MARMARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

MOBİL UYGULAMALARIN GELECEK DÖNEMDE KULLANIMININ ZAMAN SERİLERİ KULLANILARAK TAHMİN EDİLMESİ

UĞUR İNCE

YÜKSEK LİSANS TEZİ

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Tezli Yüksek Lisans Programı

DANIŞMAN

Dr. Öğretim Üyesi Zehra Aysun ALTIKARDEŞ

İSTANBUL, 2021

MARMARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

MOBİL UYGULAMALARIN GELECEK DÖNEMDE KULLANIMININ ZAMAN SERİLERİ KULLANILARAK TAHMİN EDİLMESİ

UĞUR İNCE
(523618020)

YÜKSEK LİSANS TEZİ

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Tezli Yüksek Lisans Programı

DANIŞMAN

Dr. Öğretim Üyesi Zehra Aysun ALTIKARDEŞ

İSTANBUL, 2021

TEŞEKKÜR

Tez çalışmam boyunca desteğini hiçbir zaman esirgemeyen, her ihtiyacım olduğunda yanımda olan, bilgi birikimiyle yol gösteren değerli danışmanım Sayın Dr. Öğr. Üyesi Zehra Aysun Altıkardeş'e ve tez sürecince ilgisiyle ve değerli yönlendirmeleriyle yardımını esirgemeyen Sayın Dr. Öğr. Üyesi Buket Doğan'a teşekkür ederim.

Yüksek lisans eğitimim boyunca bana güvenen ve bu süreçte beni destekleyip cesaretlendiren, tez çalışmasını başarıyla sonlandırmam için hiçbir fedakarlıktan kaçınmayan değerli eşim Sevinç İnce'ye teşekkür ederim.

Mart, 2021 **Uğur İNCE**

İÇİNDEKİLER

TEŞEKKÜR i
İÇİNDEKİLER ii
ÖZET iv
ABSTRACT v
SEMBOLLER vi
KISALTMALAR vii
ŞEKİL LİSTESİ ix
TABLO LİSTESİ xi
1. GİRİŞ 1
1.1. Literatür Araştırması 2
2. MATERYAL VE YÖNTEM 6
2.1. Zaman Serileri ve Zaman Serileri Analizi 7
2.2. Doğrusal Regresyon 9
2.2.1. Basit Doğrusal Regresyon 9
2.2.2. Çoklu Doğrusal Regresyon 10
2.3. BOX-Jenkins Yöntemi 10
2.3.1. Otoregresif Modeli (AutoRegressive) 11
2.3.2. Hareketli Ortalama Modeli (Moving Average) 11
2.3.3. Otoregresif Hareketli Ortalama Modeli (AutoRegressive Moving Average) 12
2.3.4. ARIMA (AutoRegressive Integrated Moving Average) 12
2.4. Naive Model 13
2.5. Üstel Düzleştirme Yöntemleri 13
2.5.1. Basit Üstel Düzleştirme Yöntemi 14
2.5.2. Brown'un Tek Parametreli Doğrusal Üstel Düzleştirme Yöntemi 14
2.5.3. Holt'un Çift Parametreli Doğrusal Üstel Düzleştirme Yöntemi 15
2.6. Destek Vektör Makineleri 15
2.6.1. Doğrusal Destek Vektör Makineleri 15
2.6.2. Doğrusal Olmayan Destek Vektör Makineleri 17
2.7. Yapay Sinir Ağları(ANN) 18

2.8. Tahmin Yöntemlerinde Performans Ölçümü 21
2.8.1. Ortalama Hata (Mean Error - ME) 21
2.8.2. Ortalama Mutlak Hata (Mean Absolute Error – MAE) 21
2.8.3. Ortalama Mutlak Hata Yüzdesi (Mean Absolute Percentage Error – MAPE) 21
2.8.4. Ortalama Kare Hata (Mean Squared Error – MSE) 22
2.8.5. Kök Ortalama Kare Hata (Root Mean Squared Error - RMSE) 22
2.8.6. Ortalama Yüzde Hata (Mean Percentage Error -MPE) 22
3. DENEYLER VE BULGULAR 23
3.1. Basit Doğrusal Regresyon 23
3.2. BOX-Jenkins Modeli 25
3.3. Naive Model ile tahminleme 28
3.4. Üstel Düzleştirme Yöntemleri 31
3.4.1. Basit Üstel Düzleştirme Yöntemi 31
3.4.2. Üstel Düzleştirme Yöntemi 33
3.5. Destek Vektör Makinesi Kullanılarak Tahminleme 35
3.6. Yapay Sinir Ağları Kullanılarak Tahminleme (YSA) 37
4. SONUÇLAR 40
KAYNAKLAR 42

ÖZET

MOBİL UYGULAMALARIN GELECEK DÖNEMDE KULLANIMININ ZAMAN SERİLERİ KULLANILARAK TAHMİN EDİLMESİ

Bir veya daha fazla değişkenin zamana dayalı değişiminin gösterildiği serilere zaman serisi denir. İhtiyaca göre yıllık, çeyreklik, aylık, haftalık, günlük, saatlik gibi farklı periyotlarda gösterim yapılabilir. Zaman serileri üzerindeki geçmiş dönem gözlem değerleri kullanılarak gelecek dönemle ilgili öngörüde bulunmak mümkün olabilmektedir. Günümüzde zaman serileri ekonomi, finans, meteoroloji, tarım, mühendislik gibi birçok alanda kullanılmaktadır. Örnek olarak ekonomide yıllık ihracat ve ithalat miktarları, elektrik üretimi için mevsimsel/sıcaklığa dayalı tüketim miktarları, yıllık/aylık ortalama trafik kaza sayısı, tarım ve barajlar için aylık yağış miktarı, mevsimlere göre sıcaklık değer tahmini, personel planlaması için firmaların müşteri yoğunluk periyotları, seyahat planlamaları için zamana dayalı trafik yoğunluk haritasının oluşturulması, işletmelerin dönemsel/mevsimsel üretim ve stok planlaması gibi konularda yoğunlukla kullanılmaktadır. Bu gibi birçok alanda faaliyet gösteren kurum ve şirketlerin önemli kararlar alması ve stratejilerini belirlemesi için sonraki dönemlerle ilgili öngörülere ihtiyaç duymaktadır.

Çalışmamızda Türkiye'nin ileri gelen operatörlerinden Turkcell İletişim Hizmetleri A.Ş ye ait bir mobil uygulamanın ileriki dönemlerdeki kullanımının tahmin edilmesi amaçlanmıştır. Bu bağlamda belli bir zaman dilimindeki günlük kullanıcı giriş sayısı kullanılarak sonraki dönemlerde gerçekleşebilecek günlük giriş sayısı için tahmin çalışmaları yapılmıştır. Kullanılan istatistiksel yöntemler ve makine öğrenmesi yöntemlerinin tahmin başarımları karşılaştırıldığında en yüksek başarımın SVM ile elde edildiği arkasından da YSA yönteminin geldiği görülmüştür. İstatistiksel yöntemlerin başarımları ise daha sonradan gelmiştir. Gelecek dönem kullanım eğrisinin tahmin edilmesiyle birlikte kullanılan kaynakların optimizasyonu, kapasite ve altyapı planlaması, efektif reklam planlamalarının yapılabilmesi, beklenenin üzerinde ya da altında kullanım olması halinde normal olmayan bir durumun yaşandığının tespit edilmesi (sistemsel sorun, siber saldırı, kampanya etkisi, özel gün vs.) gibi avantajların sağlanması hedeflenmektedir.

ABSTRACT

MOBILE APPLICATIONS USAGE FORECASTING USING TIME SERIES

Series that show the time-based variation of one or more variables is called the time series. Depending on the need, annual, quarterly, monthly, weekly, daily and hourly periods can be displayed. It is possible to make predictions about the future period by using past period observation values on time series. Time series are used in many fields such as economy, finance, meteorology, agriculture, engineering. For example, annual export and import amounts in the economy, seasonal / temperature-based consumption amounts for electricity generation, annual / monthly average traffic accidents, monthly precipitation amount for agriculture and dams, temperature estimation according to seasons, customer density periods for personnel planning, travel It is used extensively in subjects such as creating a time-based traffic density map for planning, periodic / seasonal production and stock planning of enterprises. In order to make important decisions and to determine their strategies, companies and companies operating in many such fields need forecasts for the next periods.

In this thesis study is aimed to forecast the usage of a mobile application in the future periods. The data set belongs to one of the leading communication company which name is Turkcell. In this context, the number of daily user logins in a certain period of time is estimated for the number of daily entries that may occur in the following periods. When the prediction performance of statistical methods and machine learning methods were compared, it was seen that the highest performance was achieved with SVM, followed by ANN method. The performance of statistical methods is lower than machine learning methods. With the estimation of the future usage trend, optimization of the resources used, capacity and infrastructure planning, effective advertisement planning, determination of an abnormal situation(anomaly) in case of use above or below expected (system problem, cyber attack, campaign effect, special day, etc.)is aimed to be provided.

SEMBOLLER

y_t **:** bağımlı değişken

x_t **:** bağımsız değişken

β_0 **:** kesme noktası

β_1 **:** regresyon katsayısı(eğim)

x_1 **:** bağımsız değişken

εt **:** hata terimi

ϕ_1 **:** model parametreleri

P **:** model derecesi

y_t **:** t dönemine ait gözlem değeri

Θ_1 **:** model parametreleri

q **:** modelin derecesi MA(q)

α **:** düzleştirme sabiti

w **:** çoklu düzlemin normali(ağırlık vektörü)

x **:** veri noktası

b **:** bias

H **:** destek vektörleri

Lp **:** Lagrange değeri

d **:** polinom derecesi

σ **:** çekirdek ölçeği

δ **:** eşik değer parametresi

φ **:** Aktivasyon fonksiyonu

x_i **:** girdi

ω **:** Ağırlık

KISALTMALAR

3G : Üçüncü Nesil(3rd Generation)
A.Ş : Anonim Şirketi
ADF : Augmented Dickey–Fuller
ANN : Artificial Neural Network
AR : Auto Regressive
ARIMA : Autoregressive Integrated Moving Average
ARMA : Auto Regressive Moving Average
ATM : Automatic Teller Machine
BJ : Box Jenkins
DVM : Destek Vektör Makineleri
DVR : Destek Vektör Regresyonu
ETS : Exponential Smoothing
GPR : Gaussian Process Regression
GB : Gigabyte
KPSS : Kwiatkowski–Phillips–Schmidt–Shin
LR : Linear Regression
LSTM : Long short-term memory
MA : Moving average
MAE : Mean Absolute Error
MAPE : Mean Absolute Percentage Error
ME : Mean Error
MLP : Multi-layer Perceptron
MLR : Multiple linear regression
MPE : Mean Percentage Error
MSE : Mean Squared Error
PDA : Personal Data Assistant
RBF : Radial Basis Function
RMSD : Root Mean Squared Deviation
RMSE : Root Mean Squared Error
SES : Simple Exponential Smoothing

SVM **:** Support Vector Machines

SVR **:** Support Vector Regression

TASM **:** Technocrat ARIMA and SVR model

TB **:** Terabyte

TL **:** Türk Lirası

ts **:** Time Series

YSA **:** Yapay Sinir Ağları

ŞEKİL LİSTESİ

Şekil 2. 1. Akış Diyagramı 7
Şekil 2. 2. Örnek zaman serisi grafiği 9
Şekil 2. 3. İki sınıfın doğrusal ayrımı 16
Şekil 2. 4. Doğrusal ayrılamayan iki sınıfın çekirdek fonksiyonu kullanılarak ayrımı 17
Şekil 2. 5. Örnek YSA Yapısı 19
Şekil 2. 6. Yaygın kullanılan aktivasyon fonksiyonları 20
Şekil 3. 1. Aralık ayı gerçek değer ve lineer model tahmin grafiği 24
Şekil 3. 2. Ocak ayı gerçek değer ve lineer model tahmin grafiği 25
Şekil 3. 3. ADF durağanlık test sonucu 26
Şekil 3. 4. KPSS durağanlık test sonucu 26
Şekil 3. 5. Mevsimsellik test sonucu 26
Şekil 3. 6. KPSS durağanlık test sonucu 26
Şekil 3. 7. 2018 Aralık ayı tahmini için ARIMA modellemesi 26
Şekil 3. 8. 2019 Ocak ayı tahmini için ARIMA modellemesi 26
Şekil 3. 9. 2018 Aralık ayı gerçek değer ve ARIMA tahmin grafiği 27
Şekil 3. 10. 2019 Ocak ayı gerçek değer ve ARIMA tahmin grafiği 28
Şekil 3. 11. 11 ay eğitim veri seti mevsimsellik test sonucu 29
Şekil 3. 12. 12 ay eğitim veri seti mevsimsellik test sonucu 29
Şekil 3. 13. 31 günlük tahminleme 29
Şekil 3. 14. 2019 Ocak ayı gerçek değer ve Naive model tahmin grafiği 30
Şekil 3. 15. 2018 Aralık ayı gerçek değer ve Naive model tahmin grafiği 30
Şekil 3. 16. 2019 Ocak ayı gerçek değer ve basit üstel düzleştirme tahmin grafiği 32
Şekil 3. 17. 2018 Aralık ayı gerçek değer ve basit üstel düzleştirme tahmin grafiği 32
Şekil 3. 18. Üstel Düzleştirme Tahmin 33
Şekil 3. 19. 2019 Ocak ayı gerçek değer ve üstel düzleştirme modeli tahmin grafiği 34
Şekil 3. 20. 2018 Aralık ayı gerçek değer ve üstel düzleştirme modeli tahmin grafiği 35
Şekil 3. 21. SVM Model Oluşturma 36
Şekil 3. 22 2019 Ocak ayı gerçek değer ve SVM modeli tahmin grafiği 36
Şekil 3. 23. 2018 Aralık ayı gerçek değer ve SVM modeli tahmin grafiği 37
Şekil 3. 24. YSA model oluşturma 38

Şekil 3. 25. YSA tahmin 38
Şekil 3. 26. 2019 Ocak ayı gerçek değer ve SVM modeli tahmin grafiği 39
Şekil 3. 27. 2018 Aralık ayı gerçek değer ve SVM modeli tahmin grafiği 39

TABLO LİSTESİ

Tablo 2. 1. ARIMA model tablosu 12
Tablo 2. 2. Temel SVM çekirdek fonksiyonları 17
Tablo 2. 3. Sinir sistemi - YSA terminolojisi 20
Tablo 2. 4. YSA modeli ile istatistiksel modellere ait kavramların karşılıkları 20
Tablo 3. 1. Doğrusal regresyon tahmini hata oranları 24
Tablo 3. 2. ARIMA tahmini hata oranları 27
Tablo 3. 3. NAIVE tahmin hata oranları 29
Tablo 3. 4. Basit üstel düzleştirme tahmini hata oranları 31
Tablo 3. 5. Üstel düzleştirme tahmini hata oranları 33
Tablo 3. 6. SVM tahmini hata oranları 36
Tablo 3. 7. YSA tahmini hata oranları 38
Tablo 4. 1. Genel tahmin hata performans tablosu 41

1. GİRİŞ

Akıllı telefon, tablet, PDA, akıllı saat gibi taşınabilir aygıtları kapsayan ve çeşitli işlevleri gerçekleştiren elektronik cihazlara genel olarak mobil cihaz adı verilir. Mobil cihazlarda çalışması amacıyla geliştirilen yazılımlara ise mobil uygulama adı verilir. Mobil uygulamalar teknolojinin gelişimiyle birlikte hemen hemen herkesin hayatına önemli ölçüde girmiştir. Hayatımıza girdiği günden beri cihazların teknik özelliklerinin de gelişmesiyle uygulamalarda her geçen gün daha akıllı hale gelmiştir. Bununla birlikte kullanıcıların mobil cihazlarıyla geçirdikleri vakit de gün geçtikçe artmaya devam etmektedir. Kullanıcılar mobil cihazlarıyla geçirdikleri vaktin çok önemli bir kısmını da mobil uygulamaları kullanarak geçirmektedir.

Başta mobil cihaz üreticileri olmak üzere birçok işletme mobil uygulamalar aracılığı ile ürünlerini tanıtmaya, pazarlamaya, bilinirliğini artırmaya ve kullanıcı ihtiyaçlarının karşılanması için çeşitli çözümler sunmaya çalışmaktadır. Mobil uygulamalar kullanıcılara alışveriş, eğlence, bankacılık, iletişim, spor, kişisel gelişim, sosyal paylaşım ve haber alma gibi alanlarda yoğunlukla hizmet etmektedir. Bu durum, tüketici talepleri doğrultusunda belirtilen hizmetlerin sağlamasına yönelik mobil uygulamaların bilişim sektöründeki işletmelerce geliştirilmesine katkı da bulunmuştur. Bu nedenle de söz konusu mobil uygulamaların sayısı her geçen gün artmış ve günümüzde çok büyük bir hacme ulaşmıştır.

Bilişim Teknolojileri ve İletişim Kurumu tarafından 2020 yılı ilk çeyreği için yayınlanan "Türkiye Elektronik Haberleşme Sektörü 3 Aylık Pazar Verileri Raporu" verilerine göre Türkiye'de yaklaşık %98,4 erişim oranına karşılık 75.371.952 4.5G abonesi ve 4.625.682 3G abonesi bulunmaktadır. 2020 yılı birinci çeyrekte kullanılan toplam mobil internet miktarının ise 1.397.404 TB olarak gerçekleştiği belirtilmektedir [1]. Bu durum, kullanıcıların web sayfalarını ziyaret etmek, alışveriş yapmak ve sosyal medya kullanımı gibi işlemlerini mobil aygıtlar üzerinden yoğunlukla gerçekleştirdiğinin bir göstergesi olarak değerlendirilebilir.

Mobil cihaz ve uygulama kullanımının günlük hayatın bir parçası haline gelmesiyle bu uygulamaları sunan firma/şirketler içinde uygulamanın mevcut kullanımının analiz edilerek gelecekteki kullanımının tahminini yapabilmek de önem kazanmıştır.

Uygulamadaki aktif kullanıcı sayısı, satış sayısı, ürün/sayfa görüntüleme ve abonelik sayısı gibi uygulamaya özel önemli işlemlerin gelecek dönemler için tahmin edilmesi birçok avantaj sağlayacağından ön plana çıkmaktadır. Platform kullanımının artış ya da azalış eğiliminde olmasına göre gelecek dönem için kapasite ve kaynak planlaması konularında avantaj sağlayacaktır. Bu sayede gelecek dönem kullanım öngörüsü yapılarak sistemin kullandığı kaynak beklenen trafiğe göre optimize edilebilir ve kaynaklarının verimli kullanılmasına imkân sağlanabilir. Uygulama üzerinden reklam yapılıyorsa reklamların daha fazla müşteriye ulaştırılabilmesi için uygulamanın kullanım yoğunluğu eğilimine göre daha efektif reklam planlaması için destek olacaktır. Ek olarak beklenenin üzerinde ya da altında kullanım olması halinde normal olmayan(anormal-anomali) bir durumun yaşandığının tespit edilebilmesine (sistemsel sorun, atak, kampanya, özel gün vs.) katkı sağlanacaktır.

Bu çalışmada, zaman serileri ile Türkiye'nin ileri gelen operatörlerinden Turkcell İletişim Hizmetleri A.Ş ye ait aktif olarak müşterilere hizmet veren bir mobil uygulamanın belli bir zaman dilimindeki günlük kullanıcı giriş sayısını içeren veri seti kullanılarak bu uygulamanın zamana dayalı kullanım yoğunluğu belirlenmiş ve sonraki dönemler için günlük giriş sayısının tahmin edilmesi üzerine çalışma yapılmıştır.

Tez kapsamında literatürde zaman serileri kullanılarak gelecek dönem tahmini yapılan çalışmalar literatür araştırması başlığı altında incelenmiştir. Materyal ve yöntem başlığı altında kullanılan veri setiyle ilgili detay bilgi verilmiş ve tahmin için kullanılan algoritma ve yöntemler incelenmiştir. Ek olarak tahmin performans ölçüm yöntemleriyle ilgili incelemeler yapılarak matematiksel ifadelerini yer verilmiştir.

Veri setinin eğitim ve test seti olarak ayrıştırılma modellerine göre her bir algoritma da giriş değerleri ve parametreler değiştirilerek yapılan deneyler ve deneylerin sonuçları ve tahmin performansları Deneyler ve Bulgular bağlığı altında el alınmıştır. Sonuçlar kısmında çalışmadan elde edilen çıktılar ve sonraki aşamalarda neler yapılabileceği konusundaki değerlendirmeler eklenmiştir.

1.1.Literatür Araştırması

Literatürde önemli yer tutan Yöktez, Ebsco, GoogleScholar, IEEE, ScienceDirect, Ulakbim gibi veri tabanlarında "time series, forecasting, time series analysis, time series

forecasting, mobile application forecasting" gibi anahtar kelimeler ile 2000 yılından sonra yapılan araştırmalar filtrelenerek incelendiğinde zaman serileri kullanılarak farklı algoritma ve yöntemlerle tahminleme yapılan birçok çalışma bulunmaktadır. Ancak telekomünikasyon sektöründe faaliyet gösteren ve aktif olarak müşterilerine hizmet veren mobil uygulamaların günlük bazda kullanımının tahmini üzerine pek fazla çalışma bulunamamaktadır. Literatürde zaman serileri analizi ve gelecek dönem tahmini için bazı örnek çalışmalar alttaki gibidir.

Tezuran'ın [2] çalışmasında zaman serileri kullanılarak çağrı merkezine gelen çağrı sayısı analiz edilmiş ve sonraki dönemler için tahmin çalışması yapılmıştır. Veri seti olarak Türk Bankalar Birliğinin sunmuş olduğu 2008 – 2018 yılları arasında 3 er aylık periyotlardaki çağrı merkezi aranma sayıları kullanılmıştır. Çalışmada çağrı sayısı tahmini için ayrıştırma modeli, üstel düzleştirme, ARIMA modeli ve YSA yöntemleri uygulanmıştır. 2019 yılı ilk çeyreği için yapılan tahminlerde en yüksek tahmin %98 doğruluk oranıyla YSA modelinde sağlanmıştır. Uygulanan diğer modellerin performansları başarıları sıralamasına göre sırasıyla ayrıştırma modeli %97, ARIMA modeli %95, ETS modeli %93 ve regresyon modeli %92 olarak gözlemlenmiştir.

Gültekin'in [3] çalışmasında bankacılık sektöründe de para çekme makinelerine (ATM) para yükleme ziyaretlerinin planlaması ve yüklenecek bakiyenin öngörülmesi üzerine inceleme yapılmıştır. ATM'ye para transferi için yapılan ziyaretlerin sıklığı ve ATM de ihtiyaçtan fazla para tutulması banka maliyetini artırırken, paranın yetersiz olması da operasyonun tıkanmasına ve müşteri memnuniyetsizliğine neden olmaktadır. Bu nedenle ATM de tutulacak paranın da geçmiş kullanımlara göre tahminlenmesi üzerine çalışmalar yapılmıştır. Türkiye'nin en büyük bankalarından birine ait 50 farklı ATM'nin 13 ay için günlük işlemlerini içeren veri seti kullanılmıştır. Bir yıl eğitim seti, 1 ay ise test veri seti olarak ayrılmıştır. Çalışmada MLR, YSA, ARIMA, GPR, DVM, LSTM kullanılmıştır. Performans ölçümü için de MAD tercih edilmiştir. Yapılan tahminlerde DVM'den (Destek Vektör Makineleri) uyarlanan DVR (Destek Vektör Regresyonu) ile 15 ATM için doğruluk en yüksek çıkarken ikinci sırada 11 ATM ile Yapay Sinir Ağları yöntemleri kullanılarak başarılı sonuçlar elde edildiği belirtilmiştir.

Singh ve arkadaşları [4] bulut sistemlerde çalışan web uygulamalarının iş yükü tahmini üzerine bir uygulama yapmıştır. Bulutta çalışan uygulamaların maliyetleri kullandığı

kaynaklara değişkenlik gösterir. Eğer iş yükü yoğun değilken fazla kaynak rezerve edilirse kaynaklar verimli kullanılmaz ve maliyetin artması anlamına gelebilir. Kaynakların etkili kullanılması ve maliyetlerin düşürülmesi için zaman serileri kullanılarak iş yükü tahmin çalışması yapılmıştır. Veri seti olarak ClarkNet ve NASA ya ait webserver loglarını ARIMA, SVR, LR, TASM yöntemleri ile analiz etmiştir. Çalışma çıktıları MAE, MSE, RMSE, MAPE yöntemlerini ile hata oranları hesaplanmıştır. Çalışma sonucunda TASM %10,7 hata ile, ARIMA ve NAIVE de %12.5 civarında hata oranı ile iş yükü tahmini gerçekleştirilmiştir.

Karakoyun'un [5] çalışmasında zaman serileri kullanılarak hisse senedi ve Bitcoin fiyat tahminlemesi yapılmıştır. İstatiksel yaklaşım için ARIMA, makine öğrenmesi için Uzun-Kısa Süreli Hafıza derin öğrenme yöntemi kullanılarak tahminleme performansları karşılaştırılmıştır. Çalışmada veri seti olarak hisse senedi verileri için Alpha Vantage dan 7 Eylül 1984 ile 14 Şubat 2018 tarihleri arasındaki 8428 adet günlük hisse kapanış fiyatı kullanılmıştır. Bitcoin ile ilgili veriler için ise Quandl veri kaynağından 28 Nisan 2013 ve 29 Ekim 2017 arasında günlük olarak alınan Bitcoin fiyatından oluşmaktadır. Çalışmada eğitim setinden sonraki 30 gün için tahmin çalışması yapılmaktadır. Performans değerlendirmesi için RMSE, MAE, MPE, MAPE kullanılmıştır. Bitcoin ve hisse senedi tahmin çalışmasında MAPE hesabına göre sırasıyla %1,4 ve %5,78 hata payıyla Uzun-Kısa Süreli Hafıza derin öğrenme mimarisinin daha başarılı olduğu gözlemlenmiştir.

Kaynar ve Taştan [6], çalışmalarında ARIMA ve Yapay Sinir Ağları yöntemlerini kullanarak gelecek dönem TL/Dolar kuru tahmin uygulaması yapmış ve sonuçlarını karşılaştırmıştır. Çalışmalarında, T.C. Merkez Bankasının elektronik veri dağıtım sisteminden aldıkları Ocak 2000 ile Haziran 2008 arasındaki günlük ve aylık döviz kuru (YTL/$) değerlerini kullanmışlardır. Performans ölçümü için MSE ve MAPE kullanılmıştır. Yapılan tahminlerde ARIMA ve MLP modellerinin MSE ve MAPE değerleri sırasıyla (0.002-2.66), (0.002-3.27), günlük veriler için ise (0.0001 - 0.67), (0.0001 - 0.62) seklinde birbirine yakın sonuçlar verildiği görülmüştür.

Çelik'in [7] çalışmasında Türkiye İstatistik Kurumu'nun sitesinde yer alan "İstatistik Göstergeler 1923-2011" ve "Karayolu Trafik Kaza İstatistikleri" kaynaklarından faydalanılarak 1955-2012 yılları arasında Türkiye'de meydana gelen trafik kazaları için

bir zaman serisi oluşturulmuş ve ARIMA yöntemi ile analiz edilmiştir. Yapılan çalışmalarda ARIMA (0, 2, 3) yönteminin ileriye yönelik en uygun kaza tahmini yapılabileceği belirlenmiştir.

Atalay'ın [8] çalışmasında makine öğrenme yöntemleri kullanılarak tüketim istatistiklerine göre talep tahmin uygulaması yapılmıştır. Elektrik enerjisi depolanabilir olmadığından ihtiyaca göre üretim sağlanması gerekmektedir. Fazla enerji üretimi maliyetin artmasına neden olurken az üretim ise kesintiye sebep olabilir. Bu durumların önüne geçebilmek için zaman serilerinden faydalanılarak tahminleme yapılmaktadır. Çalışmada beş bin ticari işletmeye ait elektrik talebi ile Doğu Slovakya'ya ait elektrik taleplerini içeren iki veri kümesi kullanılmıştır. Elektrik enerjisinin kullanım tahmini için yapılan çalışmada Box-Jenkins, kaos zaman serileri, YSA, DVM, bulanık mantık gibi çeşitli yöntemler kullanılmıştır. Hata kriteri olarak da MAPE kullanılmıştır. Uygulanan veri setleri için en başarılı sonucun Destek Vektör Makineleri yöntemiyle alındığı belirtilmiştir.

Salaudeen'in [9] çalışmasında Nijerya, Kogi eyaletinin Lokoja ilinde bulunan acil servise gelen gelen dönemsel hasta sayısı için zaman serisi oluşturulmuş ve zaman serisi analizinde otoregresif entegre hareketli ortalama modeli kullanılmıştır. ARIMA modelinin parametre belirlenmesinde 2012-2016 yılları arasında acil servise yapılan başvurular eğitim seti olarak kullanılmış, 2017 yılının ilk yedi aylık dönem ise test ve karşılaştırma için kullanılmıştır. Yapılan tahminin performansı MAPE ve RMSD ile ölçümlenmiş ve %5-10 arası hata payına sahip olduğu tespit edilmiştir.

2. MATERYAL VE YÖNTEM

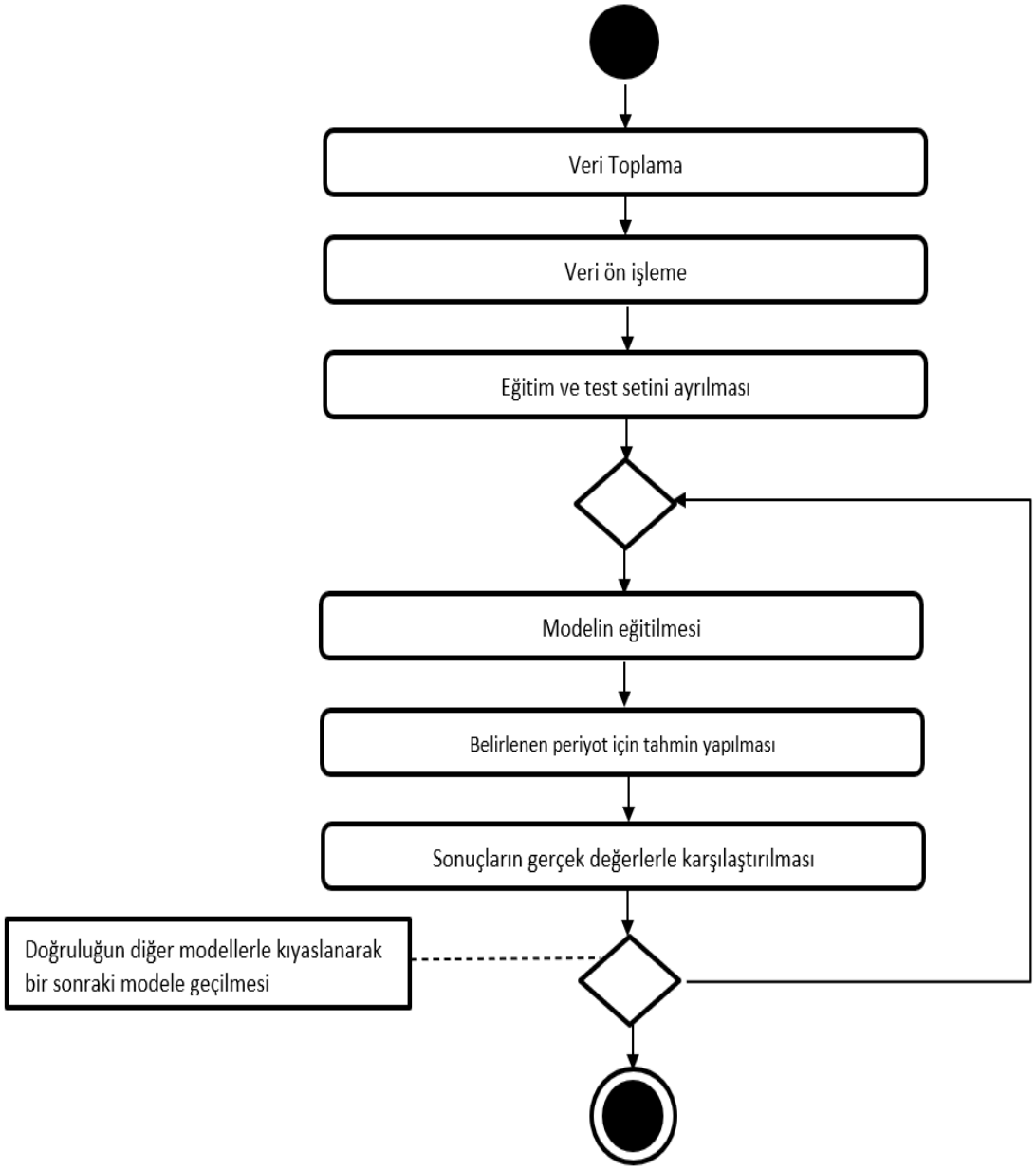
Bu bölümde zaman serisi, zaman serisi analizi, zaman serisi uygulamaları, zaman serisi bileşenleri ve tahminlemede kullanılan modeller üzerinde durulmuştur. Zaman serisi tahminleme modellerinden regresyon, üstel düzleştirme, ARIMA, destek vektör makineleri ve yapay sinir ağları modelleri üzerinde durulmuş ve bu modeller matematiksel olarak incelenmiştir.

Kullanılan veri seti Turkcell İletişim Hizmetleri A.Ş ye ait olan bir mobil uygulamanın günlük periyotta uygulama giriş sayılarını içermektedir.

Veri seti elde edilirken uygulamanın kullandığı veri tabanından 2018 yılına ait log-in sayıları çekilmiş ve günlük bazda gruplanmıştır. Ardından kronolojik olarak sıralanarak zaman serisine çevrilmiş, gün bazında kullanım örüntüsü/eğilimi elde edilmiştir.

Veri seti eğitim ve test seti olarak iki farklı şekilde ayrılarak her bir algoritmada birçok farklı deney yapılmıştır. İlk olarak bir yıl (2018 yılının tamamı) eğitim seti olarak kullanılmış ve 2019 Ocak ayı için (31 gün) tahmin çalışması yapılmıştır ve doğruluk performans değerleri ölçülmüştür. İkinci olarak da ilk on bir ay (2018 Ocak – 2018 Kasım) eğitim seti olarak kullanılmış, sonraki bir ay (Aralık) için tahmin çalışması yapılmıştır. Her iki veri seti modelinde tüm algoritmalar için giriş değerleri ve kullanılan parametreler değiştirilerek optimum doğruluk performansının bulunması için deneyler yapılmıştır.

Gelecek dönem kullanım eğrisinin tahmin edilebilmesi için temel istatistiksel yöntemler ve makine öğrenmesi yöntemleri uygulanarak sonuçlar karşılaştırılmıştır. Kullanılan yöntemlerle ilgili bilgiler altta detaylıca gösterilmiştir. Çalışmanın akış diyagramı da Şekil 2.1.'de gösterilmiştir.



Şekil 2. 1. Akış Diyagramı

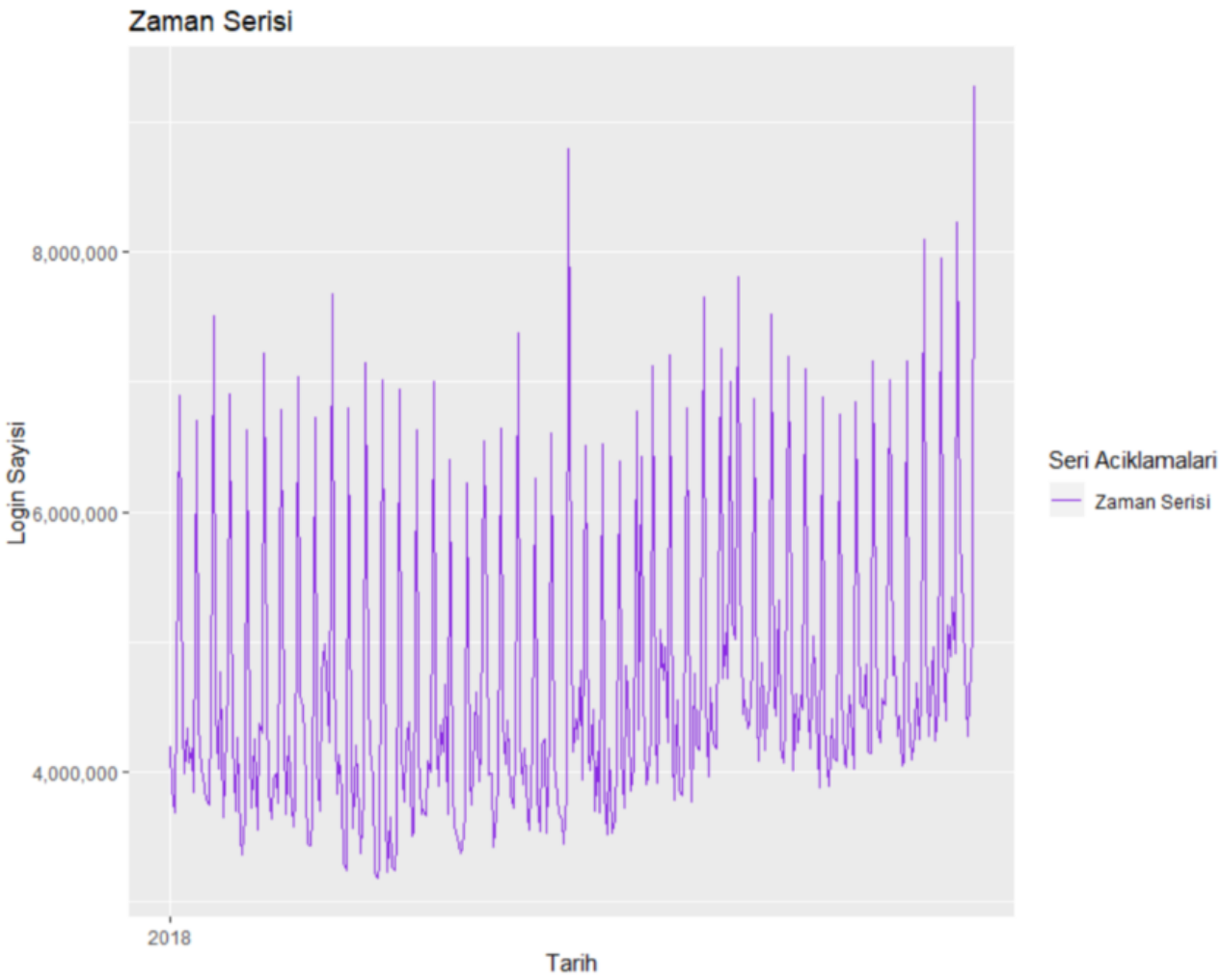
2.1. Zaman Serileri ve Zaman Serileri Analizi

Belirli bir zaman aralığındaki tutarlı periyotlardan oluşan istatistiksel verilerin ve gözlemlerin kronolojik sıraya göre (zamana göre) düzenlenmesine zaman serisi adı verilir [10,11]. Zaman serilerini zaman serisi analizi ve zaman serisi tahminleme (ön görme) olarak iki ayrı başlık altında inceleyebiliriz. Zaman serisinin analiz edilerek

anlamlı istatistiksel ve karakteristik özelliklerinin elde edilmesine zaman serisi analizi, bu serinin özellikleri ve geçmiş değerlerin kullanılarak sonraki dönemlerle ilgili tahminde bulunulması ise zaman serisi tahminleme olarak ifade edilebilir [12,13].

Zaman serileri verinin zaman içindeki değişimi ve zamana göre karakteristiği ile birlikte belli bir kalıbı (örüntüsü) varsa gösterilmesine olanak sağlar. Bu örüntünün çok dışında kalan değerler varsa yani beklenmeyen (anormal) bir değer olması halinde fark edilmesine olanak tanır ve karar vermeyi desteklemek için geleceğe dair eğilimle ilgili öngörüde bulunmamızı sağlar. Zaman serileri tek değişkenli ve çok değişkenli zaman serileri olmak üzere ikiye ayrılır. Zaman serisinde tek bir değişkenin sıralı kayıtları varsa tek değişkenli zaman serisi, birden fazla değişken verisi mevcutsa da çok değişkenli zaman serisi olarak adlandırılmaktadır.

Zaman serileri finans, istatistik, tıp, meteoroloji, tarım, mühendislik, enerji, bankacılık gibi birçok alanda geçmişe ait zamana dayalı verilerin analizinde kullanılmaktadır. Geçmiş dönem verileri kullanılarak gelecek dönem için tahminler yapılmasına ve bu tahminler doğrultusunda planlamalar ve gerekliyse önlemler alınmasına olanak sağlamaktadır [14]. Özellikle üretim planlaması, stok ve depolama planlaması, istihdam ve izin planlaması, dönemsel gelir/giderin artış/azalış eğilimlerinin takibi gibi çok önemli faydalar sağlanabilmektedir. Şekil 2.2.'de çalışmamızda kullandığımız verilerinden oluşturulan bir zaman serisi grafiğine yer verilmiştir.



Şekil 2. 2. Örnek zaman serisi grafiği

2.2.Doğrusal Regresyon

Regresyon analizi, iki veya daha fazla değişken arasındaki ilişkinin belirlenmesi ve bu ilişkiyi kullanarak o konu ile ilgili tahminler yapılması amacıyla kullanılan bir yöntemdir. Tek değişken (bağımsız) kullanılarak analiz yapılıyorsa tek değişkenli regresyon, en az iki değişken (bağımsız) kullanılarak analiz yapılıyorsa çok değişkenli regresyon analizi olarak adlandırılır. Buradaki temel amaç bilinen bağımsız değişkenlerin değerleri için bağımlı değişken değerinin tahmin edilmesidir.

2.2.1. Basit Doğrusal Regresyon

Bilinen bir değişken(bağımsız) ile ilişkili başka değişkeninin tahmin edilmesi için uygulanan regresyon modeline basit doğrusal regresyon (simple linear regression) adı verilir. Basit lineer regresyonun matematiksel formülü Denklem 2.1'de verilmiştir.

$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$ (2.1)

y_t: bağımlı değişken

β_0: kesme noktası

β_1: regresyon katsayısı(eğim)

x_t: bağımsız değişken

ε_t: hata terimi

2.2.2. Çoklu Doğrusal Regresyon

Birden fazla bağımsız değişkenden faydalanarak bir değişkenin tahmin edilmesi için kullanılması ise çoklu doğrusal regresyon (multiple linear regression) olarak tanımlanır. Birbiriyle ilişkili bağımsız değişkenler ile bağımlı değişken arasındaki ilişkiyi Denklem 2.2'de gösterildiği gibi ifade edilmektedir.

$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \cdots + \beta_k x_{k,t} + \varepsilon_t$ (2.2)

y_t: bağımlı değişken

β_0: kesme noktası

$\beta_1, \beta_2, \beta_k$: regresyon katsayısı(eğim)

x_1, x_2, x_k: bağımsız değişken

ε_t: hata terimi

2.3.BOX-Jenkins Yöntemi

ARIMA (Auto Regressive Integrated Moving Average) zaman serisi kullanılarak yapılan tahminlerde ek nitelik ve bilgi gerektirmemesinin yanında kısa ve orta dönem tahmin başarımı yüksek olması gibi nedenlerle çok yaygın kullanılan istatistiksel modellerden biridir [15]. Otoregresif (AR – Auto Regressive) ve hareketli ortalama (MA – Moving Average) yöntemlerinin birleşiminde oluşmaktadır. İsmini de bu yöntemlerin baş harflerinden almaktadır.

AR: Auto Regressive (Otoregresif)

I: Integrated (Bütünleşik)

MA: Moving average (Hareketli ortalama)

ARIMA modelleri temelde verilerdeki otokorelasyonları (autocorrelation) tanımlamayı amaçlamaktadır, George Box ve Gwilym Jenkins tarafından geliştirilmiş bu yönteme Box Jenkins (BJ) yöntemi de denilmektedir [12]. ARIMA, zaman serisinin gelecekteki değerlerinin mevcut ve geçmiş değerlere doğrusal olarak bağlı olduğu konusunda güçlü bir varsayım yapmaktadır [16]. ARIMA modeli, kısa dönem tahminlerde etkin bir performansa sahiptir [17,18].

2.3.1. Otoregresif Modeli (AutoRegressive)

Otoregresif model, bir zaman serisinin herhangi bir periyodundaki değerinin, bu değerden önce gelen belirli sayıdaki gözlem değerinin ve hata teriminin doğrusal regresyonu olarak ifade edildiği modeldir.

AR modeli içerdiği gözlem değerine göre adlandırılır. Yani geçmiş döneme ait bir adet gözlem değeri içeriyorsa birinci dereceden, iki adet gözlem değeri içeriyorsa ikinci dereceden ve genel olarak p adet geçmiş dönem gözlem değeri içeriyorsa p'inci dereceden AR modeli olarak adlandırılır.

AR(p) matematiksel ifadesi Denklem 2.3'de gösterilmiştir.

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \quad (2.3)$$

$y_{t-1}, y_{t-2}, \ldots, y_{t-p}$: gözlem değerleri

$\phi_1, \phi_2, \ldots, \phi_p$: model parametreleri

P : model derecesi

ε_t: hata terimi

2.3.2. Hareketli Ortalama Modeli (Moving Average)

Hareketli ortalama (MA) modeli, zaman serisinin herhangi bir dönemindeki değerinin, aynı döneminin hata terimi ve geçmiş dönemlerinin hata terimlerinin doğrusal bir fonksiyonu olarak ifade edilebilir. Hareketli ortalama teriminin kullanılmasının sebebi ise seriye eklenen her yeni değer ile birlikte yeni bir ortalamanın hesaplanabilmesi ve bunun öngörü olarak kullanılabilecek olmasıdır [19]. Hareketli ortalama modelleri de içerdikleri geçmiş dönem hata terimi sayısına göre isimlendirilirler. Birinci dereceden MA(1), ikinci dereceden MA(2) ve genel ifadeyle q'inci dereceden MA(q) modelleri olarak adlandırılırlar. Formülünü Denklem 2.4'de görmek mümkündür.

$y_t = \varepsilon t + \Theta_1 \varepsilon_{t-1} + \Theta_2 \varepsilon_{t-2} + \cdots + \Theta_p \varepsilon_{t-q}$ (2.4)

y_t t dönemine ait gözlem değeri

$\Theta_1, \Theta_2, \ldots, \Theta_p$ modelin parametreleri

q modelin derecesi MA(q)

2.3.3. Otoregresif Hareketli Ortalama Modeli (AutoRegressive Moving Average)

Otoregresif hareketli ortalama modeli (ARMA) genellikle durağan zaman serilerinin modellenmesi için kullanılır. Adından da anlaşılacağı gibi Otoregresif (AR) model ve Hareketli Ortalama (MA) modellerinin birleşiminden meydana gelir.

p terim içeren AR ve q terim içeren MA modelinin gösterimi ARMA(p,q) şeklindedir [20]. Formülü Denklem 2.5'de olduğu gibi ifade edilebilir.

$y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \cdots - \phi_p y_{t-p} = \varepsilon_t + \Theta_1 \varepsilon_{t-1} + \Theta_2 \varepsilon_{t-2} + \cdots + \Theta_q \varepsilon_{t-q}$ (2.5)

2.3.4. ARIMA (AutoRegressive Integrated Moving Average)

Mevsimsel ve konjonktürel dalgalanmalar, eğilim ve ön görülemeyen nedenlerle zaman serilerinin durağanlığı bozulabilir. Zaman serilerinin modellenmesi için seri durağanlığının sağlanması önemlidir. Durağanlığın sağlanması için öncelikle etkileyen faktörlerin belirlenmesi ve yok edilmesi gerekir. Eğer seri durağan değilse serinin uygun derecede farkı alınarak durağanlık sağlanabilir. Fark alma derecesi d ile ifade edilir. Otoregresyon parametresinin derecesi p ile ifade ediliyorken, hareketli ortalama parametresi derecesi de q ile ifade ediliyorsa ve d dereceden fark alınmışsa bu modele (p,d,q) dereceden otoregresif entegre hareketli ortalama modeli adı verilir ve ARIMA (p,d,q) şeklinde gösterilir ve Denklem 2.6'da gösterildiği gibi formülize edilir.

$y_{t'} = \phi_1 y_{t-1'} + \cdots + \phi_p y_{t-p'} + \Theta_1 \varepsilon_{t-1} + \cdots + \Theta_q \varepsilon_{t-q} + \varepsilon_t$ (2.6)

Tablo 2. 1. ARIMA model tablosu

Model	ARIMA Modeli
Gürültü	ARIMA(0, 0, 0)
Rassal Yürüyüş Modeli	ARIMA(0, 1, 0) sabit değer olmadan
Sürüklenme ile Rassal Yürüyüş Modeli	ARIMA(0, 1, 0) sabit değer ile
Otoregresif Model	ARIMA(p, 0, 0)
Hareketli Ortalama Modeli	ARIMA(0, 0, q)

2.4.Naive Model

Veri setindeki son gözlem değerinin bir sonraki dönem için en iyi öngörü değeri olarak kullanılabileceği varsayımını yapan modellerdir. Hisse senedi tahmin uygulamalarında sıklıkla kullanılır. En sade hali ile Denklem 2.7'de olduğu gibi gösterilebilir.

$$Y_{t+1} = Y_t \ldots\ldots \quad (2.7)$$

Y_{t+1} = Tahmin Değeri
Y_t = Son gözlem değeri

İncelediğimiz veri setindeki değerler zamana göre artış ya da azalış gösteriyorsa durağan olmayan (non stationary) seriler denir. Bu serilerde serinin ortalaması zaman içinde değişmektedir veya bir eğilimi vardır. Eğilim faktörünün de dikkate alınarak oluşturulmuş bir başka naive model de alttaki gibi ifade edilebilir. Bu modelde bir dönem öncesinin değerine önceki iki dönem arasındaki farkın yönüne göre eklenerek ya da çıkarılarak ardışık dönemler arasındaki değişim hesaba dahil edilmiştir. Formülünü Denklem 2.8'de görmek mümkündür.

$$Y_{t+1} = Y_t + (Y_t - Y_{t-1}) \ldots\ldots \quad (2.8)$$

Y_{t+1} = t+1 dönemindeki tahmin Değeri
Y_t = t dönemindeki tahmin değeri
Y_{t-1} = t dönemindeki gözlem değeri

Mevsimsellik söz konusu ise alttaki gibi ifade edilebilir.
$Y_{t-1} = Y_{t-4}$

Seri hem eğilim özellliği hem de mevsimsellik gösteriyorsa iki yaklaşım birden kullanılarak Denklem 2.9'da ifade edilen model oluşturulabilir [21].

$$Y_{t-1} = Y_{t-3} + \frac{(Y_t - Y_{t-1}) + \cdots + (Y_{t-3} - Y_{t-4})}{4} \quad (2.9)$$

2.5. Üstel Düzleştirme Yöntemleri

Hareketli ortalamalar yöntemine benzeyen ancak hareketli ortalamalardan farklı olarak geçmiş döneme ait değerlerin ağırlıklarının farklılaştırıldığı yöntemlere üstel düzleştirme yöntemleri adı verilir [22]. Farklı bir ifadeyle tahminde kullanılacak geçmiş dönem verilerinden en yüksek ağırlığa sahip olan değer en yeni değer iken değerin tarihi eskidikçe üstel olarak azalan ağırlıklar verilmektedir [23]. Hareketli ortalama modeline

göre en belirgin üstünlüğü bu özelliğidir. Bu nedenle de en yakın geçmiş verilerin geleceğe etkisi, eski dönem verilerinden daha fazla olmaktadır. Basit bir şekilde uygulanabilmesi ve düşük maliyetli olması avantajına sahiptir. Bir diğer avantajı ise gerçekleşen her bir yeni dönem verisi modele dahil edilir ve yeni dönem tahminlerine katkı sağlar [24].

2.5.1. Basit Üstel Düzleştirme Yöntemi

Belli bir eğilimi olmayan ve mevsimsellik barındırmayan zaman serilerinde gelecek zaman tahmini için basit üstel düzleştirme yöntemi kullanılabilir. Bu yöntemde yakın geçmişteki gözlem değerlerine daha yüksek ağırlık verilirken daha eski gözlem değerlerine düşük ağırlıklar verilerek hesaplamalar yapılmaktadır. Denklem 2.10'da gösterilen formülle ifade edilebilir.

$Y'_{t-1} = \alpha \cdot Y_{t-1} + (1 - \alpha) \cdot Y'_{t-1}$ (2.10)

Y'_t = t anındaki tahmin değer

Y_{t-1} = t-1 anındaki gözlem değeri

Y'_{t-1} = t-1 anındaki tahmin değeri

α = düzleştirme sabiti ($0 < \alpha < 1$)

2.5.2. Brown'un Tek Parametreli Doğrusal Üstel Düzleştirme Yöntemi

Zaman serisinin bir eğilime sahip olduğu durumlarda basit üstel düzleştirme yöntemi yerine Brown'un tek parametreli doğrusal üstel düzleştirme yöntemi daha uyumlu sonuçlar sağlamaktadır. Denklem 2.11'de gösterilmiştir.

$Y_t^1 = \alpha \cdot Y_{t-1} + (1 - \alpha) \cdot Y_{t-1}^1$ (2.11)

$Y_t^2 = \alpha \cdot Y_t + (1 - \alpha) \cdot Y_{t-1}^2$

Y_t^1: tekli üstel düzleştirme ile elde edilen değer,

Y_t^2: ikili üstel düzleştirilmiş değer

$a_t = Y_t^1 + (Y_t^1 - Y_t^2) = 2Y_t^1 - Y_t^2$

$bt = \frac{\alpha}{1-\alpha} (Y_t^1 - Y_t^2)$

sonraki m dönemin tahmini için de alttaki gibi bir model kurulabilir.

$Y'_{t+m} = a_t + b_t m$

2.5.3. Holt'un Çift Parametreli Doğrusal Üstel Düzleştirme Yöntemi

Doğrusal bir eğilimi olan zaman serilerinde tahmin için önerilen diğer bir yöntem ise Holt'un çift parametreli doğrusal üstel düzleştirme yöntemidir. Bu yöntemde eğilim değerlerine düzleştirme uygulanmaktadır. Tahmin için Denklem 2.12 de ifade edilen formülden faydalanılır.

$$Y'_{t-1} = \alpha \cdot Y_t + (1 - \alpha) \cdot (Y'_{t-1} + b_{t-1}) \quad (2.12)$$

$$b_t = \beta\,(Y'_t - Y'_{t-1}) + (1 - \beta)\, b_{t-1}$$

$$\hat{y}_{t+m} = Y'_t + b_t m$$

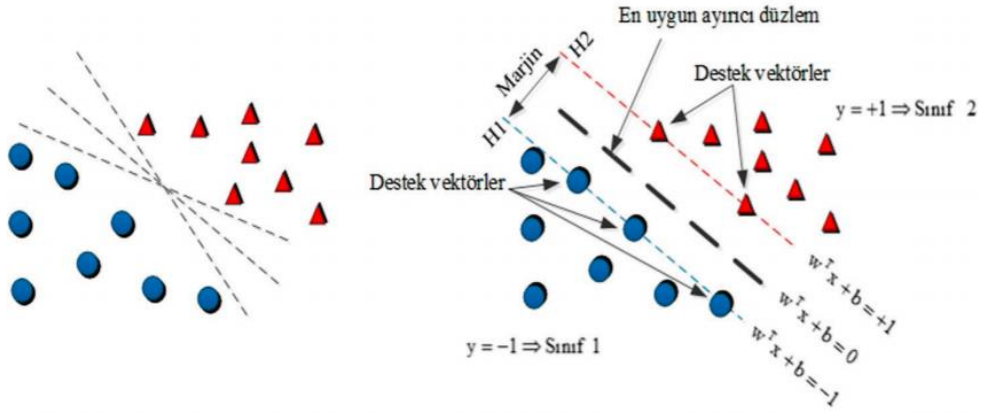
2.6. Destek Vektör Makineleri

Destek vektör makineleri Vladimir Vapnik ve Alexey Chervonenkis tarafından geliştirilmiş, temel olarak istatistiksel öğrenme teorisine dayanan bir makine öğrenmesi yöntemidir [25]. DVM'nin çalışma prensibi iki sınıfı en uygun şekilde ayırabilen hiper-düzlemin tanımlanması esasına dayanır [25,26]. DVM yapısal risk minimizasyonu ilkesini ile çalıştığından aşırı uyum (overfitting) durumunun engellenmesini sağlamakta ve yüksek genelleştirme performansı yakalanmaktadır [27, 28].

Ek olarak sınıflandırma probleminin kareli optimizasyon problemine dönüştürülerek çözülmesi öğrenme aşamasında işlem sayısı azalmakta ve diğer yöntem-algoritmalara göre daha hızlı çözüme ulaşılmaktadır [29]. Bu nedenle büyük hacimli veri setlerinde de büyük avantaj sağlamaktadır. Ayrıca optimizasyon temelli olduğundan sınıflandırma performansı, hesaplama karmaşıklığı ve kullanışlılık açısından diğer tekniklere göre daha başarılı olduğu belirtilmektedir [30, 31].

2.6.1. Doğrusal Destek Vektör Makineleri

Veri setini doğrusal olarak sınıflandırabilen hiper düzlem olarak ifade edilebilir. Sonsuz tane hiper düzlem çizebilme imkanı mevcut olmasına rağmen önemli olan birbirine en yakın noktalar arasındaki marjini maksimum yapan hiper düzlemin belirlenmesidir.



Şekil 2. 3. İki sınıfın doğrusal ayrımı

Doğrusal bir DVM de karar sınırının denklemi Denklem 2.13'de gösterildiği şekilde ifade edilebilir.

$w \cdot x + b = 0$ (2.13)

Destek Vektörlerin(H1 ve H2) formülleri de alttaki gibi verilmiştir.

$H_1 : w * x + b = 1$

$H_2 : w * x + b = 1$

w: çoklu düzlemin normali(ağırlık vektörü)

x: veri noktası

b: bias

H_1, H_2: Destek vektörleri

İki hiper düzlem arasındaki mesafe marjin olarak adlandırılır. Alttaki gibi ifade edilir.

$w \cdot (x_1 - x_2) = 2$

$d = 2 / |w|$

DVM nin temel amacı marjinin (d) maksimize edilmesidir. Bunun formülü de alttaki gibi elde edilmektedir.

$w \cdot x_i + b \geq 1$ durumunda $y_i = 1$

$w \cdot x_i + b \leq -1$ durumunda $y_i = -1$

Bu optimizasyon problemi belirlenen kısıtlar altında maksimum ya da minimum değeri vermelidir, bu da Lagrange denklemleri kullanılarak çözülebilir

$L_p = \frac{1}{2} \|w\|^2 - \sum_{i=1}^{n} a_i (N_i = 1 (y_i(wt \cdot x_i + b) - 1)$

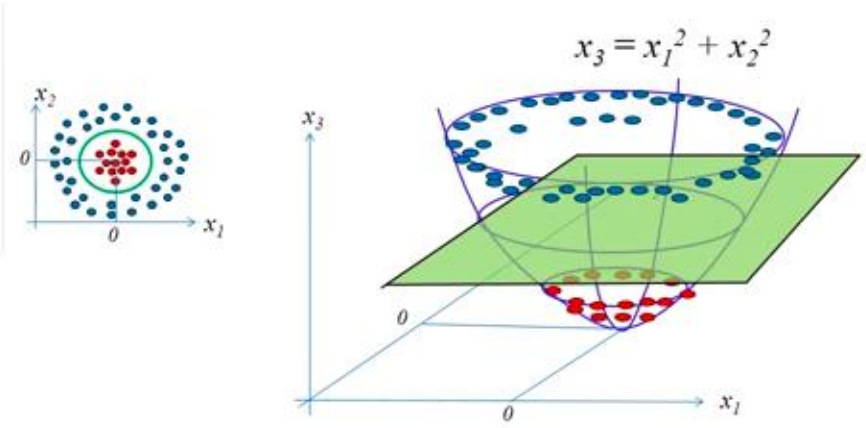
$\frac{\partial L_p}{\partial w} = 0 \Rightarrow W = \sum_{i=1}^{n} a_i y_i x_i$

$\frac{\partial L_p}{\partial b} = 0 = \sum_{i=1}^{n} a_i y_i$

$$LD = \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i=1}^{n} a_i a_j y_i y_j x_i x_j \quad (2.13)$$

2.6.2. Doğrusal Olmayan Destek Vektör Makineleri

Veri kümesinin doğrusal olmaması halinde DVM'ler ile doğrusal bir hiper-düzlem çizilemez. Bu durumlarda çekirdek fonksiyonları (kernel trick) kullanılır. Çekirdek fonksiyonu kullanımı doğrusal olmayan verilerde makine öğrenimini yüksek oranda arttırmaktadır.



Şekil 2. 4. Doğrusal ayrılamayan iki sınıfın çekirdek fonksiyonu kullanılarak ayrımı [32]

Literatürde bir çok farklı çekirdek fonksiyonu bulunmaktadır. Bunların başlıcaları Tablo 2.2.'de matematiksel formülleriyle birlikte gösterilmiştir.

Tablo 2. 2. Temel SVM çekirdek fonksiyonları

Çekirdek Fonksiyonu	Matematiksel Gösterim	Çekirdek Parametresi
Doğrusal	$k(x_i, x_j) = (x_i \cdot x_j)$	
Polinom	$k(x_i, x_j) = (x_i \cdot x_j + 1)^d$	d: polinom derecesi
Gauss(RBF)	$k(x_i, x_j) = exp(-\frac{1}{2\sigma} \lVert x_i - x_j \rVert^2)$	σ: çekirdek ölçeği
Sigmoid	$k(x_i, x_j) = \tanh(\sigma\, x_i \cdot x_j - \delta)$	σ: çekirdek ölçeği δ: eşik değer parametresi

Doğrusal Çekirdek Fonksiyonu: Doğrusal SVM ler ile kullanılır. Herhangi bir parametresi ve özelliği bulunmaz. Doğrusal olarak hesaplama yapılır.

Polinom Çekirdek Fonksiyonu: Polinomun derecesine göre adlandırılır ve polinom derecesinin pozitif bir tam sayı olması gerekir. Derecesi bir olduğunda doğrusal çekirdek fonksiyonu ile aynı olur. Derecesi iki olduğunda ikinci derece çekirdek olarak

adlandırılır ve hiper düzlem kuadratik bir yüzeydir. Derecesi 3 olduğunda kübik olarak adlandırılır.

Gauss (RBF) Çekirdek Fonksiyonu: En yaygın kullanılan yöntemlerden biridir. Radial basis olarak da adlandırılır. Çekirdek ölçeği adında bir parametresi bulunmaktadır.

Sigmoid Çekirdek Fonksiyonu: İki çekirdek parametresi vardır. Biri pozitif bir sayı olması gereken çekirdek ölçeği, diğer ise eşik değer parametresidir. İki katmanlı sinir ağı ile benzerlik göstermektedir.

2.7. Yapay Sinir Ağları (Artificial Neural Network)

Yapay sinir ağları, insan beyninin öğrenme sürecine benzetilerek geliştirilmiş, basit matematiksel modellere dayanan tahmin yöntemidir. Sinir hücrelerinden esinlenilerek tasarlanmış bir makine öğrenmesi yöntemi olarak da ifade edilebilir [33]. Yapay sinir ağlarının ilk modelleme çalışması 1943 yılında bir sinir hekimi tarafından ortaya çıkmıştır [34]. Son yıllarda tahmin problemlerimde araştırmacılar tarafından yoğunlukla tercih edilmektedir. Modelin öğrenme aşamasından eğitim algoritması tarafından otomatik belirlenmesi ve herhangi bir fonksiyonu yaklaşık olarak temsil edebilmesi iki önemli avantajıdır [35]. Ek olarak aktivasyon fonksiyonları sayesinde doğrusal olmayan ilişkilerin temsil edilmesini de sağlayabilmektedir.

YSA girdiler, ağırlıklar, toplama fonksiyonu, aktivasyon fonksiyonu ve çıkılar olmak üzere 5 bileşenden oluşur.

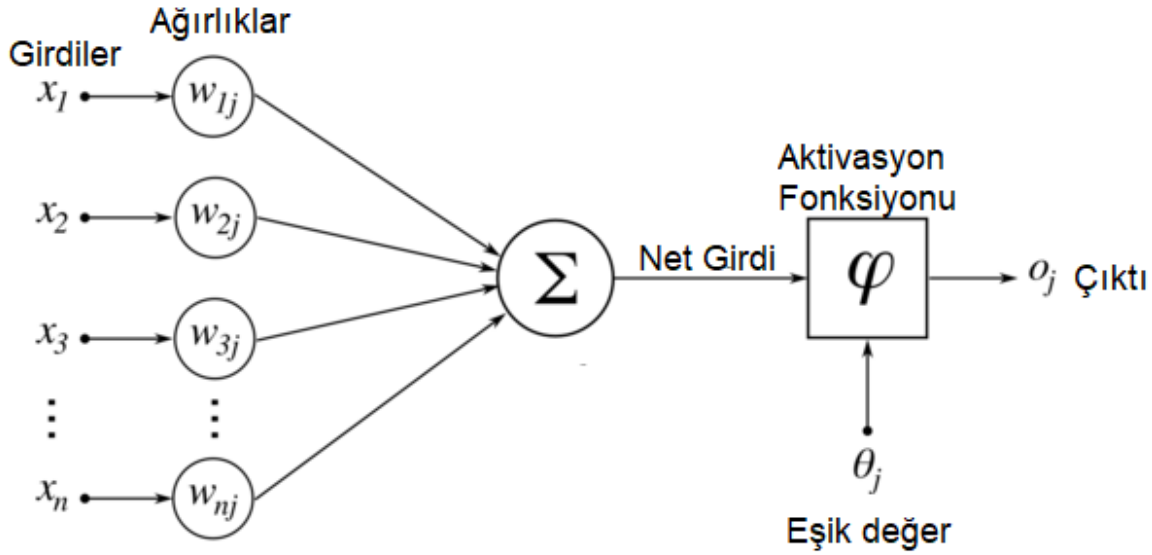
Girdiler, YSA ya dış dünyadan ya da başka hücrelerden gelen bilgilerdir.

Ağırlıklar, hücreye gelen bilginin önemini ve hücre üzerindeki etkisini gösterir.

Toplama Fonksiyonu, hücreye gelen net girdinin hesaplanması için kullanılır. Bu iş için çeşitli fonksiyonlar bulunmaktadır. En yaygın olanı ise ağırlıklı toplamıdır. Ağırlıklı toplama fonksiyonunda her gelen bilgi kendi ağırlığı ile çarpılarak toplanır ve ağa gelen net girdi bulunmuş olur.

Aktivasyon Fonksiyonu: Bu fonksiyon hücreye gelen net girdiyi işleyerek hücrenin bu girdiye karşılık üreteceği çıktıyı belirler. Şekil 2.5.'de en çok kullanılan aktivasyon

fonksiyonları listelenmiştir.



Şekil 2. 5. Örnek YSA Yapısı

YSA girdiler ile ağırlıklar çarpılarak net girdi elde edilir. Aktivasyon fonksiyonuna net girdi değeri geçirilerek, nöronun çıktısı elde edilir. Bu çıktı bir başka nöronun girdisi veya sinir ağının çıktısı olarak kullanılır [36]. Matematiksel formülü Denklem 2.14'de gösterilmiştir.

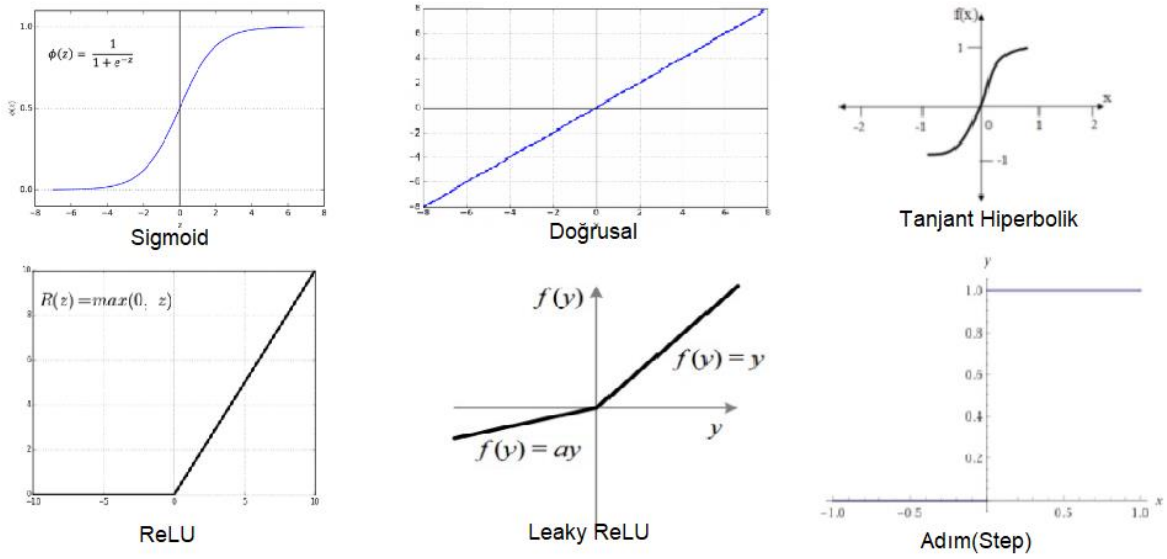
$y=\varphi(\Sigma x_i * \omega_{ij})$ (2.14)

φ: Aktivasyon fonksiyonu

x_i: girdi

ω: Ağırlık

Aktivasyon fonksiyonları



Şekil 2. 6. Yaygın kullanılan aktivasyon fonksiyonları

Sinir sistemi elemanlarının, YSA modeli terminolojisindeki karşılığı Tablo 2.3.'de gösterilmiştir.

Tablo 2. 3. Sinir sistemi - YSA terminolojisi

Sinir Sistemi	Yapay Sinir Ağı
Nöron	İşlem Elemanı
Dentrit	Toplama Fonksiyonu
Hücre Gövdesi	Aktivasyon Fonksiyonu
Akson	Eleman Çıkışı
Sinaps	Ağırlıklar

YSA modelini ile istatistiksel modellere ait kavramların karşılıkları da Tablo 2.4.'de gösterildiği gibi ifade edilebilir.

Tablo 2. 4. YSA modeli ile istatistiksel modellere ait kavramların karşılıkları

YSA Terminolojisi	İstatistik Terminolojisi
Yapay Sinir Ağı	Model
Ağırlık	Parametre
Girdi	Bağımsız değişken
Çıktı	Tahmin değeri
Hedef	Bağımlı değişken
Hata	Artık
Hata çizgisi	Güven aralığı

YSA eğer bir döngüye sahip değilse yani tek yönlü bilgi akışı mevcut ise ileri beslemeli

YSA olarak adlandırılırken, kendi için döngü varsa yani en az bir hücrenin çıkışı kendisine ya da diğer hücrelere giriş olarak veriliyorsa geri beslemeli YSA olarak adlandırılır.

Modelin en öne çıkan özelliği öğrenme sürecidir. Modele girilen bilgiyi işleyerek hafızaya alır ve bu verilerin birbirleriyle ilişkisini modelleyerek işlenen ve öğrenilen veriyi çıktı olarak sunar.

2.8. Tahmin Yöntemlerinde Performans Ölçümü

Zaman serileri kullanılarak yapılan tahminlerin doğruluğunu değerlendirilmesi için çeşitli hata hesaplama yöntemleri bulunmaktadır. Bunlardan en çok kullanılanlar alttaki başlıklar altında incelenmiştir. Çalışmamızda kullandığımız veri setinin pozitif değerler içeriyor olması ve günlük değerlerinin büyük sayılardan oluşması nedeniyle performans ölçümünde ortalama mutlak hata yüzdesi (MAPE) yöntemi tercih edilmiştir.

2.8.1. Ortalama Hata (Mean Error - ME)

Tahmin modelinin öngördüğü tahmin değerleri ile gerçek değerlerin arasındaki ortalama hatadır.

$$ME = \frac{1}{n}\sum_{i=1}^{n} Y_i - \hat{Y}_i \quad (2.15)$$

2.8.2. Ortalama Mutlak Hata (Mean Absolute Error – MAE)

İki sürekli (continuous) değişken arasındaki farkın ölçüsü ortalama mutlak hata olarak adlandırılır. MAE değeri her zaman pozitiftir, 0'dan ∞'a kadar değişebilir. MAE skoru ne kadar düşük değerlere sahip ise tahminleyicinin performansı o kadar yüksektir şeklinde yorumlanabilir.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |Y_i - \hat{Y}_i| \quad (2.16)$$

2.8.3. Ortalama Mutlak Hata Yüzdesi (Mean Absolute Percentage Error – MAPE)

Başlıca hata ölçme yöntemlerinden biridir. Ortalama Mutlak Yüzde Sapma (Mean Average Percentage Deviation – MAPD) olarak da bilinir. İstatistikte tahmin değerlerinin doğruluğunun ölçüsü olarak kullanılır ve genellikle yüzde olarak ifade edilir. Çalışmamızda büyük değerler ve günlük veriler üzerinde işlem yapıldığından performans değerlendirme ölçütü olarak MAPE kullanılmasına karar verilmiştir. Gerçek

değer sıfır olan durumlarda sıfıra bölünme olacağından dikkat edilmesi gereklidir. Denklem 2.17'de gösterilen formül ile ifade edilir.

$$\text{MAPE} = 100 \cdot \frac{1}{n} \sum_{i}^{n} \frac{|\text{Yt} - \hat{\text{Y}}\text{t}|}{|\text{Yi}|} \quad (2.17)$$

Y_t: Gerçek Değer

$\hat{Y}_t$: Tahmin Değeri

2.8.4. Ortalama Kare Hata (Mean Squared Error – MSE)

Regresyon eğrisinin bir dizi noktaya ne kadar yakın olduğu belirlenir. Değeri her zaman pozitiftir. Sıfıra yaklaştıkça model tahmininin daha iyi performansa sahip olduğunu söylenebilir.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (\text{Yi} - \hat{\text{Y}}\text{i})^2 \quad (2.18)$$

2.8.5. Kök Ortalama Kare Hata (Root Mean Squared Error - RMSE)

RMSE, ortalama kare hatanın kareköküdür. Gerçek değerin kullanılan model tarafından tahmin edilenden farkını ifade etmek için kullanılmaktadır. Tahmin modellerinde tahmin edilen değerler ile gerçek değerleri arasındaki uzaklığın bulunmasında sıklıkla kullanılan ve hatanın büyüklüğünün ölçüldüğü kuadratik bir metriktir. RMSE tahmin edilen değerlerin hatasının standart sapmasıdır. RMSE değeri 0'dan ∞'a kadar değişebilir. RMSE'nin sıfır olması modelin hiç hata yapmadığını gösterirken bu değer ne kadar sıfıra yakınlaştıkça model tahmininin daha iyi performansa sahip olduğunu söylenebilir.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\text{Yi} - \hat{\text{Y}}\text{i})^2} \quad (2.19)$$

2.8.6. Ortalama Yüzde Hata (Mean Percentage Error -MPE)

Gerçek değer ile tahmin edilen değerlin farkının gerçek değere oranının toplamı alınarak hesaplanır. Tahmin hatalarının mutlak değeri değil de gerçek değerleri kullanıldığından, pozitif ve negatif tahmin hataları birbirini dengeleyebilmektedir.

$$\text{MAPE} = 100 \cdot \frac{1}{n} \sum_{i=1}^{n} \frac{\text{Yi} - \hat{\text{Y}}\text{t}}{\text{Yi}} \quad (2.20)$$

Y_t: Gerçek Değer

$\hat{Y}_t$: Tahmin Değeri

3. DENEYLER VE BULGULAR

Çalışmamızda açık kaynak kodlu R programlama dili ve RStudio geliştirme ortamı kullanılmıştır. Dosya işlemleri, veri dönüştürme, görselleştirme, tahmin ve matematiksel işlemler için alttaki kütüphaneler kullanılmıştır [37,38]. Bu paketler R da ilk kullanım öncesi install komutuyla yüklenir, sonrasında ise library() komutuyla kullanılır.

library(readxl) : excel dosyasının içe aktarılarak işlem yapılmasını sağlar
library(tseries) : Zaman serisi modeli oluşturularak analiz imkanı sağlar.
library(ggplot2) : Görselleştirmeler için kullanılır
library(forecast) : Tek değişkenli zaman serisi tahminlerini görüntülemek ve analiz etmek için yöntemler ve araçlar sağlar.
library(fpp2) : Rob J Hyndman ve George Athanasopoulos tarafından yazılan "Tahmin: ilkeler ve uygulama" (2. baskı, 2018) kitabındaki örnekler ve alıştırmalar için gerekli tüm veri setleri ve bunları çalıştırmak için gerekli paketleri kapsar.
library(scales) : Görselleştirme için kullanılır.
library(urca) : birim kök ve eş bütünleşme testleri gerçekleştirilir.

3.1.Basit Doğrusal Regresyon

Basit doğrusal regresyon modeli için uygulanan adımlar alttaki gibidir.

- Kullanılacak kütüphaneler yüklenir. Eğer ilgili kütüphaneler daha önceden yüklenmediyse install komutu ile bir defaya mahsus yüklenmesi sağlanır.
  library(readxl), library(tseries), library(forecast), library(ggplot2)
- Öncelikle üzerinde çalışma yapılacak veri seti read_excel() fonksiyonu ile içe aktarılır.
- Eğitim ve test setleri ayrıştırılır. Deneylerimizde veri setimiz iki ayrı formda kullanılmıştır. İlk olarak 2018 yılı eğitim veri seti olarak kullanılmış, 2019 Ocak ayı (31 gün) tahmin edilmiştir. İkinci olarak da 2018 Ocak – Kasım (Dahil) eğitim seti olarak kullanılmış, Aralık ayı (31 gün) tahmin edilmiştir,
- ts() fonksiyonu kullanılarak her iki veri seti için zaman serisi verisi oluşturulur.
- tslm() fonksiyonu kullanılarak her iki veri seti için zaman serisi modeli oluşturulur.

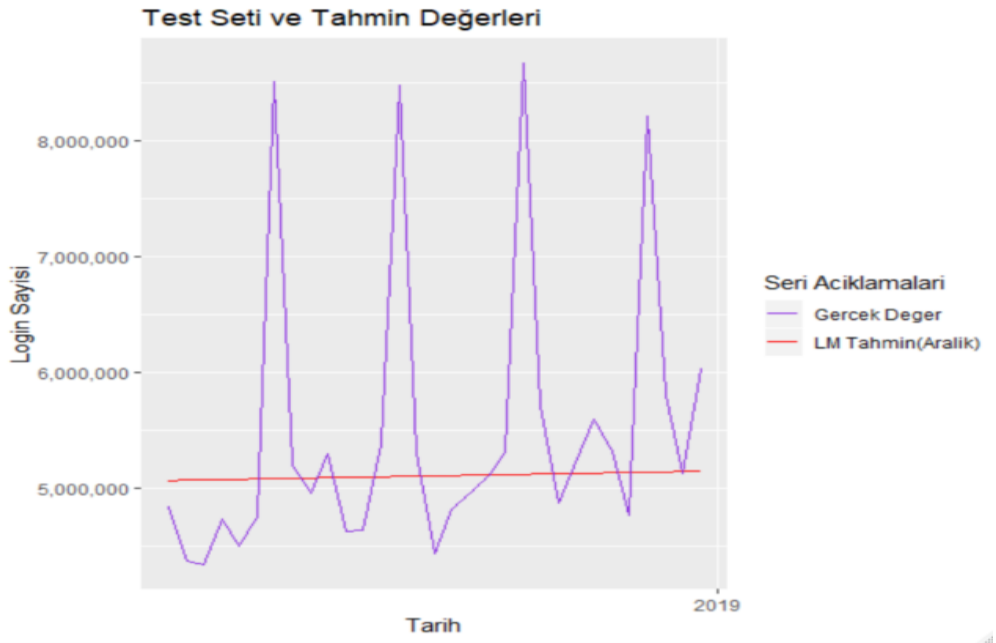
- forecast() fonksiyonu ile sonraki t periyot(bizim deneylerimizde 31 gün) için tahminleme yapılır.
- accuracy() fonksiyonuyla lineer modelin her iki veri seti üzerindeki performansı ölçümlenmiştir. Tablodaki tüm performans ölçüm yöntemleri 2.8. Tahmin Yöntemleri Performans Ölçümü başlığı altında detaylıca incelenmiştir. MAPE sonucuna göre 2019 Ocak ayı tahmini %12,78 hata oranına sahipken, 2018 Aralık ayı için yapılan tahminin %11,35 hata oranına sahip olduğu görülmüştür.

Tablo 3. 1. Doğrusal regresyon tahmini hata oranları

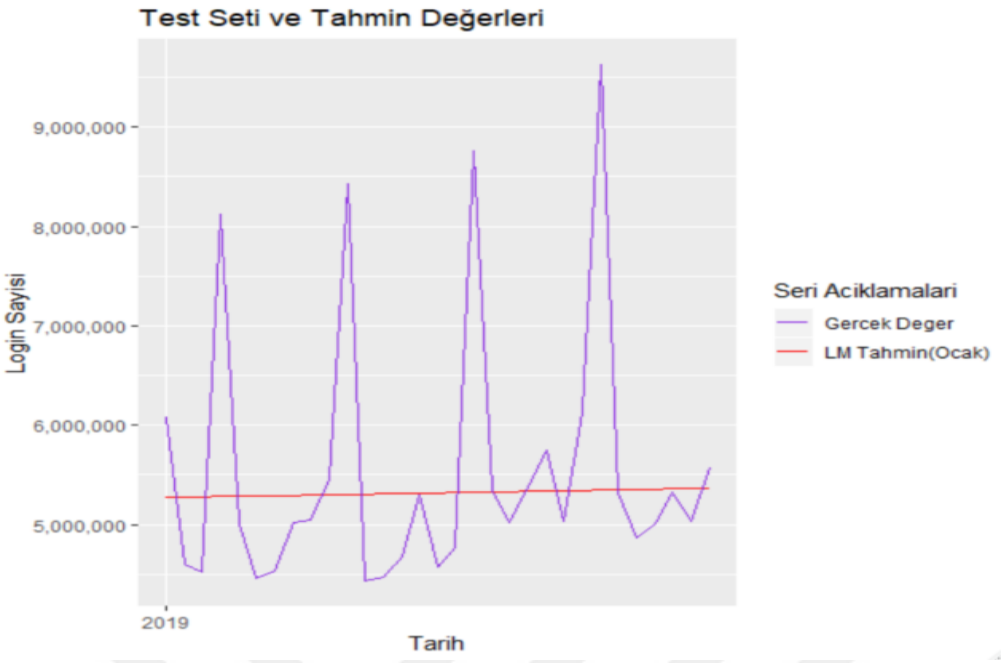
	ME	RMSE	MAE	MPE	MAPE
Doğrusal Regresyon(2019_Ocak_Tahmin)	216949,00	1334744,00	819475,40	0,13	12,78
Doğrusal Regresyon(2018_Aralık_Tahmin)	3,742926e+05	1269014,00	742117,90	3,33	11,35

- autoplot() fonksiyonu ile her iki tahmin modeli için gerçek değer ve tahmin değeri grafiksel gösterimi yapılır.

Üstte belirtilen adımlar sırasıyla uygulanarak her iki veri seti için lineer model tahmin değerleri ve gerçek değerler Şekil 3.1. ve Şekil 3.2 de gösterilmiştir. Doğrusal(lineer) regresyon modelinin doğruluk (accuracy) değeri her ne kadar %12 civarında çıksa da iniş çıkışları doğru olarak tespit edememiş, doğrusal bir çizgi olarak tahminde bulunmuştur.



Şekil 3. 1. Aralık ayı gerçek değer ve lineer model tahmin grafiği



Şekil 3. 2. Ocak ayı gerçek değer ve lineer model tahmin grafiği

3.2.BOX-Jenkins Modeli

BOX-Jenkins (ARIMA) regresyon modeli için uygulanan adımlar alttaki gibidir.

- Kullanılacak kütüphaneler yüklenir. Eğer ilgili kütüphaneler daha önceden yüklenmediyse install komutu ile bir defaya mahsus yüklenmesi sağlanır. library(readxl), library(urca), library(tseries), library(forecast), library(ggplot2), library(scales)
- Öncelikle üzerinde çalışma yapılacak veri seti read_excel() fonksiyonu ile içe aktarılır.
- Eğitim ve test setleri ayrıştırılır. Deneylerimizde veri setimiz iki ayrı formda kullanılmıştır. İlk olarak 2018 yılı eğitim veri seti olarak kullanılmış, 2019 Ocak ayı (31 gün) tahmin edilmiştir. İkinci olarak da 2018 Ocak – Kasım (Dahil) eğitim seti olarak kullanılmış Aralık ayı (31 gün) tahmin edilmiştir,
- ts() fonksiyonu kullanılarak her iki veri seti için zaman serisi verisi oluşturulur.
- adf() ve kpss() fonksiyonları ile durağanlık kontrolü yapılır, serinin durağan olduğu belirlenir. Altta görüleceği üzere durağanlık testlerinde veri seti durağan çıkmıştır. Bu nedenle fark alma gibi yöntemlerle veri setinin durağan hale getirilmesi ihtiyacı bulunmamaktadır.

```
> adf.test(data_ts_3)

        Augmented Dickey-Fuller Test

data:  data_ts_3
Dickey-Fuller = -4.2426, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(data_ts_3) : p-value smaller than printed p-value
```

Şekil 3. 3. ADF durağanlık test sonucu

```
> kpss.test(data_ts_3)

        KPSS Test for Level Stationarity

data:  data_ts_3
KPSS Level = 3.0644, Truncation lag parameter = 5, p-value = 0.01

Warning message:
In kpss.test(data_ts_3) : p-value smaller than printed p-value
>
```

Şekil 3. 4. KPSS durağanlık test sonucu

- tbats() fonksiyonu ile mevsimsellik kontrolü yapılır. Her iki veri seti için mevsimsel/dönemsel olarak değişiklik olup olmadığının kontrol edilmesi edilmesi gereklidir. Yapılan kontrollerde veri setlerinin mevsimsel değişiklik göstermediği tespit edilmiştir.

```
> mevsimsellik_kontrolu <- tbats(data_ts_1)
> Mevsimsellik <- !is.null(mevsimsellik_kontrolu$seasonal)
> Mevsimsellik
[1] FALSE
>
```

Şekil 3. 5. 11 ay eğitim veri seti mevsimsellik test sonucu

```
> mevsimsellik_kontrolu <- tbats(data_ts_3)
> Mevsimsellik <- !is.null(mevsimsellik_kontrolu$seasonal)
> Mevsimsellik
[1] FALSE
```

Şekil 3. 6. 12 ay eğitim veri seti mevsimsellik test sonucu

- auto.arima() fonksiyonu ile uygun ARIMA modeli elde edilir. Veri setimizde ARIMA(2,1,3) modeli uygun çıktığı görülmüştür.

```
Now re-fitting the best model(s) without approximations...

ARIMA(2,1,3)                with drift         : 10131.74

Best model: ARIMA(2,1,3)                with drift
```

Şekil 3. 7. 2018 Aralık ayı tahmini için ARIMA modellemesi

```
Now re-fitting the best model(s) without approximations...

ARIMA(2,1,3)                with drift         : 11074.57

Best model: ARIMA(2,1,3)                with drift
```

Şekil 3. 8. 2019 Ocak ayı tahmini için ARIMA modellemesi

- forecast() fonksiyonu ile sonraki t periyot(bizim deneylerimizde 31 gün) için

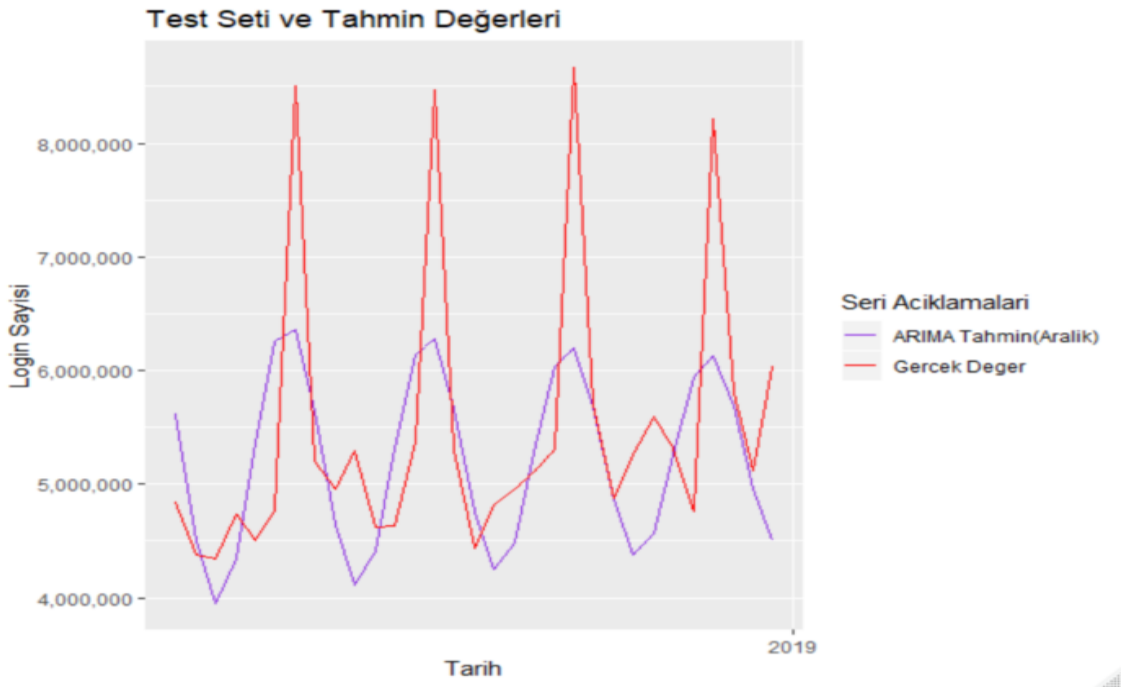
tahminleme yapılır.

- accuracy() fonksiyonuyla lineer modelin her iki veri seti üzerindeki performansı ölçümlenmiştir. Tablodaki tüm performans ölçüm yöntemleri 2.8. Tahmin Yöntemleri Performans Ölçümü başlığı altında detaylıca incelenmiştir. Her iki veri seti için performans ölçümü yapıldığında 2019 Ocak ve 2018 Aralık ayları MAPE hata oranı sırasıyla %15,83 ve %13,13 olarak belirlenmiştir.

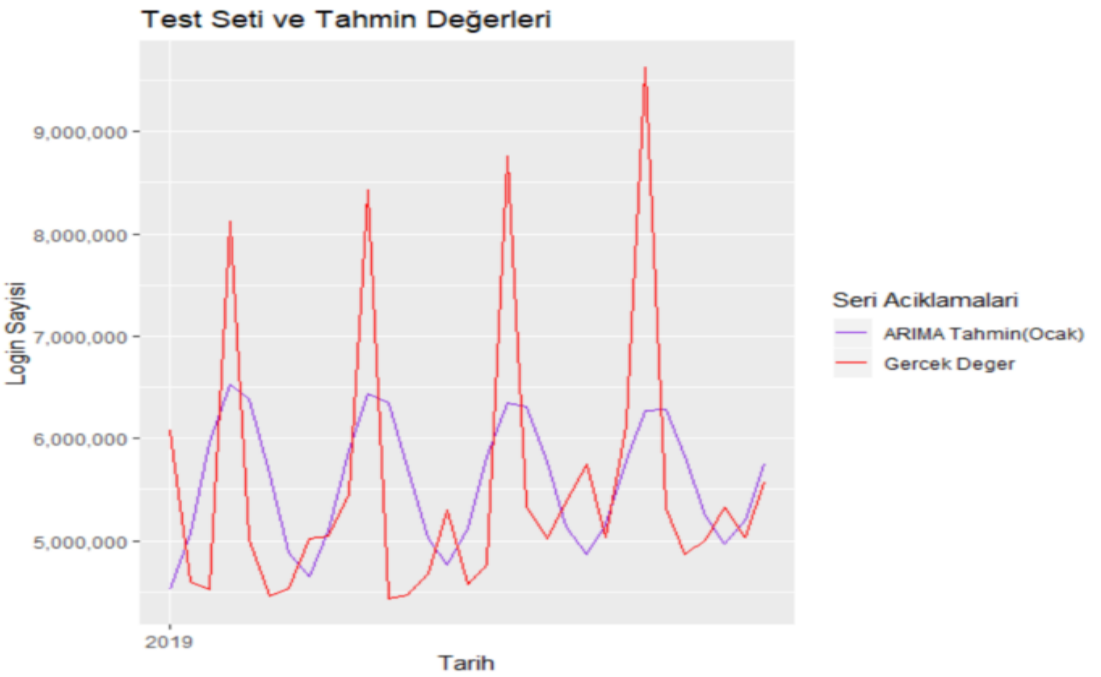
Tablo 3. 2. ARIMA tahmini hata oranları

	ME	RMSE	MAE	MPE	MAPE
ARIMA(2019_Ocak_Tahmin)	36392,20	1185743,60	915735,60	4,16	15,83
ARIMA(2018_Aralık_Tahmin)	270355,38	1039455,40	779057,90	2,66	13,13

- autoplot() fonksiyonu ile gerçek değer ve tahmin değeri grafiksel gösterimi yapılır.



Şekil 3. 9. 2018 Aralık ayı gerçek değer ve ARIMA tahmin grafiği



Şekil 3. 10. 2019 Ocak ayı gerçek değer ve ARIMA tahmin grafiği

ARIMA modeli ile yapılan deneylerde gerçek değere benzer bir örüntüye sahip olarak gelecek dönemi tahmin edilebilmiştir. Yaklaşık olarak %15 hata payıyla tahminde bulunulabilmiştir.

3.3.Naive Model ile tahminleme

Naive model kullanılarak zaman serisi tahminlemesinde uygulanan adımlar alttaki gibidir.

- Kullanılacak kütüphaneler yüklenir. Eğer ilgili kütüphaneler daha önceden yüklenmediyse install komutu ile bir defaya mahsus yüklenmesi sağlanır. library(readxl), library(tseries), library(forecast), library(ggplot2), library(scales)
- Çalışma yapılacak veri seti read_excel() fonksiyonu ile içe aktarılır.
- ts() fonksiyonu kullanılarak her iki veri seti için zaman serisi verisi oluşturulur.
- tbats() fonksiyonu ile mevsimsellik kontrolü yapılır. Her iki veri seti için mevsimsel/dönemsel olarak değişiklik olup olmadığının kontrol edilmesi gereklidir. Yapılan kontrollerde veri setlerinin mevsimsel değişiklik göstermediği tespit edilmiştir.

```
> mevsimsellik_kontrolu <- tbats(data_ts_1)
> Mevsimsellik <- !is.null(mevsimsellik_kontrolu$seasonal)
> Mevsimsellik
[1] FALSE
>
```

Şekil 3. 11. 11 ay eğitim veri seti mevsimsellik test sonucu

```
> mevsimsellik_kontrolu <- tbats(data_ts_3)
> Mevsimsellik <- !is.null(mevsimsellik_kontrolu$seasonal)
> Mevsimsellik
[1] FALSE
```

Şekil 3. 12. 12 ay eğitim veri seti mevsimsellik test sonucu

- naive() fonksiyonu ile her iki veri seti kullanılarak sonraki t periyot(31 gün) için tahminleme yapılır.

```
> naive_tahmin_1 <- naive(data_ts_1,h=31)$mean
> naive_tahmin_3 <- naive(data_ts_3,h=31)$mean
```

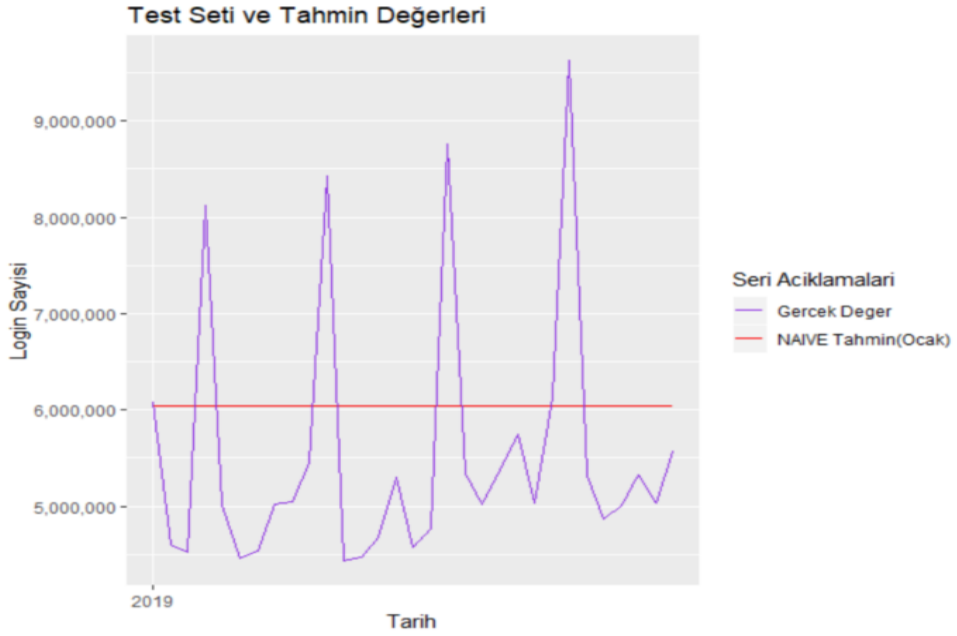
Şekil 3. 13. 31 günlük tahminleme

- accuracy() fonksiyonuyla lineer modelin her iki veri seti üzerindeki performansı ölçümlenmiştir. Tablodaki tüm performans ölçüm yöntemleri 2.8. Tahmin Yöntemleri Performans Ölçümü başlığı altında detaylıca incelenmiştir. Her iki veri seti için performans ölçümü yapıldığında 2019 Ocak ve 2018 Aralık ayları MAPE hata oranı sırasıyla %21,68 ve %75,51 olarak belirlenmiştir.

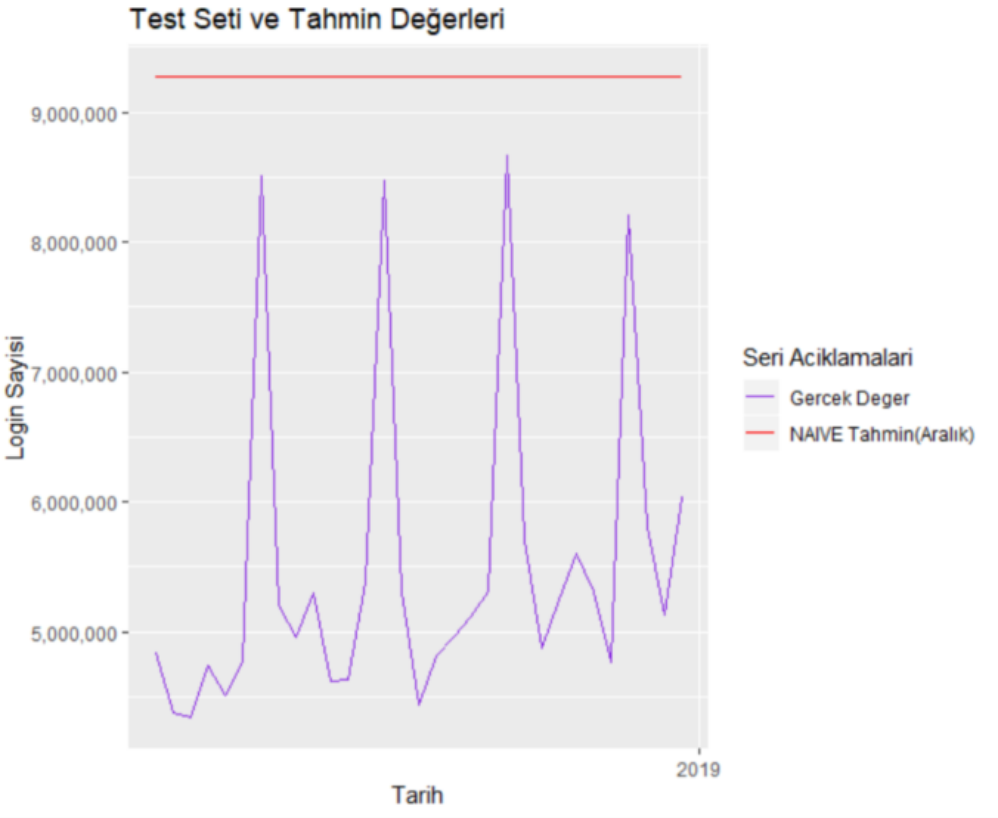
Tablo 3. 3. NAIVE tahmin hata oranları

	ME	RMSE	MAE	MPE	MAPE
NAIVE(2019_Ocak_Tahmin)	-500474,20	1410898,00	1204188,00	13,65	21,68
NAIVE(2018_Aralık_Tahmin)	3787967,00	3979110,00	3787967,00	75,51	75,51

- autoplot() fonksiyonu ile gerçek değer ve tahmin değeri grafiksel gösterimi yapılır.



Şekil 3. 14. 2019 Ocak ayı gerçek değer ve Naive model tahmin grafiği



Şekil 3. 15. 2018 Aralık ayı gerçek değer ve Naive model tahmin grafiği

Naive modelde %14 civarında bir hata payıyla tahminde bulunmuştur. Ancak doğrusal modeldeki gibi gerçek değerlerin sahip olduğu iniş çıkışı yakalayamamıştır.

3.4.Üstel Düzleştirme Yöntemleri

3.4.1. Basit Üstel Düzleştirme Yöntemi

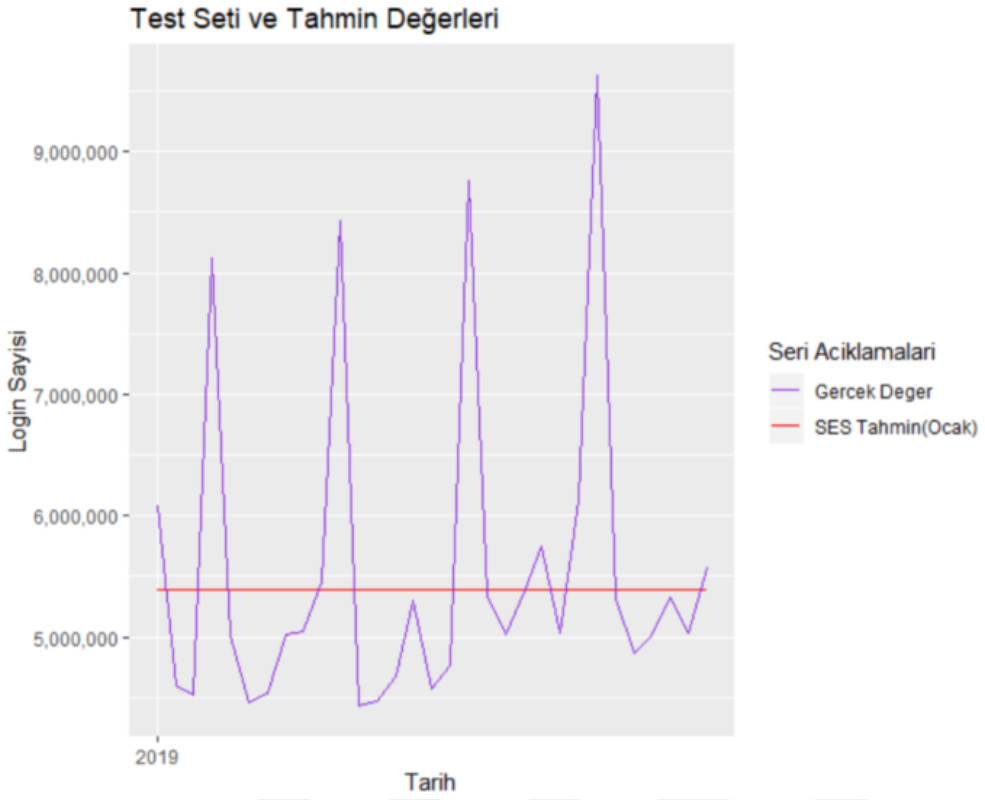
Basit üstel düzleştirme yöntemi kullanılarak zaman serisi tahmini için uygulanan adımlar alttaki gibidir.

- Kullanılacak kütüphaneler yüklenir. Eğer ilgili kütüphaneler daha önceden yüklenmediyse install komutu ile bir defaya mahsus yüklenmesi sağlanır. library(readxl), library(tseries), library(forecast), library(ggplot2), library(scales)
- Çalışma yapılacak veri seti read_excel() fonksiyonu ile içe aktarılır.
- Eğitim ve test setleri ayrıştırılır. Deneylerimizde veri setimiz iki ayrı formda kullanılmıştır. İlk olarak 2018 yılı eğitim veri seti olarak kullanılmış, 2019 Ocak ayı (31 gün) tahmin edilmiştir. İkinci olarak da 2018 Ocak – Kasım (Dahil) eğitim seti olarak kullanılmış Aralık ayı (31 gün) tahmin edilmiştir,
- ts() fonksiyonu kullanılarak her iki veri seti için zaman serisi verisi oluşturulur.
- ses() fonksiyonu ile her iki veri seti için basit üstel düzleştirme modeli oluşturulur.
- forecast() fonksiyonu ile sonraki t periyot (bizim deneylerimizde 31 gün) için tahminleme yapılır.
- accuracy() fonksiyonuyla lineer modelin her iki veri seti üzerindeki performansı ölçümlenmiştir. Tablodaki tüm performans ölçüm yöntemleri 2.8. Tahmin Yöntemleri Performans Ölçümü başlığı altında detaylıca incelenmiştir. Her iki veri seti için performans ölçümü yapıldığında 2019 Ocak ve 2018 Aralık ayları MAPE hata oranı sırasıyla %11,54 ve %13,51 olarak belirlenmiştir.

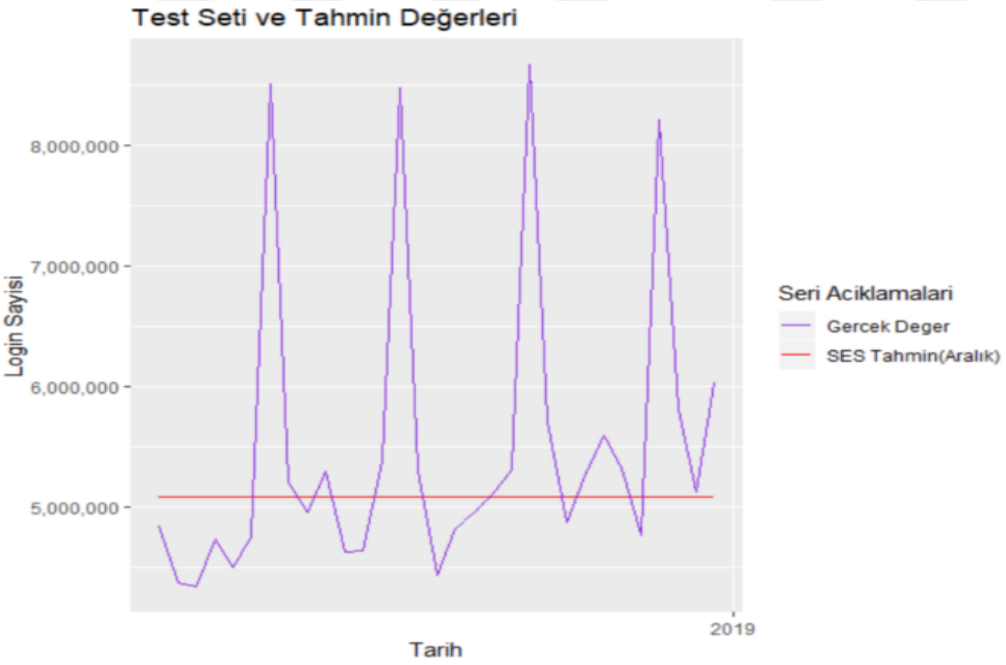
Tablo 3. 4. Basit üstel düzleştirme tahmini hata oranları

	ME	**RMSE**	**MAE**	**MPE**	**MAPE**
Basit Üstel Düzleştirme (SES_2019_Ocak_Tahmin)	141570,20	1326725,00	847796,70	1,56	13,51
Basit Üstel Düzleştirme (SES_2018_Aralık_Tahmin)	393421,40	1280391,00	754214,20	3,66	11,54

- autoplot() fonksiyonu ile gerçek değer ve tahmin değeri grafiksel gösterimi yapılır.



Şekil 3. 16. 2019 Ocak ayı gerçek değer ve basit üstel düzleştirme tahmin grafiği



Şekil 3. 17. 2018 Aralık ayı gerçek değer ve basit üstel düzleştirme tahmin grafiği

3.4.2. Üstel Düzleştirme Yöntemi

Üstel düzleştirme yöntemi kullanılarak zaman serisi tahmini için uygulanan adımlar alttaki gibidir.

- Kullanılacak kütüphaneler yüklenir. Eğer ilgili kütüphaneler daha önceden yüklenmediyse install komutu ile bir defaya mahsus yüklenmesi sağlanır. library(readxl), library(tseries), library(forecast), library(ggplot2), library(scales)
- Öncelikle üzerinde çalışma yapılacak veri seti read_excel() fonksiyonu ile içe aktarılır.
- Eğitim ve test setleri ayrıştırılır. Deneylerimizde veri setimiz iki ayrı formda kullanılmıştır. İlk olarak 2018 yılı eğitim veri seti olarak kullanılmış, 2019 Ocak ayı (31 gün) tahmin edilmiştir. İkinci olarak da 2018 Ocak – Kasım (Dahil) eğitim seti olarak kullanılmış Aralık ayı (31 gün) tahmin edilmiştir,
- ts() fonksiyonu kullanılarak her iki veri seti için zaman serisi verisi oluşturulur.
- ets() fonksiyonu (ExponenTial Smoothing) ile her iki veri seti için üstel düzleştirme modeli oluşturulur.
- forecast() fonksiyonu ile sonraki t periyot(bizim deneylerimizde 31 gün) için tahminleme yapılır.

```
> #ExponanTial Smoothing
> ets_tahmin_1 <- forecast(ets(data_ts_1), h=31)$mean
> ets_tahmin_3 <- forecast(ets(data_ts_3), h=31)$mean
```

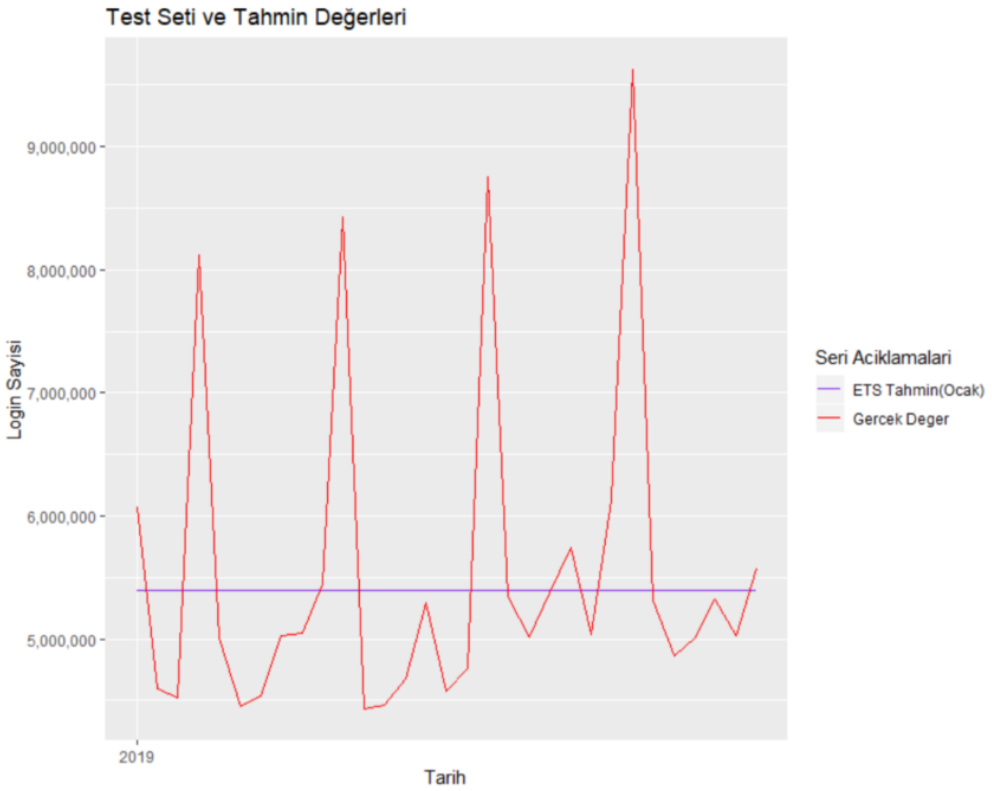
Şekil 3. 18. Üstel Düzleştirme Tahmin

- accuracy() fonksiyonuyla üstel modelin her iki veri seti üzerindeki performansı ölçümlenmiştir. Tablodaki tüm performans ölçüm yöntemleri 2.8. Tahmin Yöntemleri Performans Ölçümü başlığı altında detaylıca incelenmiştir. Her iki veri seti için performans ölçümü yapıldığında 2019 Ocak ve 2018 Aralık ayları MAPE hata oranı sırasıyla %13,50 ve %11,54 olarak belirlenmiştir.

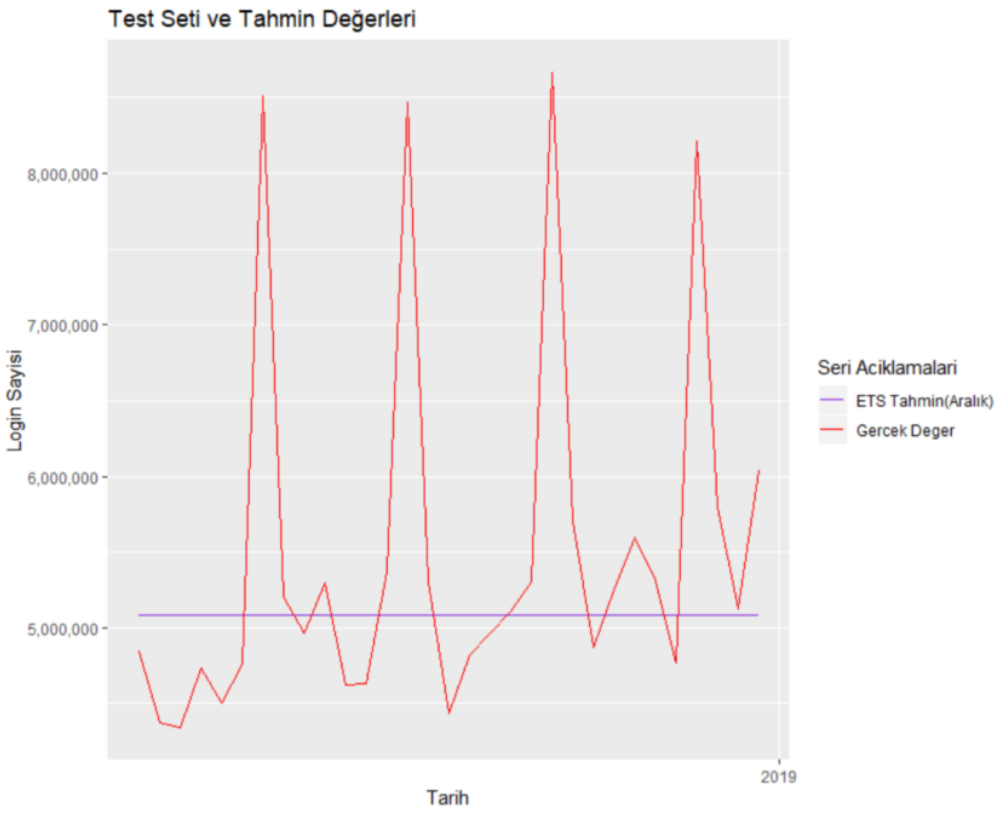
Tablo 3. 5. Üstel düzleştirme tahmini hata oranları

	ME	RMSE	MAE	MPE	MAPE
Üstel Düzleştirme (ETS_2019_Ocak_Tahmin)	142669	1326843	847336	1,54	13,50
Üstel Düzleştirme (ETS_2018_Aralık_Tahmin)	394390	1280689	754308	3,68	11,54

- autoplot() fonksiyonu ile gerçek değer ve tahmin değeri grafiksel gösterimi yapılır.



Şekil 3. 19. 2019 Ocak ayı gerçek değer ve üstel düzleştirme modeli tahmin grafiği



Şekil 3. 20. 2018 Aralık ayı gerçek değer ve üstel düzleştirme modeli tahmin grafiği

Üstel düzleştirme yöntemi için %80 ve %95 güven aralıkları içindeki ortalama tahmin değerleri kontrol edildiğinde ~%12 hata oranına sahip doğrusal bir tahmin grafiği ortaya çıkmıştır. Hata oranı iyi görünse de gerçek verilerin örüntüsü tahmin edilememiştir.

3.5.Destek Vektör Makinesi Kullanılarak Tahminleme

Destek Vektör Makinesi ile zaman serisi tahminlemesinde uygulanan adımlar alttaki gibidir.

- Kullanılacak kütüphaneler yüklenir. Eğer ilgili kütüphaneler daha önceden yüklenmediyse install komutu ile bir defaya mahsus yüklenmesi sağlanır. library(readxl), library(e1071), library(forecast), library(ggplot2), library(scales)
- Öncelikle üzerinde çalışma yapılacak veri seti read_excel() fonksiyonu ile içe aktarılır.
- Eğitim ve test setleri ayrıştırılır. Deneylerimizde veri setimiz iki ayrı formda

kullanılmıştır. İlk olarak 2018 yılı eğitim veri seti olarak kullanılmış, 2019 Ocak ayı (31 gün) tahmin edilmiştir. İkinci olarak da 2018 Ocak – Kasım (Dahil) eğitim seti olarak kullanılmış Aralık ayı(31 gün) tahmin edilmiştir,

- svm() fonksiyonu (Support Vector Machines) ile her iki veri seti için Destek Vektör Makineleri modeli oluşturulur.

```
> #Regression with SVM
> modelsvm_1 <- svm(login_data_1$adet, login_data_1, type="eps-regression",kernel="radial")
> modelsvm_2 <- svm(login_data_2$adet, login_data_2, type="eps-regression",kernel="radial")
```

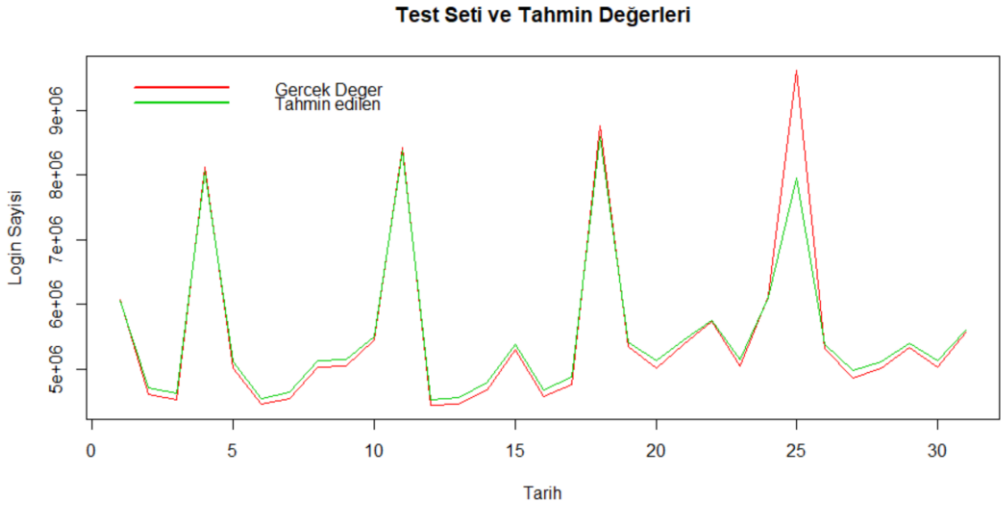
Şekil 3. 21. SVM Model Oluşturma

- predict() fonksiyonu ile sonraki t periyot(bizim deneylerimizde 31 gün) için tahminleme yapılır.
- accuracy() fonksiyonuyla lineer modelin her iki veri seti üzerindeki performansı ölçümlenmiştir. Tablodaki tüm performans ölçüm yöntemleri 2.8. Tahmin Yöntemleri Performans Ölçümü başlığı altında detaylıca incelenmiştir. Her iki veri seti için performans ölçümü yapıldığında 2019 Ocak ve 2018 Aralık ayları MAPE hata oranı sırasıyla %2,18 ve %1,70 olarak belirlenmiştir.

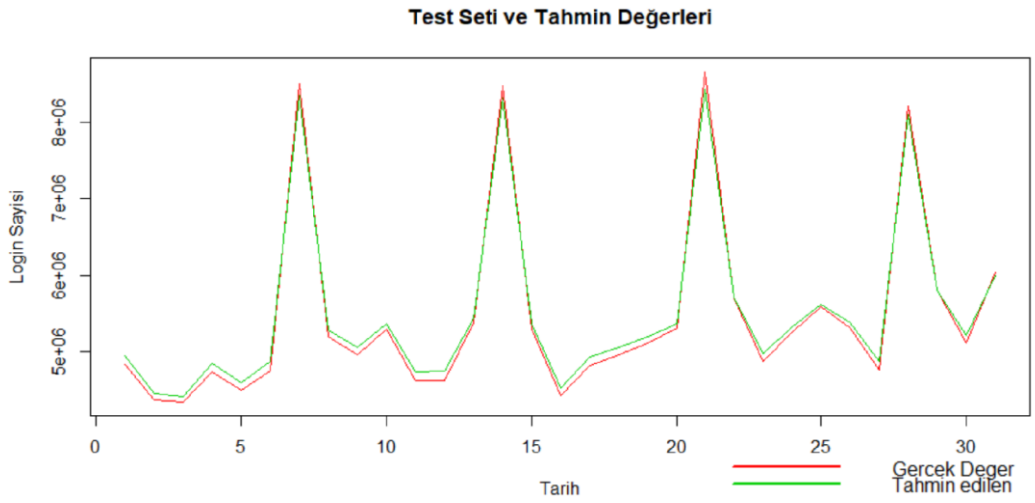
Tablo 3. 6. SVM tahmini hata oranları

	ME	RMSE	MAE	MPE	MAPE
Destek Vektör Makineleri (SVM_2019_Ocak_Tahmin)	-7031	313024	137088	-0,8	2,18
Destek Vektör Makineleri (SVM_2018_Aralık_Tahmin)	46480	102312	92445	1,15	1,70

- plot() fonksiyonu ile gerçek değer ve tahmin değeri grafiksel gösterimi yapılır.



Şekil 3. 22 2019 Ocak ayı gerçek değer ve SVM modeli tahmin grafiği



Şekil 3. 23. 2018 Aralık ayı gerçek değer ve SVM modeli tahmin grafiği

Destek vektör makineleri yöntemiyle yapılan tahminlerde gerçek değerlerin davranışı çok yüksek doğrulukla (~%98) tahmin edilebilmiştir.

3.6.Yapay Sinir Ağları Kullanılarak Tahminleme (YSA)

Yapay Sinir Ağları ile zaman serisi tahmin çalışmasında uygulanan adımlar alttaki gibidir.

- Kullanılacak kütüphaneler yüklenir. Eğer ilgili kütüphaneler daha önceden yüklenmediyse install komutu ile bir defaya mahsus yüklenmesi sağlanır. library(readxl), library(forecast), library(nnfor), library(neuralnet), library(ggplot2), library(scales)
- Öncelikle üzerinde çalışma yapılacak veri seti read_excel() fonksiyonu ile içe aktarılır.
- Eğitim ve test setleri ayrıştırılır. Deneylerimizde veri setimiz iki ayrı formda kullanılmıştır. İlk olarak 2018 yılı eğitim veri seti olarak kullanılmış, 2019 Ocak ayı (31 gün) tahmin edilmiştir. İkinci olarak da 2018 Ocak – Kasım (Dahil) eğitim seti olarak kullanılmış Aralık ayı (31 gün) tahmin edilmiştir,
- ts() fonksiyonu kullanılarak her iki veri seti için zaman serisi verisi oluşturulur.
- Yapay sinir ağları modeli için nnetar() fonksiyonu kullanılarak her iki veri seti için YSA modeli oluşturulur.

```
> #Yapay Sinir Aglari nnetar modelleri
> nn_model_1 <- nnetar(data_ts_1,size = 4, P=7,lambda = "auto", repeats = 20)
Warning message:
In nnetar(data_ts_1, size = 4, P = 7, lambda = "auto", repeats = 20) :
  Series too short for seasonal lags
> nn_model_2 <- nnetar(data_ts_2,size = 4, P=7,lambda = "auto", repeats = 20)
Warning message:
In nnetar(data_ts_2, size = 4, P = 7, lambda = "auto", repeats = 20) :
  Series too short for seasonal lags
```

Şekil 3. 24. YSA model oluşturma

- forecast() fonksiyonu ile sonraki t periyot(bizim deneylerimizde 31 gün) için tahmin yapılır.

```
> #Yapay Sinir Aglari tahmin
> nn_tahmin_1 <- forecast(nn_model_1, h=31)
> nn_tahmin_2 <- forecast(nn_model_2, h=31)
```

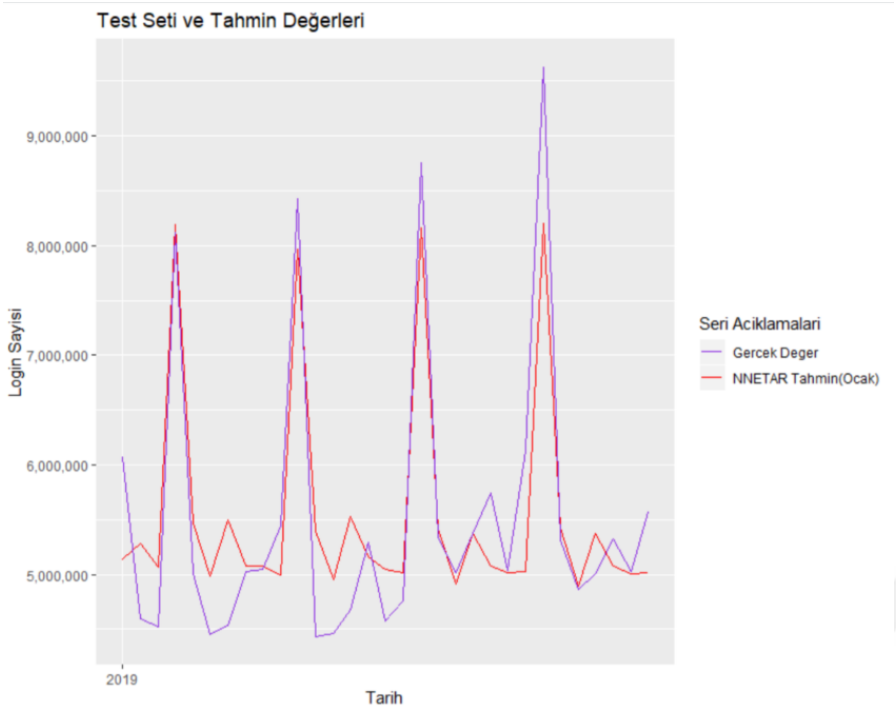
Şekil 3. 25. YSA tahmin

- accuracy() fonksiyonuyla lineer modelin her iki veri seti üzerindeki performansı ölçümlenmiştir. Tablodaki tüm performans ölçüm yöntemleri 2.8. Tahmin Yöntemleri Performans Ölçümü başlığı altında detaylıca incelenmiştir. Her iki veri seti için performans ölçümü yapıldığında 2019 Ocak ve 2018 Aralık ayları MAPE hata oranı sırasıyla %8,08 ve %7,51 olarak belirlenmiştir.

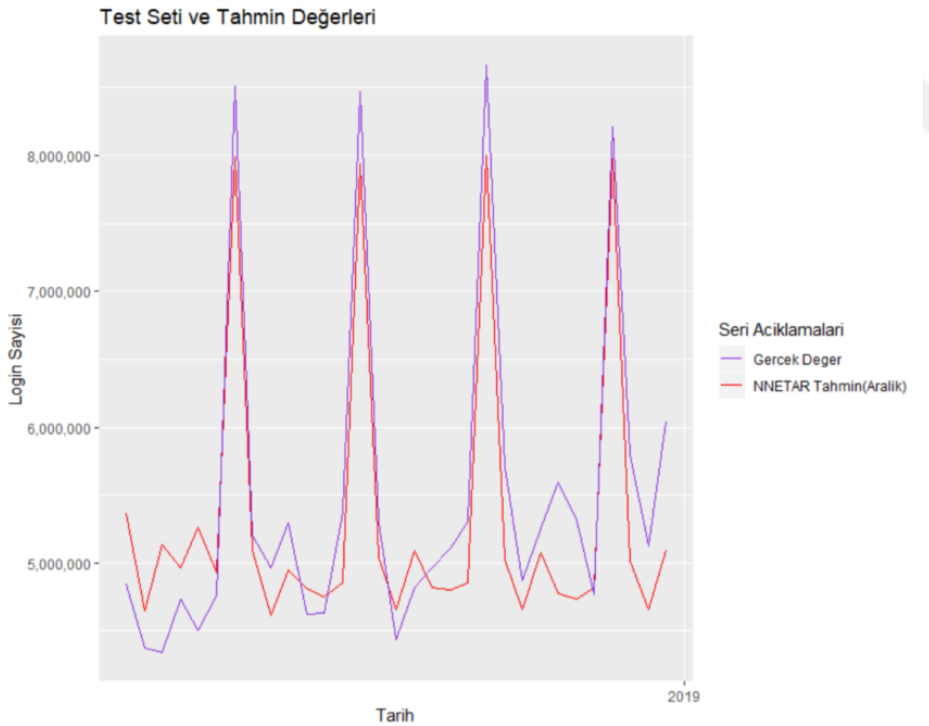
Tablo 3. 7. YSA tahmini hata oranları

	ME	**RMSE**	**MAE**	**MPE**	**MAPE**
YSA (NNETAR_2019_Ocak_Tahmin)	-8055	574763	441389	1,60	8,08
YSA (NNETAR_2018_Aralık_Tahmin)	178252	476585	409691	2,44	7,51

- plot() fonksiyonu ile gerçek değer ve tahmin değeri grafiksel gösterimi yapılır.



Şekil 3. 26. 2019 Ocak ayı gerçek değer ve SVM modeli tahmin grafiği



Şekil 3. 27. 2018 Aralık ayı gerçek değer ve SVM modeli tahmin grafiği

Yapay sinir ağları kullanılarak yapılan deneylerde ~%92 doğruluk oranıyla tahmin edilebilmiştir, genel olarak gerçek verinin sahip olduğu iniş çıkış örüntüsü de tahmin edilebilmiştir.

4. SONUÇLAR

Tez çalışmasının temel amacı istatistiksel yöntemler ve makine öğrenmesi yöntemleri ile geçmiş dönem verileri kullanılarak gelecek dönemle ilgili tahminlerde bulunmak ve kullanılan veri seti için en yüksek tahmin başarımına sahip yöntemi belirlemektedir. Veri seti olarak Turkcell İletişim Hizmetleri. A.Ş ye ait bir mobil uygulamanın 2018 yılı ve 2019 yılı Ocak ayına ait günlük giriş sayıları kullanılmıştır. Veri seti iki ayrı formda kullanılmıştır. İlk olarak 2018 yılı eğitim veri seti olarak kullanılmış, 2019 Ocak ayı (31 gün) tahmin edilmiştir. İkinci olarak da 2018 Ocak – Kasım (Dahil) eğitim seti olarak kullanılmış, Aralık ayı (31 gün) tahmin edilmiştir, Çalışmamızda Doğrusal Regresyon, ARIMA, Naive metod, üstel düzleştirme ve basit üstel düzleştirme yöntemleri, destek vektör makineleri ve yapay sinir ağları yöntemleri kullanılmıştır. Çalışmada kullanılan yöntemlerin parametreleri değiştirilerek çeşitli deneyler yapılmış sonuçları kıyaslanmıştır. Performans değerlendirme yöntemi olarak pozitif ve büyük boyutlu sayılarla işlemler yapıldığında MAPE tercih edilmiştir. Deney sonuçlarına göre en yüksek başarım oranı DVM kullanıldığında %2,18 ve %1,70 hata oranları ile elde edilmiştir. İkinci başarılı yöntem ise %8-08 ve %7,51 hata oranları ile yapay sinir ağları ile elde edilmiştir. Bunlar arkasından ise istatistiksel yöntemler olan Doğrusal Regresyon, ETS, SES birbirine yakın başarımla yaklaşık %11,5 hata oranıyla takip etmektedir. Kullanılan tüm yöntemler için başarım oranları Tablo 4.1.'de verilmiştir.

Çalışmamızda telekomünikasyon sektöründe faaliyet gösteren bir mobil uygulamanın kullanım verisi kullanılarak yüksek doğruluk oranıyla tahmin yapabildiği görülmüştür. Bu çalışma rehberliğinde hem bu sektör de hem de farklı alanlarda faaliyet gösteren benzer özellikteki mobil uygulamalar için belirlenen işlemlerin (giriş, ödeme, satın alma, görüntüleme vb.) saatlik, günlük, aylık gibi periyotlarında gelecek döneme dair tahminler yapılarak kapasite ve altyapı planlaması, reklam planlaması, ürün satış ve stok planlaması gibi birçok optimizasyon ve stratejik karar avantajı elde edilebilir. Çalışmamızda günlük veriler üzerinden tahminler yapılmıştır. Sonraki çalışmalarda periyot zamanı daraltılarak saatlik, 15 dakikalık, 5 dakikalı ya da gerçek zamanlı veri üzerinde bu yöntemlerin uygulanmasıyla tahmin çalışmasının yapılması anlık veri ihtiyacı olan alanlarda fayda sağlayacaktır. Ek olarak veri setinde mümkün olan daha fazla nitelik kullanımı ile daha yüksek doğrulukta tahminler yapılmasına olanak

sağlayacaktır.

Tablo 4. 1. Genel tahmin hata performans tablosu

Model	**MPE**	**MAPE**
Doğrusal Regresyon(2019_Tahmin)	-0,13	12,78
Doğrusal Regresyon(2018_Aralık_Tahmin)	3,33	11,35
ARIMA(2019_Tahmin)	-4,16	15,83
ARIMA(2018_Aralık_Tahmin)	2,66	13,13
NAIVE(2019_Tahmin)	13,65	21,68
NAIVE(2018_Aralık_Tahmin)	75,51	75,51
Basit Üstel Düzleştirme (SES_2019_Ocak_Tahmin)	-1,56	13,51
Basit Üstel Düzleştirme (SES_2018_Aralık_Tahmin)	3,66	11,54
Üstel Düzleştirme (ETS_2019_Ocak_Tahmin)	-1,54	13,50
Üstel Düzleştirme (ETS_2018_Aralık_Tahmin)	3,68	11,54
Destek Vektör Makineleri (SVM_2019_Ocak_Tahmin)	-0,78	2,18
Destek Vektör Makineleri (SVM_2018_Aralık_Tahmin)	-1,15	1,70
YSA (NNETAR_2019_Ocak_Tahmin)	-1,60	8,08
YSA (NNETAR_2018_Aralık_Tahmin)	2,44	7,51

KAYNAKLAR

[1] Türkiye Elektronik Haberleşme Sektörü – Üç Aylık Pazar Verileri Raporu, https://www.btk.gov.tr/uploads/pages/pazar-verileri/2020-1-16-07-20-kurum-disi.pdf Erişim Tarihi: 06/09/2020

[2] Tezuran, M. (2019) Çağrı Merkezi Performans Analizi ve Çağrı Sayılarının Mevsimsel Tahminlemesi. Yüksek Lisans Tezi, Marmara Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, Türkiye, 6, 41, 93-95

[3] Gültekin, O. G. (2019) Comparison of computational intelligence models on forecasting ATM demands. Yüksek Lisans Tezi, Marmara Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, Türkiye, 11, 13, 16, 18, 51

[4] Singh, P., Gupta, P., & Jyoti, K. (2019). Tasm: technocrat arima and svr model for workload prediction of web applications in cloud. Cluster Computing, 22(2), 619-633.

[5] Karakoyun, E. Ş. (2018) Derin Öğrenme İle Zaman Serilerinin Gerçek Zamanlı Tahmini. Yüksek Lisans Tezi, Necmettin Erbakan Üniversitesi Fen Bilimleri Enstitüsü, Konya, Türkiye, 12, 47, 65

[6] Kaynar, O., & Taştan, S. (2009). Zaman Serisianalizinde MLP Yapay Sinir Ağları ve Arıma Modelinin Karşılaştırılması. Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, (33), 161-172.

[7] ÇELİK, S. (2013). Time Series Analysis and Its Applications to Data on Traf c Accidents. Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 3(4), 43-51.

[8] Atalay, M. (2013) Makine öğrenme yöntemleri yardımıyla tüketim istatistiklerine göre talep tahmini. Yüksek Lisans Tezi, İstanbul Kültür Üniversitesi, İstanbul, Türkiye, 1, 25-28, 56, 57

[9] SALAUDEEN, A.A. (2008) Forecasting patient volume in an emergency department using time series methods: An application of a medical center in Nigeria. Yüksek lisans Tezi, Yaşar Üniversitesi Sosyal Bilimler Enstitüsü, İzmir, Türkiye, 13-22, 72-73

[10] Chatfield C., (2000). Time-Series Forecasting, CRC press, New York, USA.

[11] Seddighi, H., Lawler, K. A., Lawler, K., & Katos, A. V. (2000). Econometrics: A practical approach. Routlenge Taylor ve Francis Group, p. 252, London, UK.

[12] Adhikari, R. and Agrawal, R. K., (2013). An Introductory Study on Time Series

Modeling and Forecasting, LAP Lambert Academic Publishing, Germany.

[13] Raicharoen, T., Lursinsap, C., & Sanguanbhokai, P. (2003, May). Application of critical support vector machine to time series prediction. In Proceedings of the 2003 International Symposium on Circuits and Systems, 2003. ISCAS'03. (Vol. 5, pp. V-V). IEEE.

[14] Weigend, A, (2018). Time series prediction: forecasting the future and understanding the past, Routledge, New York, USA.

[15] AKGÜL, I. (1994) Zaman Serisi Analizi ve Öngörü Modelleri, Öneri: Marmara Üniversitesi, Sosyal Bilimler Enstitüsü Dergisi, 1(1), 52.

[16] Ediger, V. Ş., & Akar, S. (2007). ARIMA forecasting of primary energy demand by fuel in Turkey. Energy policy, 35(3), 1701-1708.

[17] Ariyo, A. A., Adewumi, A. O., & Ayo, C. K. (2014, March). Stock price prediction using the ARIMA model. In 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation (pp. 106-112). IEEE.

[18] Triphaty, N. (2017). Forecasting Gold Price with Auto Regressive Integrated Moving Avarage Model. International Journal of Economics and Finance Issues, 7(4), 324- 329.

[19] Makridakis, S. G., & Wheelwright, S. C. (1989). Forecasting methods for management. Fifth Edition, John Wiley & Sons Inc. New York, s. 67

[20] Carnot, N., Koen V. & Tissot, B. (2005). Economic Forecasting. New York: Palgrave Macmillan.

[21] HANKE, J. E. ve REITSCH, A.G., Business Forecasting, (Fourth Edition), Allyn and Bacon, Boston, 1992, s. 132

[22] Armutlulu, İ. H. (2000). İşletmelerde Uygulamalı İstatistik, Alfa Basım Yayın Dağıtım Ltd. Şti., İstanbul.

[23] Hyndman, R. and George, A. (2019). Exponential Smoothing, https://otexts.com/fpp2/expsmooth.html Erişim Tarihi: 15/12/2020

[24] Orhunbilge, N., (2002). Uygulamalı Regresyon ve Korelasyon Analizi, İÜ Yayınları, Yayın No: 281, İstanbul. s. 95

[25] Vapnik, V.N., 1995, The Nature of Statistical Learning Theory, Springer-Verlag, New York.

[26] Vapnik, V.N., 2000, The Nature of Statistical Learning Theory, 2. Baskı, Springer-Verlag, New York.
[27] Oğcu, G., Demirel, O. F., & Zaim, S. (2012). Forecasting electricity consumption with neural networks and support vector regression. Procedia-Social and Behavioral Sciences, 58, 1576-1585.
[28] Elattar, E. E., Goulermas, J., & Wu, Q. H. (2010). Electric load forecasting based on locally weighted support vector regression. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40(4), 438-447.
[29] Osowski, S., Siwek, K., & Markiewicz, T. (2004, June). Mlp and svm networks-a comparative study. In Proceedings of the 6th Nordic Signal Processing Symposium, 2004. NORSIG 2004. (pp. 37-40). IEEE.
[30] Nitze, I., Schulthess, U., & Asche, H. (2012). Comparison of machine learning algorithms random forest, artificial neural network and support vector machine to maximum likelihood for supervised crop type classification. Proc. of the 4th GEOBIA, 35.
[31] Ayhan, S., Erdogmus, S., 2014. Kernel Function Selection for the Solution of Classification Problems via Support Vector Machines. Eskisehir Osmangazi University Journal of IIBF, 9, pp. 175-198.
[32] Wang, Z. T., Zhao, N. B., Wang, W. Y., Tang, R., & Li, S. Y. (2015). A fault diagnosis approach for gas turbine exhaust gas temperature based on fuzzy c-means clustering and support vector machine. Mathematical Problems in Engineering, 2015.
[33] Alpaslan, F., Cagcag, O., Aladag, C. H., Yolcu, U., & Egrioglu, E. (2012). A novel seasonal fuzzy time series method. Hacettepe Journal of Mathematics and Statistics, 41(3), 375-385.
[34] 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.
[35] Song, K. B., Baek, Y. S., Hong, D. H., & Jang, G. (2005). Short-term load forecasting for the holidays using fuzzy linear regression method. IEEE transactions on power systems, 20(1), 96-101.
[36] Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. Computer, 29(3), 31-44.

[37] RStudio, Take control of your R code, https://rstudio.com/products/rstudio/ Erişim Tarihi: 15/08/2020

[38] Available CRAN Packages By Name,
https://cran.r-project.org/web/packages/available_packages_by_name.html
Erişim Tarihi: 15/08/2020

ÖZGEÇMİŞ

Öğrenim Durumu	
Bilgisayar Mühendisliği Yüksek Lisans	Marmara Üniversitesi Fen Bilimleri Enstitüsü (2019-)
Bilgisayar Mühendisliği Lisans	Gebze Yüksek Teknoloji Enstitüsü (2012)
Lise	Elazığ Anadolu Lisesi (2006)

İş Deneyimi	
Test Mühendisi (2012-2013)	KG Teknoloji - Turkcell Teknoloji
Sistem Mühendisi (2013 - 2015)	Turkcell İletişim Hizm. A.Ş
Uygulama ve Sistem Mühendisi (2016 - Devam ediyor)	Turkcell İletişim Hizm. A.Ş

Yabancı Dil ve Sınavlar	
İngilizce	YÖKDİL - 87.5/100
ALES	88.5/100