

Dow Jones, NASDAQ, S&P-500 ve Altın Fiyatlarının Bist-100 Üzerindeki Etkisi

Şeyhmus Aslan[20080966]

Özet

Analitik finans alanında, geçmiş verilere dayalı geleceğe yönelik tahminler oluşturmak büyük önem taşır. Son zamanlarda ilerlemiş istatistiksel yöntemler, bu konuda artan bir ilgiyi ön plana çıkarmıştır. Bu yöntemlerden biri olan zaman serisi regresyonu, belirli periyotlarda görülen değişimleri regresyon modeli yardımıyla tahmin etmeyi ve bu tahminin sonuçlarını ölçmeyi hedefler. Bu çalışmada, BIST 100 endeksi üzerindeki etkileri belirlemek için Dow Jones, NASDAQ, S&P 500 ve altın gibi fiyat endeksleri arasındaki ilişkiler çoklu regresyon analizi ile değerlendirildi. Araştırmaya, Spearman ve Pearson Korelasyon analizi kullanılarak bu değişkenler arasındaki korelasyonun incelenmesiyle başlandı. Spearman korelasyon katsayılarına dayanarak, BIST 100 ve altın arasında istatistiksel olarak anlamlı bir korelasyon saptanamadı. Ancak, BIST 100 ile Dow Jones, NASDAQ ve S&P 500 arasında güçlü ve negatif bir ilişki belirlendi. İlerleyen aşamada, regresyon analizinin görselleştirilmesi adına doğrusal ve doğrusal olmayan regresyon eğrileri çizildi. Çoklu regresyon analizi kullanılarak oluşturulan model, Dow Jones, NASDAQ, S&P 500 ve altın gibi değişkenlerin BIST 100 endeksi üzerindeki etkisi olduğu gösterildi. Modelin sonuçları, Dow Jones, NASDAQ, S&P 500 ve altın değişkenlerinin BIST 100 üzerinde negatif bir etkisinin olduğunu gösterdi. Bu bulgular istatistiksel olarak anlamlı çıktı ve güçlü bir etkiye sahip olduğu görüldü. Ancak, bir modelin doğruluğunu ve güvenilirliğini belirleme sürecinde, varsayımların kontrolünün sağlanması oldukça kritik bir aşamadır. Bu çalışmada, hata terimleri arasında otokorelasyon, heteroskedastisite ve durağanlık gibi varsayımların ihlal edildiği tespit edildi. Bu durum, modelin tahmin gücünü ve güvenilirliğini olumsuz etkileyeceğinden, gerekli koşulların sağlanması için dönüşüm, düzenlileştirme ve model seçimi teknikleri önerildi.

1 Giriş

Bu çalışmada, Türkiye'deki Borsa İstanbul 100 (Bist 100) endeksini etkileyen Dünyanın en büyük hisse senetleri olan Dow Jones, NASDAQ, S&P 500 ve altın hisse fiyatları ile analiz edilmesi amaçlanmıştır. Bu amaç doğrultusunda, Dow Jones, NASDAQ, S&P 500 ve altın hisselerin, birlikte Bist 100'e olan etkileri çok değişkenli doğrusal regresyon analizi ile incelenmiştir ve korelasyon değerleri yorumlanmıştır. Bu regresyon ve korelasyon analizinin uygulanabilmesi için bu beş endeksin 24.03.2021-24.03.2023 dönemi kapanış değerleri kullanılmıştır. Son zamanlarda Bist 100 endeksindeki uzun süreli artışının ve kısa süreli azalışının,

teorik beklentilerle uyum halinde olduğu görülmüştür. Dow Jones, NASDAQ, S&P 500 ve altın endekslerine bağlı olduğu ve bu endekslerinin birlikte, Bist 100 üzerinde istatistiksel olarak anlamlı bir etkisi olduğu bulunmuştur. Bu etki bazı endeksler için pozitifken bazıları için negatiftir. Çok değişkenli regresyon analizi ve korelasyon analizinin tutarlı sonuç verabilmesi için gerekli varsayımlar kontrol edilmiştir. İstatistiksel testler, güvenilirlik analizi ve görselleştirmeler ile bulgular yorumlanmıştır. Bu sonuçları elde edebilmek için, Bist 100, Dow Jones, NASDAQ, S&P 500 ve altın senetleri için günlük verileri içeren beş farklı veri seti, Stooq.com adresinden Excel formatında indirilmiştir. 505 gözlemden oluşan bu veriler, Date (Gün), Open (Hisse açılış fiyatı), High (Gün içindeki en yüksek fiyat), Low (Gün içindeki en düşük fiyat), Close (Hisse kapanış fiyatı), Volume (Hisseye gün içinde giren hacim) değişkenlerinden oluşmaktadır. Bu beş farklı veriden Close (Kapanış) değişkenleri seçilerek tek bir veride toplanmıştır. Ardından Date (Gün) değişkenide veriye eklenerek toplam altı değişkenli (4 bağımsız değişken (X), 1 bağımlı değişken (Y), 1 zaman değişkeni) ve 505 gözlemlili tek bir veri seti elde edilmiştir. Her endeksin kapanış değerleri üzerinden analizler gerçekleştirilmiştir.

1.1 Çalışmanın Amacı

Bist 100 endeksi, Türkiye'nin en büyük 100 şirketinin hisse senetlerinin performansını ölçen bir endekstir. Bu endeksin hareketlerinin anlaşılması ve öngörülmesi, yatırımcılar için büyük önem taşımaktadır. Ancak Bist 100 gibi diğer ülkelerin hisse senet performansları, Dow Jones, NASDAQ, S&P 500 ve altın hisse fiyatları gibi dünya genelinde önemli finansal göstergelere bağımlı hareket etmektedir. Bu finansal göstergeler diğer gelişmekte olan ülkelerin hisse senetlerini oldukça etkilemektedir. Bu durumda bu büyük hisselerin kullanıldığı çok değişkenli regresyon modeli, Bist 100 endeksinin hareketlerini görmek ve tahmin etmek için yararlı olabilir. Bu modele dahil edilen bağımsız değişkenlere, diğer borsaların hisse senetleri, dolar/TL döviz kurları, petrol endeksleri, faiz ve enflasyon oranları gibi birçok etken eklenebilir. Bu değişkenlerin her birinin Bist 100 endeksindeki hareketleri üzerindeki etkileri, çok değişkenli regresyon analizi yoluyla yorumlanabilir.

Bu analiz sonuçları, yatırımcıların Bist 100 endeksindeki muhtemel hareketleri anlamalarına ve buna göre yatırım stratejilerini ayarlamalarına yardımcı olabilir. Ayrıca, bu modelin oluşturulması, finansal piyasadaki trendleri analiz etmek ve finansal riskleri yönetmek için kullanılabilecek çok değişkenli regresyon analizinin genel kullanımı hakkında bilgi sağlayabilir.

1.2 Literatür

Dow Jones, NASDAQ, S&P 500, altın ve Bist 100 hisse fiyatları üzerinde çok değişkenli regresyon analizi uygulayarak, bu hisseler üzerinde tahmin modeli kuran oldukça fazla çalışma mevcuttur. Kao vd. (2021) makalesi altın fiyatlarını etkileyen faktörleri analiz ederek, altın fiyatlarını tahmin etmek için çok değişkenli doğrusal regresyon analizini uygular. Bu sayede, ekonomik göstergelerin altın fiyatları üzerindeki etkisini inceleyerek, gelecekteki altın fiyatlarının tahmin edilmesinde yardımcı olabilecek bir model oluşturur. Patel vd. (2020), NASDAQ hisse senedi fiyatlarını tahmin etmek için regresyon analizi ve yapay sinir ağları

yöntemlerini bir arada kullanır ve bu sayede NASDAQ fiyatlarını tahmin etmeye çalışır. Peng vd. (2020), Vijh vd. (2020), Karthikeyan vd. (2021) ve Jing vd. (2021) makaleleri ise Dow Jones, NASDAQ ve S&P 500 hisse senedi fiyatları üzerinde doğrusal regresyon analizi yaparak, gelecekteki hisse senedi fiyatlarını tahmin etmeye çalışır ve fiyatlardaki dalgalanmaların ekonomik faktörler üzerinden nedenlerini araştırır. Araştırma, Makine öğrenimi tekniklerini kullanarak, hisse senedi kapanış fiyatı tahmini için bir model oluşturur. Hasan vd. (2020) ve Yusuf (2022) makaleleri, Bist 100 endeksinin fiyat hareketlerinin tahmin edilmesi için makine öğrenmesi tekniklerinin kullanımını araştırır. Çalışmalarda, Bist 100 günlük kapanış fiyatlarının tahmini için işlem hacimleri, dolar/TL kuru, faiz oranları, altın fiyatları ve petrol fiyatları kullanılmıştır. Li vd. (2017) ve Gupta vd. (2023) makaleleri ise, hisse senedi fiyat endekslerinin tahmin edilmesi için yapay sinir ağları ve çok değişkenli regresyon modelleri arasındaki farkları incelemektedir. Bu çalışmalarda elde edilen sonuçlar, regresyon analizi yönteminin Bist 100 endeksinin tahmininde etkili bir yöntem olduğunu göstermektedir.

Bu makaleler ve literatürde yer alan diğer makaleler incelendiğinde Dow Jones, NASDAQ, S&P 500 ve altın hisse senedi fiyatları kullanılarak Bist 100 üzerinde tahmin modeli kurulan ve etkilerinin incelendiği çok az makaleye rastlanılmıştır. Dow Jones, NASDAQ, S&P 500 ve altın hisse senedi fiyatlarının bağımsız değişken olarak tutulduğunda, Bist 100 üzerinde çoklu doğrusal regresyon analizinin incelendiği bir modelin sonuçları, literatür için faydalı ve destekleyici olabilecektir. Çünkü yakın tarihte Bist 100'ün göstermiş olduğu performansın analizi, yatırımcılar için diğer etkenlere göre incelenmesi faydalı olacaktır.

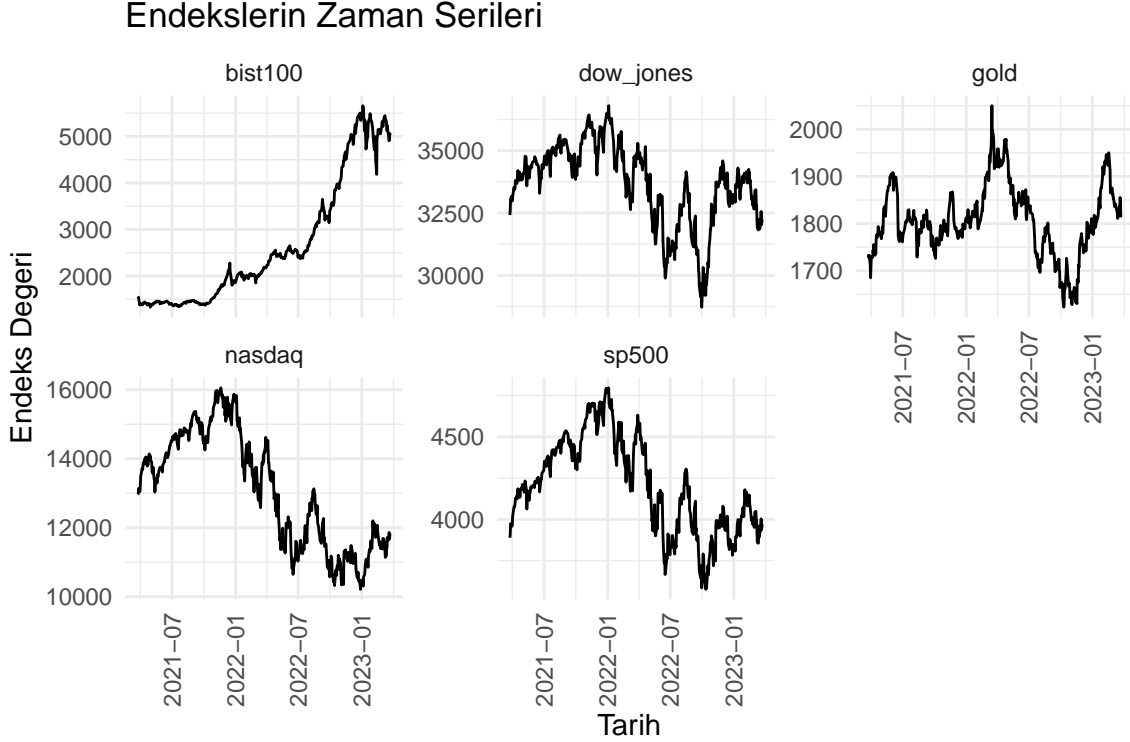
2 Veri

Analiz edilecek veri seti, <https://stooq.com> ve <https://finance.yahoo.com> sitelerinden elde edilmiş ve excel formatında bulunmaktadır. Bu veri seti, Date, Open, High, Low, Close ve Volume değişkenlerinden oluşan ve günlük toplanan beş farklı veri setini içermektedir. Tüm veriler, 24.03.2021-24.03.2023 tarihleri arasındaki dönemi kapsamaktadır. Bu verileri tek bir veri setinde birleştirmek için, her indeksin Kapanış (Close) değişkenleri R programı kullanılarak çekilmiştir. Aşağıdaki tabloda, oluşturulan veri setindeki indekslerin ortalama, standart sapma, minimum-maksimum değerlerini ve medyanını görebiliriz.

Tablo 1: Özet İstatistikler

	Ortalama	Std.Sap	Min	Medyan	Mak
bist100	2659.79	1343.52	1330.48	2175.03	5661.10
dow_jones	33678.80	1618.19	28725.51	33970.47	36799.65
gold	1806.63	75.26	1622.81	1803.51	2050.01
nasdaq	13084.99	1660.31	10213.29	13332.36	16057.44
sp500	4200.70	288.77	3577.03	4180.17	4796.56

Elde edilen betimsel istatistik sonuçlarına dayanarak indekslerin zaman serisi grafiksel görünümünü aşağıdaki şekil 1’de inceleyebiliriz.



Şekil 1: Değişkenlerin Zaman Serisi Grafikleri

3 Yöntem ve Veri Analizi

Dow Jones, NASDAQ, S&P 500 ve altın değişkenlerinin BIST 100 üzerindeki etkilerini incelemek amacıyla çoklu regresyon analizi yapılması gerekmektedir. Bu analizi uygulayıp yorumlamak için, değişkenler arasındaki korelasyonu, regresyon eğrisini ve gerekli varsayımları değerlendirmemiz ve bu konular hakkında yorum yapmamız gereklidir. İlk olarak, değişkenler arasındaki korelasyonu incelemek adına Spearman Korelasyon analizine bakalım.

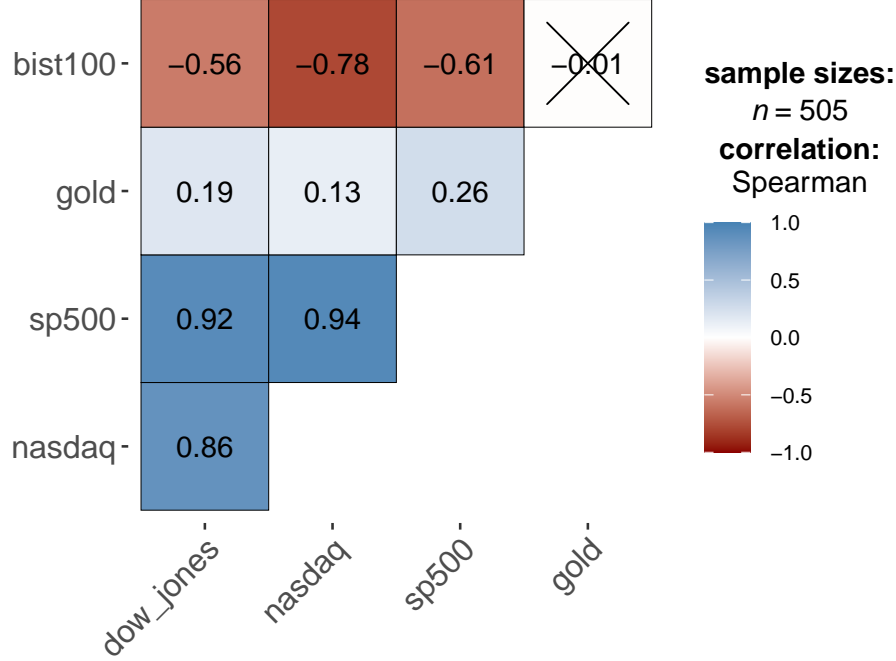
Spearman'ın sıra korelasyon katsayısı, iki değişken arasındaki monoton ilişkiyi ölçer. Bu, bir değişkendeki artışın diğer değişkendeki artışla tutarlı bir şekilde ilişkili olduğu, fakat bu artış oranının sabit olmak zorunda olmadığı anlamına gelir. Bu, korelasyonları ölçerken gerekli varsayımların yerine getirilmesi zorunluluğunu hafifletir. Pearson korelasyon katsayısı ile karşılaştırıldığında, Spearman korelasyon katsayısı değişkenlerin sıralı değerlerini kullanır ve bu da verinin normal dağılım göstermesi gerekliliğini ortadan kaldırır De Winter vd. (2016).

Spearman korelasyon katsayısı genellikle ρ (rho) sembolü ile ifade edilir ve aşağıdaki formül ile hesaplanır:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Bu formülde d_i i 'nci örneğin iki değişken arasındaki sıralama farkını temsil eder. n örneklem sayısıdır. ρ katsayısının değeri -1 ile +1 arasında değişiklik gösterir. +1 çok güçlü pozitif

korelasyonu, -1 çok güçlü negatif korelasyonu ve 0 ise ilişkinin olmadığı korelasyonu temsil eder. Değerin mutlak değeri, korelasyonun gücünü gösterir. Daha yüksek mutlak değerler, daha güçlü korelasyon anlamına gelir.



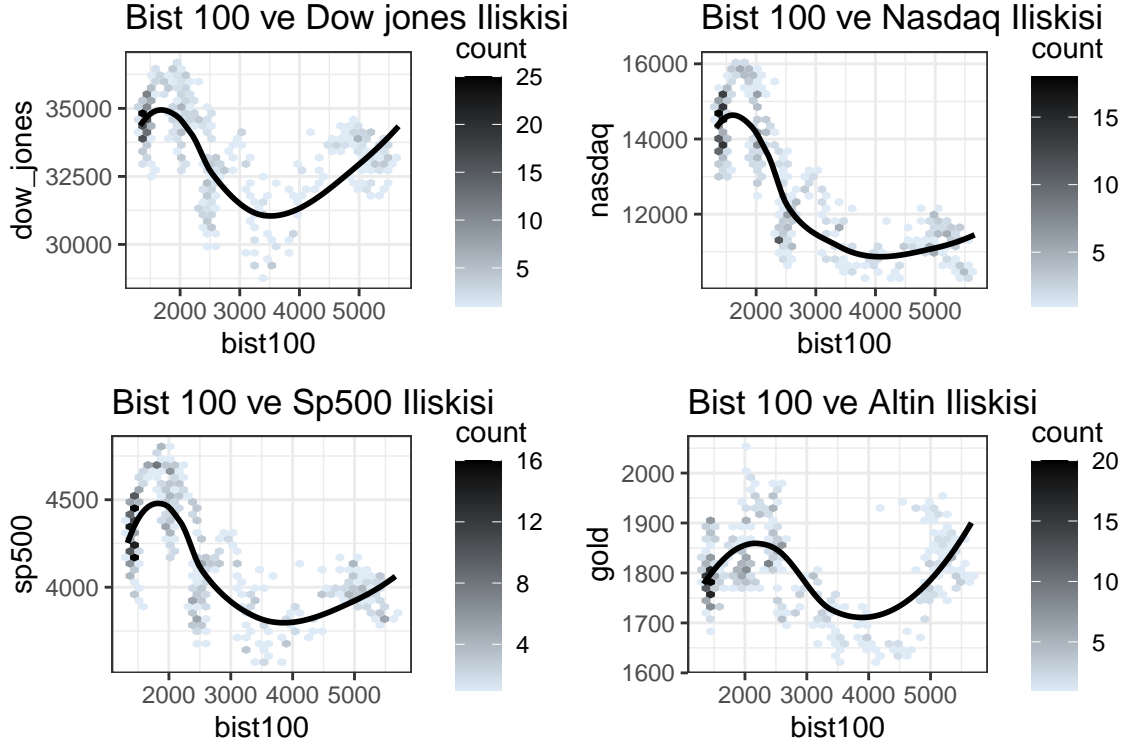
X = non-significant at $p < 0.05$ (Adjustment: Holm)

Şekil 2: Değişkenlerin Spearman Korelasyon Grafiği

Şekil 2'deki korelasyon tablosuna baktığımızda, üzerlerinde çarpı işareti bulunan ilişki değerlerinin, %95 güven düzeyinde istatistiksel anlamlılığı olmayan korelasyonları temsil ettiğini görüyoruz. BIST 100 ve altın değişkenleri arasında istatistiksel olarak anlamlı bir korelasyon tespit edilemiyor. Diğer ilişkiler incelendiğinde ise, BIST 100 ile Dow Jones, NASDAQ ve S&P 500 arasında negatif ve güçlü bir ilişki olduğu görülüyor. Özellikle NASDAQ ile BIST 100 arasındaki -0.78 değerindeki korelasyon, NASDAQ değeri arttıkça BIST 100 değerinde belirgin bir azalma olacağını gösteriyor. Grafikte en güçlü korelasyon ise NASDAQ ve S&P 500 arasında yer alıyor. Bu, korelasyonun pozitif olup, 0.94'lük bir etkileşim sağladığını ifade ediyor.

Bu korelasyonları grafiksel olarak ve doğrusal olmayan regresyon eğrisi üzerinden incelemek isterseniz, WVPlots kütüphanesinde bulunan HexBinPlot fonksiyonunu kullanarak regresyon grafiklerini çizebilirsiniz.

Şekil 3 incelendiğinde, koyu renkli alanların ilişkinin yoğun ve güçlü olduğu bölgeleri temsil ettiğini görüyoruz. Korelasyon tablosunda, altın ile BIST 100 arasında belirgin bir korelasyon belirlenmemiştir. Ancak, grafiksel analiz sonucunda, BIST 100 değeri 3500'e ulaştığında altın fiyatlarında bir düşüş, 6000'e ulaştığında ise altın fiyatlarında bir artış görülüyor. Genel bir bakışla, altın ve BIST 100 arasında belirli bir simetri olduğunu gözlemleyebiliyoruz. Ancak bu durum, belirgin bir korelasyonun olduğu anlamına gelmez. Dolayısıyla, iki değişken



Şekil 3: Değişkenler Arası Doğrusal Olmayan Regresyon Grafiği

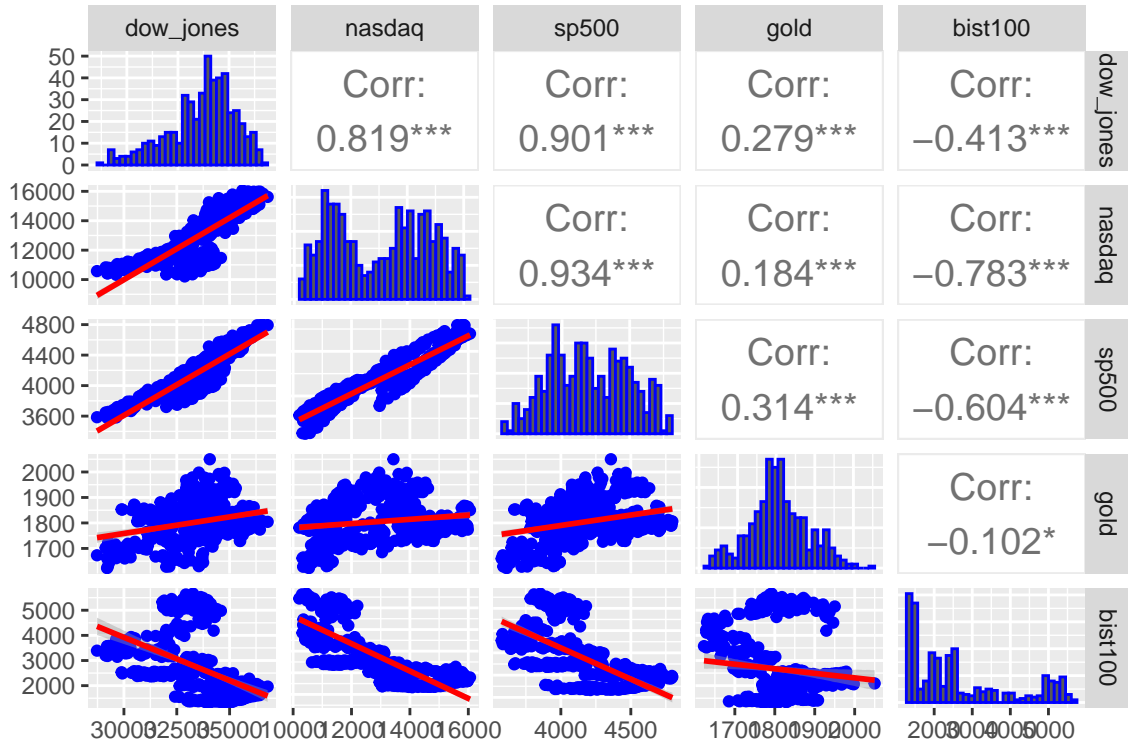
arasında anlamlı bir korelasyon olduğunu söylemek doğru olmaz. Ancak, farklı bir korelasyon analizi olan Pearson korelasyon analizini denemekte fayda olabilir. Bu korelasyon türü, normallik varsayımı yerine getirildiğinde kullanışlıdır. Ayrıca, bu korelasyon ilişkisi doğrusal regresyon eğrisi üzerinden de yorumlanabilir.

Pearson korelasyon katsayısı, r , aşağıdaki formülle hesaplanır:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

Bu formülde x_i ve y_i veri setlerindeki gözlem sayılarıdır. Yani, x_i x veri setinin i 'inci gözlemi ve y_i y veri setinin i 'inci gözlemidir. \bar{x} ve \bar{y} ise ilgili veri setlerinin ortalamalarını temsil eder. n , toplam örneklem sayısını temsil eder.

Şekil 4'te ise, her değişkenin Pearson korelasyon ilişkisini, dağılımını ve regresyon eğrilerini görebiliriz. Altın ve BİST 100 arasındaki anlamlı olmayan Spearman korelasyon ilişkisi, bu Pearson korelasyon ilişkisinde %90 güven anlamlı çıkmaktadır. Bu korelasyon ise 0.10'luk bir negatif etki yaratmaktadır. Özetle her bağımsız değişken BİST 100 üzerinde negatif ve anlamlı bir ilişkiye sahip olduğunu söyleyebiliriz. Ayrıca değişkenler arasındaki doğrusal regresyon eğrilerinin yönleri hakkında fikir sahibi olabiliriz. Her bağımsız değişkenin doğrusal regresyon eğrisi BIST 100 üzerinde negatif yönlü olduğu görülmektedir. Bu koşullar altında çok değişkenli regresyon analizi uygulanabilir.



Şekil 4: Değişkenler Arası Doğrusal Regresyon Grafiği, Pearson Korelasyon Değerleri ve Dağılımları

Çok değişkenli regresyon analizi, birden fazla bağımsız değişkenin bir bağımlı değişken üzerindeki etkisini inceleyen bir istatistiksel tekniktir. Çok değişkenli regresyon modeli genel formunda aşağıdaki gibidir:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon$$

Bu modelde y bağımlı değişkendir ve hedef değişken olarak tanımlanır. β_0 y -kestirimidir, tüm bağımsız değişkenler sıfır olduğunda beklenen Y değeridir. $\beta_1, \beta_2, \dots, \beta_p$ bağımsız değişkenlerin katsayılarıdır ve bu değişkenlerin bağımlı değişken üzerindeki etkisini belirtirler. x_1, x_2, \dots, x_p bağımsız değişkenlerdir. ε hata terimi olup, modelin tahminleri ile gerçek değerler arasındaki farkı ifade eder.

Çok değişkenli regresyon analizi, genellikle en küçük kareler yöntemi (OLS - Ordinary Least Squares) kullanılarak hesaplanır. Bu yöntem, hata terimi ε 'in kareler toplamını minimize etmeye çalışır. Bu, aşağıdaki formülle ifade edilir:

$$\min_{\beta_0, \beta_1, \dots, \beta_p} \sum_{i=1}^n (y_i - (\beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi}))^2$$

Bu denklemde, $\beta_0, \beta_1, \dots, \beta_p$ katsayılarını bulmak için karelerin toplamını minimize ederiz. Bu sonuçlar ışığında bist 100 değişkenine olan etkinin çok değişkenli regresyon modeli kurulabilir. Bu modelin sonuçları aşağıda incelenmektedir.

```
##
## Call:
## lm(formula = bist100 ~ dow_jones + nasdaq + sp500 + gold, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1315.15  -442.88   -87.79   321.14  2018.53
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -933.37549   858.04121  -1.088    0.277
## dow_jones     0.40284    0.03909   10.306 < 2e-16 ***
## nasdaq       -1.39989    0.04891  -28.621 < 2e-16 ***
## sp500         2.83801    0.37501    7.568 1.84e-13 ***
## gold         -1.98048    0.40088   -4.940 1.07e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 607.7 on 500 degrees of freedom
## Multiple R-squared:  0.797, Adjusted R-squared:  0.7954
## F-statistic: 490.9 on 4 and 500 DF, p-value: < 2.2e-16
```


$$Bist100 = -933.37549 + 0.40284DowJones - 1.39989Nasdaq + 2.83801SP500 - 1.98048Altın$$

Bu sonuçlara bakarak, katsayılar tablosu modelin regresyon katsayılarını içermektedir. Her bir katsayı, diğer tüm değişkenlerin sabit tutulduğunda, ele alınan bağımsız değişkenin bağımlı değişkene etkisini temsil eder.

Intercept (Y kestirimi) yaklaşık -933.37549. Bu değer, bütün bağımsız değişkenlerin değeri 0 olduğunda beklenen 'bist100' değeridir. Ancak, bu yorum, bağımsız değişkenlerin gerçekten 0 olabileceği durumlarda geçerlidir. Dow jones, Nasdaq, Sp500 ve Gold değişkenlerinin katsayıları sırasıyla yaklaşık 0.40284, -1.39989, 2.83801 ve -1.98048' olarak elde edilmiştir. Bu değerler, diğer tüm değişkenler sabit tutulduğunda, ilgili bağımsız değişkenin bir birimlik değişiminin Bist100 üzerindeki etkisini gösterir. Örneğin Dow Jones değişkeni dışında diğer tüm değişkenler sabit tutulduğunda Dow Jones değişkeninin 1 birimlik artışı bist100 değişkenini 0.40284 birim arttırmaktadır. Bu etkilerin istatistiksel olarak anlamlılığını kontrol etmek için p anlamlılık değerleri incelenir. $Pr(>|t|)$ değerleri, bağımsız değişkenin katsayısının istatistiksel olarak anlamlı olup olmadığını gösterir. Bu durumda, tüm bağımsız değişkenler (Dow jones, Nasdaq, Sp500 ve Gold) %95 güven düzeyinde istatistiksel olarak anlamlıdır çünkü p-değerleri 0.05 değerinin altındadır.

Residual standard error değeri modelin standart hatasıdır ve modelin verilere ne kadar iyi uyduğunu gösterir. Bu örnekte, hata terimi 607.7'dir. Multiple R-squared değeri ise, modelin bağımlı değişkenin varyansını ne ölçüde açıkladığını gösteren R kare değeridir. Bu model, bist100 değişkeninin varyansının yaklaşık %79.7'sini açıkladığını söylemektedir. Tüm bağımsız değişkenler Bist100 değişkeninin üzerindeki değişimin %79.7'sini açıklamaktadır.

Son olarak F-statistic ve p-value değerleri, yani F istatistiği ve ilgili p-değeri, modelin bütünüyle anlamlı olup olmadığını test eder. Bu durumda, p-değerinin 0.05'in çok altında olması modelin istatistiksel olarak %95 güven düzeyinde anlamlı olduğunu göstermektedir.

Özetle, oluşturulan model iyi bir uyum göstermektedir ve bağımsız değişkenlerin Bist100 değerleri üzerinde önemli bir etkiye sahip görünmektedir. Ancak, modelin doğruluğunu doğrulamak için daha fazla tanı testi yapmamız gerekmektedir. Örneğin, hataların eşvaryansı ve otokorelasyonu, çoklu bağlantı sorunu gibi. Bu varsayımların ihlali durumunda, modelin tahminleri yanıltıcı olabilir ve yanlış sonuçlara neden olabilir. Dolayısıyla, varsayımları test etmek ve modelin uygunluğunu değerlendirmek önemlidir. Özellikle zaman serisi verileri söz konusu olduğunda, otokorelasyon ve durağanlık önemli hususlardır.

3.1 Varsayımlar

Gerekli varsayımlardan birisi otokorelasyon durumu, yani hata terimlerinin bağımsızlığıdır. Otokorelasyon, zaman serisi analizinde önemli bir konudur, çünkü varsayımlardan biri hataların bağımsız olduğudur. Otokorelasyon varsa, bu varsayım ihlal edilir ve modelin tahmin gücü, standart hatalarının doğru hesaplanması ve istatistiksel çıkarımların güvenilirliği etkilenebilir.

Eğer otokorelasyon varsa (yani hatalar birbiriyle ilişkiliyse), modelin hata terimleri yanıltıcı olabilir ve standart hataların altında tahmin edilebilir. Bu durum, t istatistiklerinin ve dolayısıyla p-değerlerinin yanıltıcı olmasına yol açabilir Osborne ve Waters (2002).

Tablo 2: Durbin-Watson Test Sonucu

	Otokorelasyon	İstatistik	p.Değeri	lag	hipotez
1	0.96	0.06	0.00	1.00	rho != 0

Durbin-Watson testinin çıktısı hatalar arasındaki otokorelasyonu ölçer. Lag 1'deki otokorelasyon 0.9593232, bu oldukça yüksek ve pozitif bir otokorelasyonu gösterir. Durbin-Watson istatistiği ise 0.06276502 olduğundan, otokorelasyon problemine sahip olduğunu gösterir çünkü Durbin-Watson istatistiği otokorelasyon olmaması için 2 değerine yakın olması gerekmektedir.

p-değeri ise 0 olup, bu da otokorelasyonun var olduğunu (yani hataların bağımsız olmadığını) gösterir. Bu, otokorelasyon varlığını reddetmek için kullanılan hipotezin (hatalar arasında otokorelasyon yok) reddedildiğini gösterir.

Bu durumda, modelinizin hatalarında yüksek bir otokorelasyon olduğunu söyleyebiliriz. Bu genellikle, modelinizin belirli bir trendi veya döngüyü yakalamadığı veya bazı önemli değişkenlerin eksik olduğu anlamına gelir. Bu durum, modelin tahminlerinin güvenilirliğini azaltabilir ve genellikle modelin düzeltilmesi veya iyileştirilmesi gerektiğini gösterir.

Ayrıca şekil 5'te de görüleceği üzere hataların sabit varyanslı olmadığını yani homoscedasticity probleminin olduğu görülmektedir. Sonuçların doğrusal olmadığı görülmektedir.

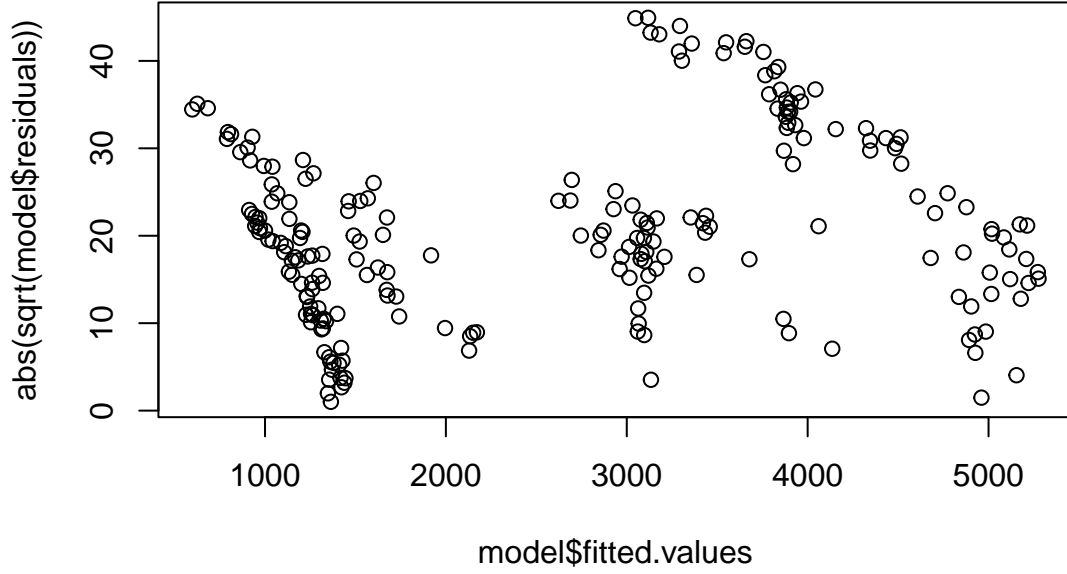
Diğer bir varsayım durağanlık durumudur. Zaman serisi verisi genellikle durağan olmalıdır, yani zamanla değişmeyen bir ortalaması ve varyansı olmalıdır. Eğer durağan değilse (örneğin, bir trend veya mevsimsellik varsa), modelin tahminleri yanıltıcı olabilir. Augmented Dickey-Fuller testi veya KPSS testi, durağanlık varsayımını kontrol etmek için kullanılabilir.

Tablo 3: Dickey-Fuller Test Sonucu

	Değişken	Dickey.Fuller.Test.İstatistiği	P.Değeri
1	dow_jones	-2.83	0.23
2	nasdaq	-2.35	0.43
3	sp500	-2.63	0.31
4	gold	-2.50	0.37
5	bist100	-1.87	0.63

ADF testi sonuçlarına göre, her bir değişkenin p-değerleri 0.05 anlamlılık düzeyinde durağanlık hipotezini desteklememektedir. Bu durumda, değişkenlerin zaman serileri durağan olmayabilir ve zaman içindeki değişikliklere sahip olabilirler.

Son olarak çoklu regresyon analizinde, Varyans Inflation Factor (VIF) değerleri bağımsız değişkenler arasındaki çoklu doğrusal ilişkiye (multicollinearity) bakmamız gerekmektedir.



Şekil 5: Residual vs Fitted Grafiği

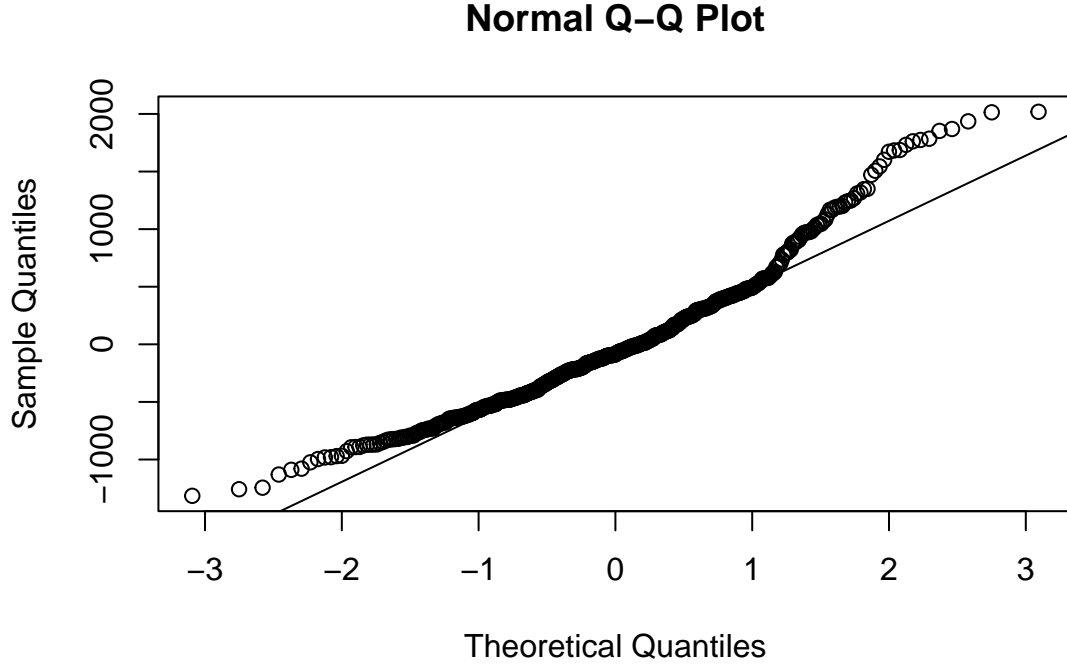
VIF değeri 5 veya 10'un üzerindeyse, genellikle bu, bağımsız değişkenler arasında yüksek düzeyde çoklu doğrusallık olduğu anlamına gelir Daoud (2017).

Tablo 4: VIF Sonuçları

	Değişken	VIF
1	dow_jones	5.46
2	nasdaq	9.00
3	sp500	16.01
4	gold	1.24

VIF değerlerinin 5'ten büyük olması çoklu bağlantı sorunun olduğunu göstermektedir. Sp500 ve nasdaq değişkenleri arasında ciddi bir çoklu bağlantı sorunu olduğu görülmektedir. Bu durumda, regresyon analizinde bu değişkenlerin birlikte kullanılması güvenilir sonuçlar üretmeyebilir. Çoklu bağlantı sorununu çözmek için bu değişkenlerden birini çıkarmak veya dönüşümler uygulamak gibi yöntemler kullanılabilir. Bu durum, bu iki değişkenin çok güçlü bir korelasyona sahip olması durumuna bağlı olabilir.

Son olarak şekil 6 ile hata terimlerinin normal dağılımdan geldiği varsayımı kontrol edilebilir. Bu grafik, noktaların doğrusal çizginin dışına çıkması ve doğrusallığı net bir şekilde sağlayamaması sonucunda hata terimlerinin normal dağılımdan gelmediği sonucunu ortaya koymaktadır.



Şekil 6: Normal Q-Q Grafiği

Yukarıdaki istatistiksel sonuçlar, analizler ve grafikler, çoklu değişkenli regresyon analizi ve zaman serisi verileri için gerekli varsayımlar hakkında bilgileri sağlar. Ancak, daha karmaşık modeller ve senaryolar için ek testler ve modelleme teknikleri gerekebilir.

Bu durumda, regresyon modelini analiz ederken varsayımlar konusunu göz önünde bulundurmanız önemlidir. Yukarıda bahsi geçen varsayımların sağlanmaması model yapısını oldukça etkiler. Her ne kadar modelin anlamlılığı, etki değerleri ve korelasyon değerleri anlamlı çıksada, varsayımların sağlanmaması bu sonuçların yorumlanmasını oldukça zorlaştırır ve güvenilir sonuçlar üretmemizi etkiler.

Varsayım eksikliği, modelinizin tahminlerinin güvenilirliğini ve geçerliliğini etkilediğinden varsayımları sağlamak için dönemsel farklar, dönemsel bileşenler veya diğer istatistiksel dönüşümleri (standardize etme, logaritmik fark alma) kullanarak verileri dönüştürebilir veya alternatif modeller ve yöntemler kullanılabilir. Örneğin Lasso, Ridge veya Elastic Net gibi regresyon yöntemlerini kullanarak modele cezalandırma parametresi eklemek ve regresyon sonuçlarını elde etmek oldukça faydalı olacaktır.

4 Sonuç

Çoklu regresyon analizi sonucumuzda, BIST 100 endeksinin üzerinde etkili olan Dow Jones, NASDAQ, S&P 500 ve altın fiyatları değişkenlerini inceledik. Her bir değişkenin BIST 100 üzerindeki etkisi, modelimizin regresyon katsayıları ile belirlendi. Regresyon denklemine göre

NASDAQ ve Altın fiyatlarının etkisi negatifken Dow Jones ve S&P 500 fiyatları pozitif etki yaratmaktadır. Ayrıca Spearman ve Pearson korelasyon değerlerine göre her bağımsız değişken BIST 100 üzerinde negatif etki yaratmaktadır. Bu sonuçlar istatistiksel olarak anlamlı bulunmaktadır. Bu bulgular, modelimizin BIST 100 endeksinin davranışını öngörme kabiliyeti üzerine bir dizi etkisi olan değişkenleri yakaladığını göstermektedir. Ancak, modelin tahmin doğruluğunu artırmak ve daha güvenilir sonuçlar almak için, belirli varsayımların ihlal edilmesiyle ilgili sorunların ele alınması ve çözülmesi gerekmektedir. Öncelikle, hatalarımızın bağımsız olduğu varsayımı otokorelasyon sorunu nedeniyle ihlal edildi. Ayrıca, hataların sabit bir varyansa sahip olması gereken homoscedasticity varsayımını da ihlal edildi. Son olarak, bağımsız değişkenler arasında yüksek bir çoklu doğrusallık sorunu tespit ettik. Bu durumlar, modelin tahmin gücünü etkileyebilir çünkü modelin her bir bağımsız değişkenin BIST 100 üzerindeki etkisini tek başına belirlemesi zorlaşabilir.

Tüm bu faktörler, analizimizin sonuçlarının yorumlanmasında dikkatli olmamız gerektiğini göstermektedir. Otokorelasyon, heteroscedasticity ve çoklu doğrusallık gibi varsayım ihlalleri, modelimizin tahminlerinin güvenilirliğini etkileyebilir ve modelin doğru bir şekilde kullanılabilmesi için bu sorunların düzeltilmesi gerekebilir. Modelin iyileştirilmesi için, gelişmiş istatistiksel teknikler kullanabiliriz. Ayrıca, bu sorunları çözmek için değişken seçim teknikleri, düzenleme (regularization) ve dönüşüm (logaritmik, standardize) teknikleri uygulayabiliriz. Son olarak cezalandırma yöntemleri olan regresyon modelleri kullanılabilir. Bir finansal varlığın yatırım kararı verirken fiyatını tahmin etmek ve risk ile beklenen getiri arasındaki ilişkiyi ölçmek, günümüz finans ve bilişim sektöründe popüler hale geldiğinden bu modeller üzerinde geliştirme yapmak ve daha başarılı ve güvenilir analizler gerçekleştirmek finans, akademik ve sektörel çalışan kimselerin daha başarılı sonuçlar elde etmesine yardımcı olacaktır.

5 Kaynakça

- Daoud, J. I. (2017). Multicollinearity and regression analysis. *Journal of Physics: Conference Series* içinde (C. 949, s. 012009). IOP Publishing.
- De Winter, J. C., Gosling, S. D. ve Potter, J. (2016). Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychological methods*, 21(3), 273.
- Gupta, R., Nel, J. ve Pierdzioch, C. (2023). Investor confidence and forecastability of US stock market realized volatility: evidence from machine learning. *Journal of Behavioral Finance*, 24(1), 111-122.
- Hasan, A., Kalipsız, O. ve Akyokuş, S. (2020). Modeling traders' behavior with deep learning and machine learning methods: evidence from BIST 100 index. *Complexity*, 2020, 1-16.
- Jing, N., Wu, Z. ve Wang, H. (2021). A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction. *Expert Systems with Applications*, 178, 115019.
- Kao, W.-T., Hsu, S.-T., Chou, F. H.-C., Chou, L.-S., Hsieh, K.-Y., Li, D.-J., ... Huang, J.-J. (2021). The societal influences and quality of life among healthcare team members during the COVID-19 pandemic. *Frontiers in Psychiatry*, 12, 706443.
- Karthikeyan, C., Anandan, P., Prabha, R., Mohan, D., Babu, V., vd. (2021). Predicting Stock Prices Using Machine Learning Techniques. *2021 6th International Conference on Inventive Computation Technologies (ICICT)* içinde (ss. 1-5). IEEE.
- Li, J., Bu, H. ve Wu, J. (2017). Sentiment-aware stock market prediction: A deep learning method. *2017 international conference on service systems and service management* içinde (ss. 1-6). IEEE.
- Osborne, J. W. ve Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical assessment, research, and evaluation*, 8(1), 2.
- Patel, K., Mehta, D., Mistry, C., Gupta, R., Tanwar, S., Kumar, N. ve Alazab, M. (2020). Facial sentiment analysis using AI techniques: state-of-the-art, taxonomies, and challenges. *IEEE Access*, 8, 90495-90519.
- Peng, Y., Chen, W., Wei, P. ve Yu, G. (2020). Spillover effect and Granger causality investigation between China's stock market and international oil market: A dynamic multiscale approach. *Journal of Computational and Applied Mathematics*, 367, 112460.
- Vijh, M., Chandola, D., Tikkiwal, V. A. ve Kumar, A. (2020). Stock closing price prediction using machine learning techniques. *Procedia computer science*, 167, 599-606.
- Yusuf, A. (2022). Analysis of price volatility in BIST 100 index with time series: comparison of Fbprophet and LSTM model. *Avrupa Bilim ve Teknoloji Dergisi*, (35), 89-93.