
HSIC-LASSO-LSTM	0.3571	0.8125	0.5770	0.6000
HSIC-LASSO-GRU	0.8571	0.3750	0.5778	0.6000
MRMR-LSTM	0.5000	0.6250	0.5652	0.5667
MRMR-GRU	0.5000	0.5625	0.5333	0.5333
RFE-LSTM	0.4286	0.5625	0.4983	0.5000
RFE-GRU	0.3571	0.7500	0.5488	0.5667

Çizelge 4.2’de TCELL için tahmin performans metrikleri verilmiştir. LSTM ve GRU modelleri toplam doğruluk baz alındığında kabul edilebilir aralıkta çıktı üretse de long-short tahminlerinde dengesiz performans göstermiştir. Boyut indirgeme yöntemleri ile tasarlanmış tüm modellere bakıldığında yalnızca RFE ile GRU üzerinde kabul edilebilir performans iyileşmesi sağlanmıştır.

Çizelge 4.2 : TCELL 100 adımlı dizileme performans sonuçları.

MODEL	Long Doğruluk	Short Doğruluk	Ağırlıklı F1 Skor	Toplam Doğruluk
LSTM	0.2143	0.8750	0.5116	0.5667
GRU	0.2143	0.9375	0.5365	0.6000
HSIC-LASSO-LSTM	0.6429	0.4375	0.5292	0.5333
HSIC-LASSO-GRU	0.4286	0.5625	0.4983	0.5000
MRMR-LSTM	0.3571	0.6875	0.5204	0.5333
MRMR-GRU	0.7857	0.1250	0.3648	0.4333
RFE-LSTM	0.2857	0.6875	0.4794	0.5000
RFE-GRU	0.4286	0.8750	0.6475	0.6667

Çizelge 4.3’te ise FROTO için performans çıktıları verilmiştir. mRMR-GRU tüm metrikler yönünden kabul edilebilir en başarılı tahmin performansını göstermiştir. LSTM modeli de belirlenen başarı kriterlerini sağlayan diğer yöntem olmuştur.

Çizelge 4.3 : FROTO 100 adımlı dizileme performans sonuçları.

MODEL	Long Doğruluk	Short Doğruluk	Ağırlıklı F1 Skor	Toplam Doğruluk
LSTM	0.5385	0.5882	0.5681	0.5667
GRU	0.3846	0.7059	0.5560	0.5667
HSIC-LASSO-LSTM	0.6923	0.4118	0.5271	0.5333
HSIC-LASSO-GRU	0.7692	0.4706	0.5946	0.6000
MRMR-LSTM	0.6923	0.4118	0.5271	0.5333
MRMR-GRU	0.6154	0.5882	0.6018	0.6000
RFE-LSTM	0.4615	0.6471	0.5642	0.5667
RFE-GRU	0.3077	0.5882	0.4593	0.4667

4.2 Geri Bakma Değeri 200 Olan Dizilemede Tahmin Performansı

Çizelge 4.4'teki AKBNK tahminleri LSTM ve HSIC-Lasso-GRU modellerinin dengeli short-long doğruluk performansıyla toplamda sırasıyla %70 ve %66 doğruluk ile çıktı ürettiğini göstermektedir. Genel olarak bakıldığında tüm GRU modelleri belirlenen başarı kriterlerinden kötü bir performans sergilememiş, yatırım kararlarında kullanılabilecek çıktılar üretmiştir.

Veri kümesinde özellik seçim yöntemleri uygulanmış LSTM modelleri ise tüm verinin girdi olarak kullanıldığı baz LSTM modelinin performansının gerisinde kalmıştır. HSIC-Lasso-LSTM yöntemi LSTM model çıktılarıyla karşılaştırıldığında long doğrulukta aynı performans alınsa da short doğrulukta ciddi bir düşüş olmuştur. MRMR-LSTM modelinde de short doğrulukta performans kaybı yaşanmasa da long doğrulukta düşüş gözlenmiştir. RFE-LSTM yönteminde ise baz modele göre tüm metriklerde performans kaybı gözlenmektedir.

GRU üzerinde uygulanmış özellik seçim yöntemleri özelinde bakıldığında HSIC-Lasso yönteminin baz model performansında iyileşme sağladığı görülmektedir. Long doğrulukta ciddi bir performans iyileşmesi sağlanmış ve toplam doğruluğa bakıldığında GRU modelleri arasında en başarılı çıktıyı sağlamıştır.

Çizelge 4.4 : AKBNK 200 adımlı dizileme performans sonuçları.

MODEL	Long Doğruluk	Short Doğruluk	Ağırlıklı F1 Skor	Toplam Doğruluk
LSTM	0.5714	0.8125	0.6949	0.7000
GRU	0.5000	0.6875	0.5964	0.6000
HSIC-LASSO-LSTM	0.5714	0.5625	0.5671	0.5667
HSIC-LASSO-GRU	0.7143	0.6250	0.6667	0.6667
MRMR-LSTM	0.4286	0.8125	0.6183	0.6333
MRMR-GRU	0.6429	0.5000	0.5652	0.5667
RFE-LSTM	0.4286	0.6875	0.5592	0.5667
RFE-GRU	0.2857	0.7500	0.5065	0.5333

Çizelge 4.5'teki TCELL metriklerine bakıldığında geri bakma değeri 100 olan dizilemede olduğu gibi RFE yöntemi ile GRU modeli üzerinde ciddi iyileştirme sağlanmıştır; ancak long doğrulukta performans yatırım kabul kriterlerini sağlayamamıştır. Toplam doğruluğa bakıldığında HSIC-Lasso-LSTM modeli genel LSTM performansını iyileştirmiş ve yatırım kararlarında kabul edilebilir çıktılar sağlamıştır. mRMR özellik seçim yöntemi kullanılarak tasarlanan modellere bakıldığında hem LSTM hem GRU baz modellerine göre long doğruluklar artırılrsa da short doğrulukta kötüleşme olmasından dolayı mRMR ile yatırım kabul kriterlerini sağlayan bir performans alınamamıştır.

Çizelge 4.5 : TCELL 200 adımlı dizileme performans sonuçları.

MODEL	Long Doğruluk	Short Doğruluk	Ağırlıklı F1 Skor	Toplam Doğruluk
LSTM	0.4286	0.6250	0.5291	0.5333
GRU	0.1429	0.6875	0.3897	0.4333
HSIC-LASSO-LSTM	0.5714	0.5625	0.5671	0.5667
HSIC-LASSO-GRU	0.3571	0.5625	0.4618	0.4667
MRMR-LSTM	0.5000	0.4375	0.4667	0.4667
MRMR-GRU	0.5000	0.3125	0.3946	0.4000
RFE-LSTM	0.3571	0.5625	0.4618	0.4667
RFE-GRU	0.4286	0.8750	0.6475	0.6667

Çizelge 4.6’da da FROTO için tahmin çıktıları gösterilmektedir. GRU ile sağlanan başarılı sınıflandırma performansı RFE-GRU ile long doğruluğun da iyileşmesiyle daha da artmıştır. LSTM’e göre short doğruluğu iyileştiren HSIC-Lasso-LSTM başarı kriterlerini sağlayan diğer model olmuştur.

Çizelge 4.6 : FROTO 200 adımlı dizileme performans sonuçları.

MODEL	Long Doğruluk	Short Doğruluk	Ağırlıklı F1 Skor	Toplam Doğruluk
LSTM	0.5385	0.4118	0.4667	0.4667
GRU	0.5385	0.7059	0.6312	0.6333
HSIC-LASSO-LSTM	0.5385	0.6471	0.6000	0.6000
HSIC-LASSO-GRU	0.3077	0.7059	0.5155	0.5333
MRMR-LSTM	0.6154	0.4118	0.4972	0.5000
MRMR-GRU	0.7692	0.2353	0.4296	0.4667
RFE-LSTM	0.3846	0.7059	0.5560	0.5667
RFE-GRU	0.6154	0.7059	0.6667	0.6667

4.3 Geri Bakma Değeri 300 Olan Dizilemede Tahmin Performansı

AKBNK tahmin çıktıları Çizelge 4.7’de gösterilmiştir. mRMR ve RFE yöntemleri ile sağlanan tahminler baz LSTM ve GRU’ya göre genel performansta kötüleşmeye neden olmuştur. HSIC-Lasso yöntemi ile LSTM’e göre F1 skor ve toplam doğruluk iyileştirilse de long ve short doğruluk tarafındaki dengesizlik artmıştır. LSTM modeli genel başarı kriterlerine yakın bir performans sergilemiştir.

Çizelge 4.7 : AKBNK 300 adımlı dizileme performans sonuçları.

MODEL	Long Doğruluk	Short Doğruluk	Ağırlıklı F1 Skor	Toplam Doğruluk
LSTM	0.5000	0.6250	0.5652	0.5667
GRU	0.3571	0.8750	0.6051	0.6333
HSIC-LASSO-LSTM	0.4286	0.7500	0.5889	0.6000
HSIC-LASSO-GRU	0.5000	0.3750	0.4314	0.4333
MRMR-LSTM	0.4286	0.5000	0.4667	0.4667
MRMR-GRU	0.5000	0.4375	0.4667	0.4667
RFE-LSTM	0.3571	0.6250	0.4914	0.5000
RFE-GRU	0.3571	0.8125	0.5770	0.6000

Çizelge 4.8’deki TCELL çıktıları en iyi tahmin performansının LSTM ile sağlandığını göstermektedir. Buna ek olarak HSIC-Lasso-LSTM ve mRMR-LSTM modelleri de toplam doğrulukta kabul edilebilir bir azalma ile başarılı performans gösteren diğer modellerdir. GRU modelleri ise genel olarak kabul edilebilir yatırım karar kriterlerinin dışında kalarak kötü bir performans sergilemiştir.

Çizelge 4.8 : TCELL 300 adımlı dizileme performans sonuçları.

MODEL	Long Doğruluk	Short Doğruluk	Ağırlıklı F1 Skor	Toplam Doğruluk
LSTM	0.5714	0.6250	0.6000	0.6000
GRU	0.1429	0.6875	0.3897	0.4333
HSIC-LASSO-LSTM	0.5714	0.5625	0.5671	0.5667
HSIC-LASSO-GRU	0.0714	1.0000	0.4415	0.5667
MRMR-LSTM	0.5714	0.5625	0.5671	0.5667
MRMR-GRU	0.2143	0.5625	0.3833	0.4000
RFE-LSTM	0.3571	0.6250	0.4914	0.5000
RFE-GRU	0.3571	0.8125	0.5770	0.6000

Geri bakma değeri 300 olan dizileme akışında son incelenen FROTO çıktıları ise Çizelge 4.9’da gösterilmiştir. LSTM tüm metriklerle bakıldığında yatırım karar kriterlerine göre en iyi kabul edilebilir tahmin performansını göstermiştir. Toplam doğruluk kriterine bakıldığında en iyi çıktıları veren RFE temelli yöntemler ise long-short dengesi bakımından kriterlerin dışında kalmıştır. Benzer şekilde mRMR yöntemleri de long-short doğruluklarına bakıldığında dengesiz bir dağılım sergilemiştir.

Çizelge 4.9 : FROTO 300 adımlı dizileme performans sonuçları.

MODEL	Long Doğruluk	Short Doğruluk	Ağırlıklı F1 Skor	Toplam Doğruluk
LSTM	0.6923	0.5882	0.6346	0.6333
GRU	0.3846	0.7059	0.5560	0.5667
HSIC-LASSO-LSTM	0.7692	0.4118	0.5565	0.5667
HSIC-LASSO-GRU	0.4615	0.5882	0.5333	0.5333
MRMR-LSTM	0.3846	0.7647	0.5847	0.6000
MRMR-GRU	0.6923	0.4706	0.5643	0.5667
RFE-LSTM	0.4615	0.8235	0.6539	0.6667
RFE-GRU	0.3077	0.9412	0.6243	0.6667

4.4 Geri Bakma Değeri 400 Olan Dizilemede Tahmin Performansı

Çizelge 4.10'daki veriler de AKBNK üzerindeki en başarılı genel performansın RFE-GRU ile sağlandığını göstermektedir. RFE-GRU ile GRU'ya kıyasla long doğruluk ciddi bir şekilde artırılmış ve toplam doğruluk %70 seviyesine gelmiştir. Bunun yanında LSTM'ler ile de yatırım kararlarında kullanılabilecek başarılı çıktıların edinildiği görülmektedir. HSIC-Lasso-LSTM ile LSTM'e göre hem short hem de long doğrulukları artırılarak performans iyileşmesi sağlandığı görülmektedir. mRMR modelleri ise kriterlerin dışında kalmıştır.

Çizelge 4.10 : AKBNK 400 adımlı dizileme performans sonuçları.

MODEL	Long Doğruluk	Short Doğruluk	Ağırlıklı F1 Skor	Toplam Doğruluk
LSTM	0.5000	0.6250	0.5652	0.5667
GRU	0.4286	0.7500	0.5889	0.6000
HSIC-LASSO-LSTM	0.5714	0.6875	0.6321	0.6333
HSIC-LASSO-GRU	0.5000	0.5000	0.5006	0.5000
MRMR-LSTM	0.4286	0.6250	0.5291	0.5333
MRMR-GRU	0.4286	0.6250	0.5291	0.5333
RFE-LSTM	0.4286	0.6250	0.5291	0.5333
RFE-GRU	0.6429	0.7500	0.6990	0.7000

TCELL verileri Çizelge 4.11'e göre incelendiğinde HSIC-Lasso-LSTM ile LSTM performansı kabul edilebilir performans değerlerine çıkmıştır. Tüm GRU mimarileri long-short doğruluklarına bakıldığında dengesiz kötü performans göstermiştir. mRMR yöntemleri short doğrulukları iyileştirse de long doğruluk performansları ile kriterleri sağlayamamıştır.

Çizelge 4.11 : TCELL 400 adımlı dizileme performans sonuçları.

MODEL	Long Doğruluk	Short Doğruluk	Ağırlıklı F1 Skor	Toplam Doğruluk
LSTM	0.5000	0.5000	0.5006	0.5000
GRU	0.2857	0.6875	0.4794	0.5000
HSIC-LASSO-LSTM	0.5714	0.5625	0.5671	0.5667
HSIC-LASSO-GRU	0.0714	1.0000	0.4415	0.5667
MRMR-LSTM	0.2857	0.6250	0.4519	0.4667
MRMR-GRU	0.2143	0.7500	0.4615	0.5000
RFE-LSTM	0.6429	0.4375	0.5292	0.5333
RFE-GRU	0.0714	0.8750	0.4022	0.5000

Son olarak Çizelge 4.12’de FROTO tahmin performansları verilmiştir. LSTM modeli hem short hem de long doğrulukta %60 üzerinde performans sergilemiştir. HSIC-Lasso-LSTM ve mRMR-GRU modellerinin çıktıları aynı ve kabul edilebilir yatırım kriterleri arasındadır. RFE-GRU ve RFE-LSTM performans metriklerine bakıldığında long doğruluk üzerinde GRU ve LSTM’e göre ciddi kötüleşmeye neden olmuştur.

Çizelge 4.12 : FROTO 400 adımlı dizileme performans sonuçları.

MODEL	Long Doğruluk	Short Doğruluk	Ağırlıklı F1 Skor	Toplam Doğruluk
LSTM	0.6154	0.6471	0.6346	0.6333
GRU	0.4615	0.8235	0.6539	0.6667
HSIC-LASSO-LSTM	0.6154	0.5882	0.6018	0.6000
HSIC-LASSO-GRU	0.4615	0.6471	0.5642	0.5667
MRMR-LSTM	0.4615	0.7059	0.5944	0.6000
MRMR-GRU	0.6154	0.5882	0.6018	0.6000
RFE-LSTM	0.3846	0.7647	0.5847	0.6000
RFE-GRU	0.1538	1.0000	0.5437	0.6333

4.5 Geri Bakma Değerleri İçin Performans Karşılaştırması

AKBNK hisse senedi için sınıflandırma performanslarının geri bakma değerlerine göre karşılaştırılması yapıldığında en yüksek genel başarının LSTM için %70 toplam doğruluk ile 200 adımlı dizileme ile sağlandığı görülmektedir. GRU modelleri tüm dizilerde benzer performanslar göstermiştir; ancak başarı kriterlerine en uygun olan model %60 toplam doğruluk çıktısı veren 200 adımlı dizileme içindir. HSIC-Lasso-LSTM modeli ise 400 adımlı dizileme için en iyi performansını toplam doğrulukta %63,33 olarak vermiştir. HSIC-Lasso-GRU modellerine bakıldığında LSTM ve GRU’ya benzer şekilde en yüksek doğruluk ile tahminlerin 200 adımlı dizilemede %66,67 olarak sağlandığı görülmektedir. mRMR-LSTM modeli ise belirlenen başarı kriterlerini sağladığı 100 adımlı dizide %56,67 toplam doğruluk performansı göstermiştir. mRMR-GRU’ya bakıldığında 200 adımlı dizilemede gerekli kriterleri sağlamış ve %56,67 toplam tahmin doğruluk performansı sergilemiştir. RFE-LSTM modeli hiçbir dizi için beklenen kriterleri sağlayamamıştır. RFE-GRU ise sadece 400 adımlı dizilemede başarı kriterlerini sağlamış ve toplam %70 tahmin doğruluğuna ulaşmıştır. Genel olarak bakıldığında tüm dizileme senaryolarında en optimum

performans LSTM ile sağlanmıştır. Tekil olarak bakıldığında ise %70 tahmin doğruluğu ve dengeli short-long doğruluk oranıyla en başarılı model 400 adımlı dizileme kullanan RFE-GRU olmuştur. Kullanılan tüm modeller baz alındığında ise 200 adımlı dizileme yaklaşımının AKBNK hisse senedi tahmini için en başarılı yöntem olduğu görülmüştür.

TCELL hisse senedi için sınıflandırma performanslarının geri bakma değerlerine göre karşılaştırılması yapıldığında LSTM modelinde 300 adımlı dizilemede kabul kriterlerini sağladığı ve %60 toplam doğruluk ile en iyi çıktının alındığı görülmektedir. GRU modelleri tüm dizileme senaryolarında kötü performans göstermiş ve istenen kriterleri sağlayamamıştır. HSIC-Lasso-LSTM modeli ise 200, 300 ve 400 adımlı dizilemelerde de kriterleri sağlamış 3 senaryoda da %56,67 toplam doğruluk ile tahmin yapmıştır. HSIC-Lasso-GRU modeli de GRU gibi hiçbir dizide beklenen performansı gösterememiştir. 300 ve 400 adımlı dizilemelerde long doğruluğun çok fazla düştüğü 100 ve 200 adımlı dizilemeler için daha iyi sonuçlar alındığı görülmektedir. mRMR-LSTM modeline bakıldığında kriterleri sağlayan en iyi sonucun %56,67 toplam doğruluk ile 300 adımlı dizilemede alındığı görülmektedir. mRMR-GRU modeli de tüm dizi yaklaşımlarında beklenen kriterlerin dışında kalmıştır. Belirli bir dizi kalıbının etkisi görülmemiştir. RFE-LSTM modeli tüm dizileme yaklaşımlarında istenen performansı göstermemiştir. Ancak F1 skorlarına bakıldığında en iyi performansı gösterdiği senaryo 400 adımlı dizileme olmuştur. RFE-GRU modeli de tüm diziler için kriterleri sağlayamamıştır, benzer şekilde F1 skorlarına bakıldığında 100 ve 200 adımlı dizilemelerde 300 ve 400'e göre daha iyi performans sergilemiştir.

FROTO hisse senedi için sınıflandırma performanslarının geri bakma değerlerine göre karşılaştırılması yapıldığında LSTM modelinde 300 ve 400 adımlı dizilemelerde long-short doğruluk performans kriterleri sağlanmış, aynı F1 skor alınarak %63,33 toplam doğruluk oranı yakalanmıştır. GRU modeline bakıldığında ise tüm performans metrikleri göz önüne alındığında en iyi genel performans 200 adımlı dizileme için %63,33 toplam doğruluk oranıyla yakalanmıştır. HSIC-Lasso-LSTM modeli ise en iyi çıktıları 200 ve 400 adımlı dizilerde vermiştir. Bu senaryolar için yakın F1 skorlar alınmış ve %60 toplam doğruluk elde edilmiştir. HSIC-Lasso-GRU modeli ise 100 adımlı dizileme için short doğruluk tarafında belirlenen kriteri sağlamasa da 200, 300 ve 400 adımlı dizilemelere göre daha iyi performans göstermiştir. mRMR-LSTM de

belirlenen yatırım kriterlerini hiçbir senaryoda sağlayamasa da F1 skorları baz alındığında en iyi performansını 400 adımlı dizileme için toplam %60 doğruluk ile göstermiştir. mRMR-GRU modeli de 100 ve 400 adımlı dizilemeler için aynı performansı başarımlı kriterlerini sağlayarak %60 toplam doğruluk oranı ile vermiştir. RFE-LSTM modeline bakıldığında 300 adımlı dizilemede en iyi performansı gösterdiği görülmektedir. RFE-GRU modeli ise tüm başarımlı kriterlerini 200 adımlı dizilemede %66,67 toplam başarımlı oranıyla sağlamıştır.

4.6 Özellik Seçim Yöntemleri İçin Performans Karşılaştırması

Bu bölümde LSTM ve GRU model performanslarını yatırım kabul kriterlerini sağlayarak daha da geliştirmiş özellik seçim yöntemleri dizileme adımları dikkate alınarak belirtilecektir. Tek bir performans metriğinde iyileşme sağlayan yöntemler kapsam dışı bırakılmıştır. Genel olarak uygulanan yöntem ile ilgili yatırım kabul kriteri sağlanmışsa ve baz modelin toplam doğruluk performansı iyileştirilmişse karşılaştırma kapsamına alınmıştır.

AKBNK hisse senedi için sınıflandırma performanslarının özellik seçim yöntemlerine göre karşılaştırılması yapıldığında 100 adımlı dizilemede mRMR boyut indirgeme yönteminin LSTM'in toplam doğruluk performansını %53,33'ten %56,67'ye çıkardığı görülmektedir. 200 adımlı dizilemeye bakıldığında ise HSIC-Lasso özellik seçimi ile GRU toplam doğruluğunun %60'tan %66,67'ye çıktığı görülmektedir. 400 adımlı dizileme senaryosunda ise RFE ile indirgenen veri kümesi sonrasında GRU model performansı toplam doğruluk baz alındığında %60'tan %70'e çıkmıştır. Aynı akışta HSIC-Lasso yöntemi de LSTM'de toplam doğruluğu %56,67'den %63,33'e çıkarmıştır.

TCELL hisse senedi için sınıflandırma performanslarının özellik seçim yöntemlerine göre karşılaştırılması yapıldığında 200 adımlı dizilemede HSIC-Lasso yöntemi ile LSTM toplam doğruluk çıktısı %53,33'ten %56,67'ye yükselmiştir. 400 adımlı dizileme için bakıldığında da aynı özellik seçim yöntemi ile LSTM doğruluk performansının %50'den %56,67'ye çıktığı görülmektedir.

FROTO hisse senedi için sınıflandırma performanslarının özellik seçim yöntemlerine göre karşılaştırılması yapıldığında 200 adımlı dizilemede HSIC-Lasso yönteminin

LSTM performansını ciddi ölçüde artırarak toplam doğruluk oranını %46,67'den %60'a çıkardığı görülmektedir. Buna ek olarak RFE yöntemiyle kullanılan GRU modelinde de toplam tahmin doğruluk oranı %63,33'ten %66,67'ye yükselmiştir. 100 adımlı dizileme baz alındığında da mRMR yöntemi ile boyut indirgeme yapılmış veri kümesi üzerinde GRU modelinin toplam doğruluk performansı %56,67'den %60'a çıkmıştır.





5. SONUÇLAR VE ÖNERİLER

Finansal piyasalarda karar vermek piyasaları etkileyen birçok farklı faktör göz önüne alındığında zor bir süreci ifade etmektedir. Zaman serisi olarak modellenen finansal verilerin doğrusal olmayan dinamik ve kaotik karakteristiği geleceğin tahmin edilebilmesini oldukça zorlaştırmaktadır.

Bu çalışmada yinelemeli sinir ağları kullanılarak Borsa İstanbul’da işlem gören farklı sektörlerden 3 şirketin hisse senedi yön hareketlerinin tahmini amaçlanmıştır. Derin öğrenme modellerinin tahmin performansının, zaman serisi olarak modellenen finansal veri üzerinde farklı adımlı dizileme yaklaşımları ve hisse senedi veri kümeleri üzerinde doğrusal olmayan özellik seçim yöntemleri kullanılarak yapılacak boyut indirgeme yöntemleriyle sağlanan çıktıların karşılaştırılmasının hedeflendiği tasarımlar belirli bir metodolojiyi takip etmiştir.

Çalışma Akbank, Turkcell ve Ford Otosan için veri kümelerinin Reuters Haber Ajansı kaynak alınarak dönemsel şirket finansal verileri, teknik indikatörler, ulusal ve uluslararası borsa endeks verileri, emtia ve döviz fiyatları, makroekonomik göstergeler, finansal oranlar gibi çok çeşitli verilerin sıralı 1270 piyasa işlem günü için hazırlanması ile başlamaktadır.

Veri işleme süreçlerinden sonra hazırlanan zaman serisi verilerine bir sonraki işlem günündeki yön hareketine göre “AL” ve “SAT” anlamını taşıyan ikili etiketleme uygulanarak verinin sınıflandırma algoritmalarına elverişli hale getirilmesi sağlanmıştır. Etiketleme çalışmasının tamamlanmasının ardından tüm özellik uzayı sayısal değerlerden oluşan veri kümeleri, sinir ağları üzerinde optimum sonuçlara daha performanslı ulaşabilmek amacıyla normalizasyon işlemiyle 0 ve 1 arasında ölçeklendirilmiştir.

Veri normalizasyonundan sonra doğrusal olmayan finansal veri karakteristiğindeki bağımlılıkları tespit ederek yapılacak boyut indirgemenin tahmin performansları üzerindeki etkisini inceleyebilmek amacıyla doğrusal olmayan 3 özellik seçim yöntemi, hazırlanan veri kümeleri üzerinde uygulanmış ve yapılan stres testleri neticesinde özellik uzayı 70 olacak şekilde indirgenmiştir.

Tahmin tasarımının devamında zaman serisi verileri derin sinir ağlarının girdilerini oluşturacak zaman dizilerine ayrılmıştır. Dizileme çalışmasında 100, 200, 300 ve 400 sıralı zaman adımlarıyla oluşturulan iki boyutlu zaman dizileri, tüm zaman serisi üzerinde birer adım kaydırılan pencereler yöntemi ile oluşturulmuştur. Dizi etiketlemesinin de beraberinde yapıldığı bu yöntemle gelecek işlem günündeki hisse senedi yönünün tahmini için sıralı geçmiş zaman serisi gözlem gruplarıyla eğitilen modellerden yararlanılmıştır.

Tahminlemenin son aşamasında zaman dizileri, tasarlanan LSTM ve GRU ağlarına girdi olarak verilerek sınıf etiketlerine göre eğitim aşaması gerçekleştirilmiştir. 1240 işlem günü üzerinde farklı geri bakma değerleri ile hazırlanan diziler model eğitimi için kullanılmış ve gelecek 30 işlem günündeki hisse senetlerinin yön hareketleri tahminlemesi yapılmıştır.

Çalışmada elde edilen model tahminleme çıktıları AKBNK, TCELL ve FROTO hisse senetleri için 4 farklı geri bakma değeri ile sağlanan dizilemeler için sunulmuş, devamında geri bakma değerleri ve özellik seçim yöntemleri ayrı olarak ele alınarak performans karşılaştırmalarına yer verilmiştir. Performans karşılaştırmaları belirlenen 4 metrik çerçevesinde yapılmıştır. Bu metrikler “AL” kararını ölçen long doğruluk, “SAT” kararını ölçen short doğruluk ve genel model başarımlarını ölçen ağırlıklı F1 skor ile toplam tahmin doğruluğunu ölçen toplam doğruluk metrikleridir. Karar verme aşamasında modelin yatırım kabul kriteri olarak long ve short doğrulukta %50 üzerinde dengeli performans sergilemesi ve toplam tahmin doğruluğunda %55 seviyesinin üzerine çıkabilmesi belirlenmiştir. Performans metrikleri bu açıdan değerlendirilmiş, en az bir metrikte yatırım kabul kriterini sağlamayan model başarısız olarak değerlendirilmiş ve performans açısından iyileşmeye açık olarak yorumlanmıştır.

Tasarlanan yapı çerçevesinde sunulan yöntem ve modellerin farklı sektörden 3 şirketin hisse senetleri üzerinde de yatırım kararlarında yararlanılabilecek yüksek tahmin performansları ortaya koyabileceğini göstermiştir. Özellik seçim yöntemleri ile LSTM ve GRU derin sinir ağı mimarilerinin sağladığı tahmin performansının iyileştirilebileceği, sunulan akışın portföy yönetimi kararlarında uygulanabilirliği ortaya koyulmuştur. Kaotik, dinamik ve doğrusal olmayan finansal verilerin, geçmiş gözlem noktalarının dizilenmesi ile derin sinir ağları üzerindeki eğitim aşamalarından sonra tahmin edilebilirliğinin gösterildiği çalışma ayrıca ortaya farklı gelişim alanları da koymaktadır. LSTM ve GRU model hiper parametrelerinin optimizasyonu ile elde edilen sonuçlar iyileştirilebilir. Tahmin stratejisi yönünden ise modeller belirli bir gün için tahmin yaptıktan sonra güncel veri ile yeniden eğitilerek yinelemeli bir tahmin yöntemi ile hisse senedi yön tahmin başarımları daha da iyileştirilebilir.



KAYNAKLAR

- [1] Zahedi, J., & Rounaghi, M. M. (2015). Application of artificial neural network models and principal component analysis method in predicting stock prices on Tehran Stock Exchange. *Physica A*, 438, 178–187. <https://doi.org/10.1016/j.physa.2015.06.033>
- [2] Leung, M. T., Daouk, H., & Chen, A.-S. (2000). Forecasting stock indices: a comparison of classification and level estimation models. *International Journal of Forecasting*, 16(2), 173–190. [https://doi.org/10.1016/s0169-2070\(99\)00048-5](https://doi.org/10.1016/s0169-2070(99)00048-5)
- [3] Adhikari, R., & Agrawal, R. K. (2013). *An introductory study on time series modeling and forecasting*. <https://doi.org/10.48550/ARXIV.1302.661>
- [4] Raicharoen, T., Lursinsap, C., & Sanguanbhokai, P. (2003). Application of critical support vector machine to time series prediction. Proceedings of the 2003 International Symposium on Circuits and Systems, 2003. ISCAS '03.
- [5] Pai, P.-F., & Lin, C.-S. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega*, 33(6), 497–505. <https://doi.org/10.1016/j.omega.2004.07.024>
- [6] Menon, V. K., Chekravarthi Vasireddy, N., Jami, S. A., Pedamallu, V. T. N., Sureshkumar, V., & Soman, K. P. (2016). Bulk price forecasting using spark over NSE data set. In *Data Mining and Big Data* (pp. 137–146). Springer International Publishing.
- [7] Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159–175. [https://doi.org/10.1016/s0925-2312\(01\)00702-0](https://doi.org/10.1016/s0925-2312(01)00702-0)
- [8] Corberán-Vallet, A., Bermúdez, J. D., & Vercher, E. (2011). Forecasting correlated time series with exponential smoothing models. *International Journal of Forecasting*, 27(2), 252–265. <https://doi.org/10.1016/j.ijforecast.2010.06.003>
- [9] Ma, D., Yang, T., Liu, L., & He, Y. (2022). Analysis of factors influencing stock market volatility based on GARCH-MIDAS model. *Complexity*, 2022, 1–10. <https://doi.org/10.1155/2022/6176451>

- [10] Yadav, S., & Sharma, N. (2019). Forecasting of Indian stock market using time-series models. In *Lecture Notes in Networks and Systems* (pp. 405–412). Springer Singapore.
- [11] Sethia, A., & Raut, P. (2019). Application of LSTM, GRU and ICA for stock price prediction. In *Information and Communication Technology for Intelligent Systems* (pp. 479–487). Springer Singapore.
- [12] Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, 38(1), 34. <https://doi.org/10.1086/294743>
- [13] Fama, E. F. (1995). Random walks in stock market prices. *Financial Analysts Journal*, 51(1), 75–80. <https://doi.org/10.2469/faj.v51.n1.1861>
- [14] Basak, S., Kar, S., Saha, S., Khaidem, L., & Dey, S. R. (2019). Predicting the direction of stock market prices using tree-based classifiers. *The North American Journal of Economics and Finance*, 47, 552–567. <https://doi.org/10.1016/j.najef.2018.06.013>
- [15] Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. *Expert Systems with Applications*, 124, 226–251. <https://doi.org/10.1016/j.eswa.2019.01.012>
- [16] Siامي-Namini, S., Tavakoli, N., & Siامي Namin, A. (2018). A comparison of ARIMA and LSTM in forecasting time series. *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*.
- [17] Ghosh, A., Bose, S., Maji, G., Debnath, N., & Sen, S. (2019). *Stock price prediction using LSTM on Indian share market*.
- [18] Ji, G., Yu, J., Hu, K., Xie, J., & Ji, X. (2022). An adaptive feature selection schema using improved technical indicators for predicting stock price movements. *Expert Systems with Applications*, 200(116941), 116941. <https://doi.org/10.1016/j.eswa.2022.116941>
- [19] Zhong, X., & Enke, D. (2019). Predicting the daily return direction of the stock market using hybrid machine learning algorithms. *Financial Innovation*, 5(1). <https://doi.org/10.1186/s40854-019-0138-0>
- [20] Moghar, A., & Hamiche, M. (2020). Stock market prediction using LSTM recurrent neural network. *Procedia Computer Science*, 170, 1168–1173. <https://doi.org/10.1016/j.procs.2020.03.049>
- [21] Budiharto, W. (2021). Data science approach to stock prices forecasting in Indonesia during Covid-19 using Long Short-Term Memory (LSTM). *Journal of Big Data*, 8(1), 47. <https://doi.org/10.1186/s40537-021-00430-0>

- [22] Kwon, Y.-K., & Sun, H.-D. (2011). A hybrid system integrating a piecewise linear representation and a neural network for stock prediction. *Proceedings of 2011 6th International Forum on Strategic Technology*.
- [23] Borovkova, S., & Tsiamas, I. (2019). An ensemble of LSTM neural networks for high-frequency stock market classification. *Journal of Forecasting*, *for*.2585. <https://doi.org/10.1002/for.2585>
- [24] Gao, Y., Wang, R., & Zhou, E. (2021). Stock prediction based on optimized LSTM and GRU models. *Scientific Programming*, 2021, 1–8. <https://doi.org/10.1155/2021/4055281>
- [25] Faurina, R., Winduratna, B., & Nugroho, P. 2018. Predicting Stock Movement Using Unidirectional LSTM and Feature Reduction: The Case of An Indonesia Stock. In 2018 International Conference on Electrical Engineering and Computer Science (ICEECS), (pp.180–185).
- [26] M, H., E.a., G., Menon, V. K., & K.p., S. (2018). NSE stock market prediction using deep-learning models. *Procedia Computer Science*, 132, 1351–1362. <https://doi.org/10.1016/j.procs.2018.05.050>
- [27] Liu, Y., & Liu, X. (2018). A trend-based stock index forecasting model with gated recurrent neural network. *2018 IEEE International Conference on Progress in Informatics and Computing (PIC)*.
- [28] Shah, R., Tambe, A., Bhatt, T., & Rote, U. (2020). Real-time stock market forecasting using ensemble deep learning and rainbow DQN. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3586788>
- [29] Nezhad, M. T. F., & Rezaei, M. (2022). Stock price prediction using intelligent models, Ensemble Learning and feature selection. *2022 Second International Conference on Distributed Computing and High Performance Computing (DCHPC)*.
- [30] Niu, T., Wang, J., Lu, H., Yang, W., & Du, P. (2020). Developing a deep learning framework with two-stage feature selection for multivariate financial time series forecasting. *Expert Systems with Applications*, 148(113237), 113237. <https://doi.org/10.1016/j.eswa.2020.113237>
- [31] Lee, J. S., Sung Cho, H., Chung, K. I., & Sang Park, J. (2020). Feature Selection for Stock forecasting using Multivariate Convolution Neural Network. *2020 International Conference on Information and Communication Technology Convergence (ICTC)*.
- [32] Liu, J., Sun, H., Li, Y., Fang, W., & Niu, S. (2020). An improved power system transient stability prediction model based on mRMR feature selection and WTA ensemble learning. *Applied Sciences (Basel, Switzerland)*, 10(7), 2255. <https://doi.org/10.3390/app10072255>

- [33] A stable SVM-RFE Feature Selection method for gene expression data. (2019). *International Journal of Engineering and Advanced Technology*, 8(6), 2110–2115. <https://doi.org/10.35940/ijeat.f8482.088619>
- [34] Ghosh, P., Azam, S., Jonkman, M., Karim, A., Shamrat, F. M. J. M., Ignatious, E., Shultana, S., Beeravolu, A. R., & De Boer, F. (2021). Efficient prediction of cardiovascular disease using machine learning algorithms with relief and LASSO feature selection techniques. *IEEE Access: Practical Innovations, Open Solutions*, 9, 19304–19326. <https://doi.org/10.1109/access.2021.3053759>
- [35] Guha Majumdar, S., Rai, A., & Mishra, D. C. (2020). Integrated framework for selection of additive and nonadditive genetic markers for genomic selection. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, 27(6), 845–855. <https://doi.org/10.1089/cmb.2019.0223>
- [36] McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4), 115–133. <https://doi.org/10.1007/bf02478259>
- [37] Chen, X., & Sun, L. (2021). Bayesian temporal factorization for multidimensional time series prediction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PP, 1–1. <https://doi.org/10.1109/TPAMI.2021.3066551>
- [38] Adhikari, R., & Agrawal, R. K. (2013). *An introductory study on time series modeling and forecasting*. <https://doi.org/10.48550/ARXIV.1302.6613>
- [39] Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408. <https://doi.org/10.1037/h0042519>
- [40] Minsky, M., Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA, USA: MIT Press.
- [41] P. Werbos, “Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences,” Ph.D. Thesis, Harvard University, Cambridge, 1974.
- [42] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536. <https://doi.org/10.1038/323533a0>
- [43] Hochreiter, S. (1991). Untersuchungen zu dynamischen neuronalen Netzen. Diploma thesis, Institut für Informatik, Lehrstuhl Prof. Brauer, Technische Universität München
- [44] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>

- [45] Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE. Institute of Electrical and Electronics Engineers*, 86(11), 2278–2324. <https://doi.org/10.1109/5.726791>
- [46] Gurney, K. (2018). *An Introduction to Neural Networks*. CRC Press.
- [47] Sherstinsky, A. (2020). Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D. Nonlinear Phenomena*, 404(132306), 132306. <https://doi.org/10.1016/j.physd.2019.132306>
- [48] Zaremba, W., Sutskever, I., and Vinyals, O. Recurrent neural network regularization. arXiv preprint arXiv:1409.2329, 2014.
- [49] Li, S., Li, W., Cook, C., Zhu, C., & Gao, Y. (2018). Independently recurrent neural network (IndRNN): Building A longer and deeper RNN. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*.
- [50] Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). *Learning phrase representations using RNN encoder-decoder for statistical machine translation*. <https://doi.org/10.48550/ARXIV.1406.1078>
- [51] Lee, M.-C. (2022). Research on the feasibility of applying GRU and Attention mechanism combined with technical indicators in stock trading strategies. *Applied Sciences (Basel, Switzerland)*, 12(3), 1007. <https://doi.org/10.3390/app12031007>
- [52] Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., & Liu, H. (2018). Feature selection: A data perspective. *ACM Computing Surveys*, 50(6), 1–45. <https://doi.org/10.1145/3136625>
- [53] Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, 67(2), 301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>
- [54] Liu, H., & Motoda, H. (Eds.). (2007). *Computational methods of feature selection*. Chapman and Hall/CRC.
- [55] Forman, G. (2002). Choose your words carefully: An empirical study of feature selection metrics for text classification. In *Principles of Data Mining and Knowledge Discovery* (pp. 150–162). Springer Berlin Heidelberg.
- [56] Pai, N., & Ilango, V. (2020). LSTM Neural Network Model with Feature selection for Financial Time series Prediction. *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*.

- [57] Chowdhury, S., Dong, X., & Li, X. (2019). Recurrent neural network based feature selection for high dimensional and low sample size micro-array data. *2019 IEEE International Conference on Big Data (Big Data)*.
- [58] Saeys, Y., Inza, I., & Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics. *Bioinformatics (Oxford, England)*, 23(19), 2507–2517. <https://doi.org/10.1093/bioinformatics/btm344>
- [59] Yadav, S., & Sharma, N. (2019). Forecasting of Indian stock market using time-series models. In *Lecture Notes in Networks and Systems* (pp. 405–412). Springer Singapore.
- [60] Yamada, M., Jitkrittum, W., Sigal, L., Xing, E. P., & Sugiyama, M. (2014). High-dimensional feature selection by feature-wise kernelized Lasso. *Neural Computation*, 26(1), 185–207. https://doi.org/10.1162/NECO_a_00537
- [61] Takahashi, Y., Ueki, M., Yamada, M., Tamiya, G., Motoike, I. N., Saigusa, D., Sakurai, M., Nagami, F., Ogishima, S., Koshihara, S., Kinoshita, K., Yamamoto, M., & Tomita, H. (2020). Improved metabolomic data-based prediction of depressive symptoms using nonlinear machine learning with feature selection. *Translational Psychiatry*, 10(1), 157. <https://doi.org/10.1038/s41398-020-0831-9>
- [62] Zhao, Z., Anand, R., & Wang, M. (2019). Maximum relevance and minimum redundancy feature selection methods for a marketing machine learning platform. *2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*.
- [63] Ding, C., & Peng, H. (2004). Minimum redundancy feature selection from microarray gene expression data. *Computational Systems Bioinformatics. CSB2003. Proceedings of the 2003 IEEE Bioinformatics Conference. CSB2003*.
- [64] Radovic, M., Ghalwash, M., Filipovic, N., & Obradovic, Z. (2017). Minimum redundancy maximum relevance feature selection approach for temporal gene expression data. *BMC Bioinformatics*, 18(1), 9. <https://doi.org/10.1186/s12859-016-1423-9>
- [65] Kim, H.-J., Baek, J.-W., & Chung, K. (2021). Associative knowledge graph using fuzzy clustering and min-max normalization in video contents. *IEEE Access: Practical Innovations, Open Solutions*, 9, 74802–74816. <https://doi.org/10.1109/access.2021.3080180>

ÖZGEÇMİŞ

Ad-Soyad : Muhiddin Cağlar Eren

ÖĞRENİM DURUMU:

- **Lisans** : 2015, İstanbul Teknik Üniversitesi, Elektrik - Elektronik Fakültesi, Elektronik ve Haberleşme Mühendisliği

MESLEKİ DENEYİM VE SERTİFİKALAR:

- 2014-2015: Mercedes Benz – Sistem Proje Asistanı
- 2015-2015: Turkcell – Orta Katman Mühendisi
- 2015-2017: Avea – Sistem Mühendisi
- 2018-2019: Garanti BBVA Teknoloji – Sistem Mühendisi
- 2019-2022: Yapı Kredi Teknoloji – Sistem Mühendisi
- 2022- ... : Trendyol – DevOps Mühendisi