

Twitter ile Hisse Senetleri Oynaklığı Tahmin Edilebilir mi?

Gürkan BOZMA (<https://orcid.org/0000-0003-4047-9012>), Department of Economics, İğdır University, Turkey;
e-mail: gurkan.bozma@igdir.edu.tr

Sinan KUL (<https://orcid.org/0000-0002-7824-756X>), Department of Computer Using, Atatürk University, Turkey; e-mail: sinan.kul@atauni.edu.tr

Can Twitter Forecast Uncertainty of Stocks?

Abstract

Academic studies have shown that there is a relationship between emotional analysis results of tweets and stock price movements, and then stock prices can be estimated using this relationship. In this study, in which the effect of tweets on the volatility of the stock is estimated, the volatility scores and the emotion scores between the stocks were also revealed. In the scope of the study, sentiment analysis with Naive Bayes was performed on Turkish tweets shared by three phone companies (Alcatel, Turkcell and Vestel) which are in Borsa Istanbul and whose products are sold in Turkey. According to the results of the analysis, it was found that sentiment scores obtained for Turkcell and Vestel significantly increased Alcatel's conditional variance statistically.

Keywords : Sentiment Analysis, Turkish Tweet, Natural Language Processing, Data Mining, Text Mining, Stock Market Prediction, Machine Learning, Naive Bayes, BEKK-GARCH-X.

JEL Classification Codes : C11, C32, D91.

Öz

Tweetlerin duygu analizi ile hisse senedi fiyat hareketleri arasında ilişki olduğu hatta hisse senedi fiyatlarının bu yolla tahmin edilebildiği yapılan çalışmalarla ortaya konmuştur. Tweetlerin, hisse senetlerinin oynaklığa etkisinin tahmin edilmesi üzerine yapılan bu çalışmada ayrıca, hisse senetlerinin birbirleri arasındaki ve duygu puanlarının hisse senetleri arasındaki oynaklık geçişkenlikleri de ortaya çıkarılmıştır. Çalışma kapsamında, ürünlerinin Türkiye’de satışı yapılan ve Borsa İstanbul’da faaliyet gösteren 3 telefon markası (Alcatel, Turkcell ve Vestel) hakkında paylaşılan Türkçe tweetler üzerinde Naive Bayes ile duygu analizi yapılmıştır. Analiz sonuçlarına göre Turkcell ve Vestel için elde edilen duygu puanlarının Alcatel’in koşullu varyansını istatistiki olarak anlamlı bir şekilde artırdığı tespit edilmiştir.

Anahtar Sözcükler : Duygu Analizi, Türkçe Tweet, Doğal Dil İşleme, Veri Madenciliği, Metin Madenciliği, Borsa Tahmini, Makine Öğrenmesi, Naive Bayes, BEKK-GARCH-X.

1. Giriş

1950'lerin başından günümüze kadar davranışsal finans kapsamında karar vericilerin yatırım kararlarını etkileyen faktörler üzerine birçok çalışma yapılmıştır (Thaler, 1991, 1993; Wood, 1995; Olsen, 1998; Ricciardi & Simon, 2000, Shefrin, 2002; Barberis & Thaler, 2003; Shiller, 2003; Chan vd., 2004; Baker, 2010). Bu çalışmalarda elde edilen sonuçlara göre bireylerin duyguları, karar verme davranışlarını etkileyebilmektedir.

Diğer taraftan, Malkiel ve Fama (1970)'nin ortaya koyduğu etkin piyasa hipotezine göre fiyatlar, erişebildiği her bilgiye tepki vermektedir. Buradan sosyal medyanın fiyatlar üzerinde etkisi olduğunu ifade etmek yanlış olmayacaktır. Son yıllarda özellikle Facebook, Twitter ve Google trend verilerini kullanarak finansal varlıkların getirilerini ve oynaklıklarını tahmin eden çalışmalar literatürde yaygınlaşmaktadır (Kaminski & Gloor, 2014; Mao vd., 2012; Mittal & Goel, 2012; Rao & Srivastava, 2012; Zhang vd., 2011; Ye & Wu, 2010; Bollen vd., 2011, Nisar & Yeung, 2018). Örneğin Liu vd. (2015) ve Jin et vd. (2016) tarafından yapılan çalışmalarda sosyal medyanın borsa hacmi ve getirileri üzerindeki etkileri incelenmiş ve sosyal medyanın borsa hacmini artırırken, getirileri azalttığı tespit edilmiştir. Zhang vd. (2011), atılan tweetlerin içerdiği duygulara göre Dow Jones, Nasdaq ve S&P 500 gibi indeksler üzerindeki etkisini araştırmışlardır. Elde edilen bulgular sonucunda, atılan tweetlerin çok fazla umut, korku ve endişe içermesi durumunda Dow Jones endeksinin düştüğünü, daha az umut, korku ve endişe içermesi durumunda ise Dow Jones endeksinin yükseldiğini ifade etmişlerdir.

Twitter verilerini kullanarak borsa fiyatlarını tahmin eden çalışmalar incelendiğinde Bollen vd. (2011), Bollen ve Mao (2011) ve Vu vd. (2012) tarafından yapılan çalışmalar görülmektedir. Oliveira vd. (2017) ise 2012'nin Aralık ayından 2015'in Ekim ayına kadar olan dönemde 3800 hisse senediyle alakalı olarak atılan 31 milyon adet tweeti kullanarak hisse senetleri üzerine tahminler yapmışlardır. Dolayısıyla sosyal medya verilerinin analizinde kullanılan en yaygın yöntemlerden birinin duygu analizi olduğu ifade edilebilir.

Herhangi bir konuda paylaşılan görüşlerin hangi duyguyu yansıttığının tespitine odaklanan duygu analiziyle içerikler, bilgisayar yazılımları marifetiyle otomatik olarak pozitif ve negatif gibi sınıflara atanmaktadır. Duygu analizi çalışmaları, kullanılan yöntemlere göre sözlük tabanlı yaklaşım ve istatistiksel (makine öğrenmesi) yaklaşımlar olmak üzere iki ana yöntem etrafında şekillenmiştir.

Sözlük tabanlı yaklaşımda, önceden duygu ağırlıkları belirlenmiş kelimelerin bulunduğu bir tablodan faydalanılmaktadır. Makine öğrenmesi yönteminde ise metinlerden bir kısmını elle etiketleyerek eğitim veri setinin oluşturulmasına yardımcı bir yöntem olarak duygu simgeleri (Çoban & Özyer, 2015) ve duygu belirleyici ifadeler (Claster vd., 2010) de kullanılabilir.

Türkçe metinlerin çözülmesi ve işlenmesi, Türkçenin sondan eklemeli dil yapısı dolayısıyla çeşitli zorluklar içermektedir. Türkçenin yapısından kaynaklanan zorluğa ek olarak sosyal medya yazışmalarındaki kurlsız veri girişi dolayısıyla sosyal medyada

kendine özgü jargonun gelişmesi ve sosyal medyanın imla/dilbilgisi açısından zayıf metinler barındırması da eklenince duygu analizi çalışmalarında Türkçeye daha az yer verilmektedir (Türkmenoglu & Tantug, 2014). Diğer taraftan İngilizcenin yaygın kullanımı dolayısıyla duygu analizi çalışmalarının, çoğunlukla İngilizce metinler üzerinde yapıldığı görülmektedir.

Twitter ile hisse senedi fiyatları arasındaki ilişkiyi araştıran çalışmalardan birinde (Wei vd., 2016), paylaşılan tweet sayısı ile hisse senedi fiyat hareketleri arasında kısa dönemde doğrudan bir ilişki olduğu tespit edilebilmiştir. Başka bir çalışma (Aich vd., 2017) ise Samsung firmasının 5 günlük tweetlerini (200 bin tweet) analiz ederek tweetler ile hisse senedi fiyatları arasında korelasyon olduğunu tespit etmiştir. Twitter verilerinin duygu analizi puanlarıyla hisse senedi fiyatları arasındaki ilişkiyi araştıran Şimşek ve Özdemir (2012) ise bu korelasyonu %45 olarak tespit ederken; Pagolu vd. (2016) tarafından bu ilişki %71,8 olarak gözlemlenmiştir.

Attigeri vd. (2015) ise tweet analizine (2 firma hakkında paylaşılan 1029 tweet), 865 haber makalesini dâhil ederek hisse senedi fiyatlarıyla arasında (lojistik regresyon yöntemi kullanarak) %70'lik bir ilişki olduğunu tespit etmiştir.

Tweetlerin duygu analizi ile hisse senedi fiyat hareketlerini tahmin etmeye odaklanan bazı çalışmalar (Si vd., 2014; Bollen vd., 2011) borsanın Twitter'den etkilendiğini tespit etmiştir. Qasem vd. (2015), dört firmanın (Twitter, Google, Facebook ve Tesla) Twitter verilerini (42 bin tweet) analiz ederek hisse senedi fiyatlarını %58 doğrulukla tahmin edebilmiştir. New York borsasındaki 30 şirket üzerinde yaptıkları çalışmalarında Bing vd. (2014) ise bazı şirketlerin hisse senedi fiyatlarının %76,12 gibi yüksek bir ortalama doğrulukla tahmin edilmesinin mümkün olduğunu bulmuştur. Cakra ve Trisedya (2015) ise Endonezya'daki 13 firma üzerine yaptıkları çalışmada lineer regresyon kullanarak Dow Jones Borsası Endüstri Endeksini %87,6 doğrulukla tahmin edebilmiştir.

Skuza ve Romanowski (2015) ise Apple firmasının New York borsasındaki hisse senedi fiyatlarını, 300 bin tweet analiz ederek, 5, 15, 30 ve 60 dakikalık dönemler için tahmin etmeye çalışmış; 5 dakikada biriken tweet miktarının azlığı dolayısıyla tahminlerin daha zor olduğunu belirlemiştir. Hu vd. (2014), Elon Musk gibi önemli Twitter kullanıcıları tarafından atılan tweetleri analiz ederek aynı işlem günündeki hisse senedi fiyat hareketlerini mukayese ettikleri çalışmalarında kurdukları modelin zaman serisi etkisini 1 günden 5 güne kadar ölçmüşler ve en iyi eşleşmenin (%60 doğruluk) 3 gün için olduğunu bulmuşlar. Nofer vd. (2015), 100 milyon tweet üzerinden yapılan analize (Twitter takipçilerinin sayısının ruhsal bulaşıcılığı göz önünde bulundurarak) Twitter takipçilerinin sayılarını dahil ettiklerinde anlamlı bir ilişki olduğunu görmüşlerdir. Ağırlıklandırılmış sosyal duygu endeksini hesaplarken de takipçi sayılarını da dikkate almışlardır.

Bu çalışmada, Türkiye'de satışı yapılan ve Borsa İstanbul'da faaliyet gösteren firmalar tarafından üretilen telefon markaları (Alcatel, Turkcell ve Vestel) hakkında paylaşılan 6 aylık Twitter metinleri üzerinde duygu analizi yapılmıştır. Naive Bayes yöntemi (makine öğrenmesi ile sınıflandırma yöntemlerinden biri) ile sınıflandırma öncesinde eğitim

veri setinin oluşturulmasında elle etiketlenen tweetlerden, duygu belirleyici ifadelerden ve duygu simgelerinden faydalanılmıştır. Türkçe kelimelerin ek ve köklerine ayrılmasında ve yazım yanlışlığı bulunan kelimelerin düzeltilmesinde yaygın kullanımı dolayısıyla Zemberek kütüphanesinden faydalanılmıştır. Çalışma, literatüre birkaç yoldan katkı sağlamaktadır. Öncelikle, literatürde tweetlerden yararlanılarak hisse senetlerindeki oynaklığın tahmin edilmesi üzerine yapılan çalışmaların sınırlı sayıda olmasından dolayı literatürdeki boşluk giderilmeye çalışılmıştır. İkinci olarak, hisse senetlerinin birbirleri arasındaki oynaklık geçişkenlikleri ortaya çıkarılmıştır. Son olarak kullanılan ekonometrik yöntem gereği, hisse senetleri için elde edilen duygu puanlarının hisse senetleri arasındaki oynaklığa etkisi ölçülmüştür.

2. Duygu Analizi

Çalışma kapsamında Türkiye'deki e-ticaret sitelerinde satışı yapılan telefon markaları hakkında paylaşılan ve Kul (2018) tarafından toplanan Türkçe tweetler (2017, Şubat - Temmuz arası dönem) içinden Alcatel, Turkcell ve Vestel isimlerinin geçtiği tweetler (154176 adet) filtrelenmiştir.

Binlerce tweet okunduktan sonra, herhangi bir tweetin hangi duygu sınıfına ait olduğunu doğrudan belirlemeye yardımcı olan kelime veya kelime gruplarının bir listesi hazırlanmıştır. Bu listeye, Simşek ve Özdemir (2012) tarafından hazırlanan 113 duygu belirleyici ifade ve sosyal medya yazışmalarında sık kullanılan ve tweetler içinde duyguyu belirlemede yardımcı olan simgeler (☺, ☹ vb.) de eklendiğinde yaklaşık 1800 adet ifade listesi oluşturulmuştur.

Eğitim veri seti oluşturulurken öncelikle el yordamıyla okunarak işaretlenen yaklaşık 2000 tweet kullanılmıştır. Sonrasında duygu belirleyici ifadeler ve duygu simgeleri kullanılarak sınıflandırılabilen tweetler eğitim veri setine dâhil edilmiştir. İfadelerin, sınıflandırma başarısı Weka programında Destek Vektör Makinesi kullanılarak test edilmiş ve tweetlerin %84 başarıyla sınıflandırılabildiği görülmüştür.

3. Her Bir Markanın Günlük Duygu Puanının Hesaplanması

Her bir marka hakkında paylaşılan olumlu ve olumsuz tweet sayılarının gün bazında hesaplanarak birbirinden çıkarılması gerçek duygu ağırlığını verememektedir. Bazı tweetler yaygın olarak paylaşılıp okunurken bazı tweetler atıl kalabilmektedir. Bundan dolayı markaların günlük sosyal medya skorları hesaplanırken, tweetlerin okunma sayılarını hesaplamaya katmanın daha doğru ve gerçekçi sonuç vereceği, Kul (2018) tarafından yapılan çalışmayla ortaya konmuştur. Aynı çalışmada kullanılan etki ağırlığı hesaplamada kullandığı eşitlik (1) ile her bir markanın duygu puanı hesaplanmıştır. Bu formülde, pc (positif count - pozitif tweet sayısı), plc (positif listed count - pozitif tweetlerin okunma sayısı), nc (negatif count - negatif tweet sayısı), tc (total count - toplam tweet sayısı) ve tlc (total listed count - toplam tweet okunma sayısı) değerleri kullanılmıştır.

$$f = \frac{pc+plc-(nc+nlc)}{tc+tlc} \quad (1)$$

4. Duygu Puanlarıyla Hisse Senedi Getiri İlişkisi

Bu çalışmada Borsa İstanbul'da işlem gören Alcatel, Turkcell ve Vestel hisse senetleri arasındaki oynaklık ilişkileri ile hisse senetleri için tweet verilerinden elde edilen duygu puanlarının oynaklığa etkisi incelenmiştir. Bu bağlamda oynaklık ilişkilerinin ortaya çıkarılması için Engle ve Kroner (1995) tarafından geliştirilen BEKK-GARCH(1,1)-X yöntemi kullanılmıştır. Bu yöntemde getirilerin koşullu varyansları arasındaki ilişkilerin yayılımı incelenebilmektedir. Bu modelde için kullanılan ortalama denklemleri eşitlik (2)'teki gibidir:

$$R_t = \Phi + \sum_{i=1}^p \Gamma_i R_{t-i} + \varepsilon_t$$

$$\varepsilon_t \sim (0, H_t) \text{ ve } H = \begin{pmatrix} h_{11,t} & \cdots & h_{1n,t} \\ \vdots & \ddots & \vdots \\ h_{m1,t} & \cdots & h_{mn,t} \end{pmatrix}, \text{ burada } m=n=3$$

(2)

$$R_t = \begin{bmatrix} R_{1,t} \\ R_{2,t} \\ R_{3,t} \end{bmatrix}; \varepsilon_t = \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{2,t} \end{bmatrix}; \Phi_t = \begin{bmatrix} \Phi_1 \\ \Phi_2 \\ \Phi_3 \end{bmatrix}; \Gamma_i = \begin{pmatrix} \Gamma_{11}^{(i)} & \cdots & \Gamma_{1n}^{(i)} \\ \vdots & \ddots & \vdots \\ \Gamma_{m1}^{(i)} & \cdots & \Gamma_{mn}^{(i)} \end{pmatrix}$$

Eşitlik (1)'de R_t ürünlere ait getiri vektörünü, Γ_i bir dönem önceki getirilere ait parametre matrisini, h_i değişkeninin ortalama denklemindeki volatilitesi (oynaklığı), Φ her bir getiriye ait sabit terim parametresini ve ε_t hata terimlerini göstermektedir. BEKK-GARCH(1,1)-X modelinin cebirsel gösterimi eşitlik (3)'teki gibidir:

$$H_t = CC' + \sum_{k=1}^l \rho_{kk}^2 E_k E_k' + \sum_{k=1}^l \varphi_{kk}^2 F_k F_k' + \sum_{k=1}^l \lambda_{kk}^2 G_k G_k' + \sum_{j=1}^f B_j' H_{t-j} B_j + \sum_{k=1}^g A_k' \varepsilon_{t-k} \varepsilon_{t-k}' A_k$$

(3)

İlk model için Eşitlik (2)'de H, C, A ve B 3x3 boyutunda matrislerdir. A ve B matrisi sırasıyla kısa dönem şokları ve uzun dönem volatilitiyi göstermektedir. E, F ve G matrisleri ise duygu puanlarının koşullu varyans üzerindeki etkisi gösteren parametreler olarak ortaya çıkmaktadır. Burada E, F ve G matrisleri sırasıyla tweet verilerinden elde edilen Alcatel'in, Turkcell'in ve Vestel'in duygu puanlarını koşullu varyans üzerindeki etkisini ifade etmektedir. Duygu puanları burada sabit bir terim gibi yalnızca koşullu varyansı artıran bir faktördür. Çalışmada analizler getiri serileri elde edilerek yapılmıştır. Getiri serileri;

$$R_{t,1,t+1} = 100 * \ln \left(\frac{P_{i,t+1}}{P_{i,t}} \right)$$

olacak şekilde elde edilmiştir. R burada getiriyi, P ise fiyatı göstermektedir. Tablo 1’de getiri serilerine ait bazı tanımlayıcı istatistikler verilmiştir.

Tablo: 1
Betimleyici İstatistikler

| İstatistikler | Alcatel (R _{1,t}) | Turkcell (R _{2,t}) | Vestel (R _{3,t}) |
|---|-----------------------------|------------------------------|----------------------------|
| Ortalama | -0,003 | -0,004 | 0,007 |
| Std. dev. | 3,596 | 1,966 | 2,307 |
| Eğiklik | -0,569 | -0,332 | -9,165 |
| Basıklık | 10,192 | 6,814 | 118,010 |
| Jarque-Bera | 1104,439 (0,000) | 492,248 (0,000) | 149755,304 (0,000) |
| Getiri Düzeylerindeki otokorelasyonların Test Edilmesi: | | | |
| Ljung-Box Q(5) | 21,467 (0,000) | 6,616 (0,250) | 2,521 (0,773) |
| Ljung-Box Q ² (5) | 103,669 (0,000) | 2,127 (0,831) | 0,002 (0,999) |
| Getiri Düzeylerindeki ARCH Etkisinin Test Edilmesi: | | | |
| ARCH-LM(5) | 18,843 (0,000) | 0,402 (0,847) | 0,006 (0,999) |
| ARCH-LM ² (5) | 10,251 (0,000) | 0,053 (0,998) | 0,004 (1,000) |
| Getirileri için Durağanlık Birim Kök Testi: | | | |
| ADF | -9,437*** | -18,069*** | -15,784*** |

Not: ***, **, * sırasıyla %1, %5 ve %10 seviyelerinde boş hipotezin reddedildiğini göstermektedir. ARCH-LM ve ARCH etkileri için Lagrange gösterirken, Ljung-Box Q kalıntılarda ve / veya kareli kalıntılarda sıralı bağımlılık testini göstermektedir. ADF, sabit ve trend değişkenleri dikkate alarak Artırılmış Dick-Fuller testini belirtir. Parantez içinde ilişkili p-değerleri bulunur.

Tablo 1’e göre incelenen üç hisse senedi içerisinde sadece Vestel’in ortalama getirisi pozitif olarak karşımıza çıkmaktadır. Alcatel ve Turkcell’in getiri ortalamaları sırasıyla -0,003 ve -0,004 olarak tespit edilmiştir. Getiri serileri içinde ise en fazla koşulsuz standart sapmanın Alcatel’de olduğu görülmektedir. Turkcell ve Vestel’in koşulsuz standart sapmaları, 1,966 ve 3,307 olarak karşımıza çıkmaktadır. İncelenen üç seri içinde eğiklik katsayısının ise negatif olduğu dikkat çekmektedir. Buna göre, üç getirisi seri içinde negatif değerlerin pozitif değerlere göre gelme olasılığı daha fazladır. Basıklık değerleri incelendiğinde ise serilerin leptokurtic yani sivri uçlu ve şişman kuyruklu bir dağılım sergilediği ifade edilebilir. Diğer bir ifadeyle getiri serilerinin normal dağılmamaktadır. Jarque-Bera istatistiği de bu durumu doğrular niteliktedir.

Getiri serilerinde otokorelasyon ve değişen varyans problemlerinin incelenmesi için Ljung-Box Q ve ARCH-LM testleri kullanılmıştır. Elde edilen sonuçlara göre çeşitli anlamlılık düzeylerinde, serilerde otokorelasyon ve değişen varyans problemleri olduğu tablo üzerinden görülmektedir. Son olarak serilerin durağanlık analizleri ADF (Augmented Dickey-Fuller) birim kök testi kullanılarak yapılmıştır. Elde edilen bulgulara göre getiri serileri I(0) seviyesinde durağan olarak tespit edilmiştir.

Tablo 2’de VAR(1) BEKK - GARCH (1, 1)-X modelinin getiri serilerinin ortalama ve varyans denklem sonuçları verilmiştir.

Tablo 2
VAR(1) BEKK-GARCH (1,1)-X Modeli Sonuçları

| Tahmin | Alcatel (R _{1,t}) | Turkcell (R _{2,t}) | Vestel (R _{3,t}) |
|------------------------------|-----------------------------|------------------------------|----------------------------|
| <i>Ortalama Eşitliği</i> | | | |
| <i>Sabit</i> | -0,003 (-0,021) | 0,200** (2,314) | 0,160*** (5,491) |
| Γ_{i1} | -0,129*** (-3,159) | 0,025 (0,980) | -0,042*** (-3,056) |
| Γ_{i2} | -0,065 (-0,799) | -0,209*** (-4,612) | 0,079*** (2,686) |
| Γ_{i3} | 0,032 (0,418) | -0,013 (-0,255) | -0,209*** (-14,236) |
| <i>Varyans Eşitliği</i> | | | |
| <i>c_{1i}</i> | 1,895*** (24,265) | | |
| <i>c_{2i}</i> | 0,425*** (4,122) | 1,493*** (32,616) | |
| <i>c_{3i}</i> | 0,060 (0,945) | 0,152** (2,342) | -0,000 (-0,000) |
| <i>a_{1i}</i> | 0,563*** (14,249) | 0,022 (0,542) | 0,080*** (5,262) |
| <i>a_{2i}</i> | -0,396*** (-3,526) | 0,157** (2,422) | -0,339*** (-8,517) |
| <i>a_{3i}</i> | -0,409*** (-4,683) | 0,360*** (2,950) | 1,120*** (63,757) |
| <i>b_{1i}</i> | 0,407*** (14,923) | -0,074** (-2,134) | -0,089*** (-7,470) |
| <i>b_{2i}</i> | -0,118 (-0,962) | 0,303*** (5,605) | -0,247*** (-7,516) |
| <i>b_{3i}</i> | 0,455*** (6,056) | 0,240*** (4,830) | 0,670*** (36,257) |
| <i>e_{1i}</i> | 0,529 (0,427) | | |
| <i>e_{2i}</i> | -0,434 (-0,587) | -0,000 (-0,000) | |
| <i>e_{3i}</i> | -0,151 (-0,152) | -0,000 (-0,000) | -0,000 (-0,000) |
| <i>f_{1i}</i> | -6,056*** (-7,108) | | |
| <i>f_{2i}</i> | 0,785 (0,580) | -1,010 (-0,712) | |
| <i>f_{3i}</i> | -0,538 (-0,447) | 0,274 (0,295) | 0,000 (0,000) |
| <i>g_{1i}</i> | 20,449*** (8,680) | | |
| <i>g_{2i}</i> | 1,831 (1,396) | 0,000 (0,000) | |
| <i>g_{3i}</i> | -3,314*** (-4,764) | 0,000 (0,000) | 0,000 (0,000) |
| <i>Tamsal Testler</i> | | | |
| Ljung-Box Q(5) | 8,143 [0,148] | 1,964 [0,854] | 7,740 [0,171] |
| Ljung-Box Q ² (5) | 8,504 [0,130] | 1,916 [0,860] | 3,105 [0,683] |
| ARCH-LM(5) | 1,576 [0,167] | 0,359 [0,876] | 0,608 [0,694] |
| ARCH-LM ² (5) | 0,208 [0,958] | 0,040 [0,999] | 0,018 [0,999] |

*Not: ***, **, * sırasıyla %1, %5 ve %10 seviyelerinde boş hipotezin reddedildiğini göstermektedir. ARCH-LM ve ARCH etkileri için Lagrange gösterirken, Ljung-Box Q kalıntılarda ve / veya kareli kalıntılarda sıralı bağımlılık testini göstermektedir. Parantez içinde ilişkili t-değerleri; köşeli parantez içindekiler ilişkili p-değerlerini göstermektedir.*

Tablo 2'de görüleceği üzere ortalama eşitlikler için Alcatel'in bir dönem gecikmesinde meydana gelecek %1'lik artış, kendi getirisini %0,12 azaltmaktadır. Benzer olarak Turkcell'in bir dönem gecikmesinde %1'lik artış da kendi getirisini %0,20

düşürmektedir. Diğer taraftan, Vestel'in getirisi kendi bir dönem gecikmesinin yanı sıra Alcatel ve Turkcell'in bir dönem gecikmelerinden istatistiki olarak anlamlı bir şekilde etkilenmektedir. Alcatel'in bir dönem gecikmesinde meydana gelecek %1'lik artış ise Vestel'in getirisini %0,04 azaltırken, Turkcell'in ki %0,07 artırmaktadır. Son olarak Vestel'in kendi gecikmesinde ki artışlar getirisi üzerinde negatif etkiler yaratmaktadır. Kısaca ifade edilecek olursa, Alcatel ve Turkcell'in gecikmeli getirilerinde meydana gelecek değişimler Vestel'in getirisi üzerinde önemli etkilere neden olmaktadır.

Getiri serilerinin varyans eşitliklerine bakıldığında Alcatel'in kısa dönem şokları kendi koşullu varyansını pozitif olarak etkilemektedir. BEKK-GARCH-X modelinin yapısı gereği delta metodu kullanarak marjinal etkiler incelendiğinde Alcatel'in kısa dönem şoklarının kendi koşullu varyansı üzerindeki etkisi 0,316 olarak karşımıza çıkmaktadır. Diğer taraftan Alcatel'in koşullu varyansı kendi kısa şokları ile Turkcell'in kısa şoklarının çarpımından negatif (-0,382) olarak etkilenmektedir. Benzer durum Alcatel ile Vestel arasındaki (-0,457) çarpım durumu içinde geçerlidir. Turkcel (-0,396²) ve Vestel'in (-0,409²) kısa dönem şokları Alcatel'in koşullu varyansını pozitif olarak etkilemektedir. Diğer yandan, Alcatel'in kısa dönem şokları Turkcell'in koşullu varyansı üzerinde istatistiki olarak anlamlı bir etkiye sahip değilken Vestel'in kısa dönem şokları Turkcell'in koşullu varyansını artırmaktadır. Son olarak, Alcatel'in (0,080²) ve Turkcell'in (-0,339²) kısa dönem şokları Vestel'in koşullu varyansını artırmaktadır. Uzun dönem oynaklıkların koşullu varyans üzerindeki etkileri incelendiğinde Alcatel'in kendi uzun dönem oynaklığı kendi koşullu varyansını (0,165) artırmaktadır. Turkcell'in uzun dönem oynaklığının Alcatel'in koşullu varyansı üzerinde istatistiki olarak anlamlı bir etkisi yokken Vestel'in uzun dönem oynaklığı istatistiki olarak anlamlı bir şekilde Alcatel'in koşullu varyansını artırmaktadır. Son olarak Turkcell'in ve Vestel'in koşullu varyansı kendi uzun dönem oynaklıklarının yanında diğer getirilerin uzun dönem oynaklıklarından istatistiki olarak anlamlı ve pozitif bir şekilde etkilenmektedir.

Çalışmanın da ana konusunu oluşturan tweet verilerinden elde edilen duygu puanlarının Alcatel, Turkcell ve Vestel'in getirilerinin koşullu varyansları üzerindeki etkisi incelendiğinde, Alcatel için elde edilen duygu puanlarının hem kendi hem de diğer getirilerin koşullu varyansları üzerinde istatistiki olarak anlamlı bir etkiye sahip olmadığı tablo üzerinden görülmektedir. Diğer taraftan, Turkcell için elde edilen duygu puanları Alcatel'in koşullu varyansını (36,675) artırmaktadır. Buradan Turkcell hakkında atılan pozitif tweetlerin Alcatel de belirsizlik yarattığı ifade edilebilir. Benzer durum Vestel'in duygu puanlarının Alcatel üzerindeki etkisi içinde geçerlidir. Son olarak, Vestel'in duygu puanları kendi koşullu varyansını (10,982) artırmaktadır.

Davranışsal finans perspektifinden bu durum incelenecek olursa, 3 hisse senedinin de telefon üretici olduğu gerçeğinden yola çıkarak bir piyasadaki bilgi akışının diğerlerine doğru yayıldığı ifade edilebilir. Özellikle Turkcell ve Vestel için atılan tweetlerin Alcatel üzerinde etkili olması, Tversky ve Kahneman (1979) ve Damasio (1994) çalışmalarında da vurgulandığı gibi duyguların insan davranışları üzerinde etkisi olduğunu açıkça göstermektedir. Bu çalışmadan elde edilen bulgulara göre yatırımcıların birbirine ikame olacak sektörlerdeki hisse senetleri arasındaki bilgi akışını dikkate alarak risk

minimizasyonuna yönelmesi ve bu bağlamda hisse senetlerinde oynaklığın artmasının hem davranışsal finans (Olsen 1998; Ricciardi ve Simon 2000, Shefrin, 2002; Barberis ve Thaler 2003; Shiller 2003; Frankel ve Kothari 2004; Baker 2010) hem de etkin piyasa hipotezi (Fama, 1970) bağlamında tutarlı olduğu söylenebilir. Diğer yandan, Merkezi Kayıt Kuruluşu (2020)'na göre Borsa İstanbul'da 2019 yılı Mart ayından, 2020 Şubat ayına yatırımcı sayısındaki yaklaşık olarak 345 bin artış dikkat çekmektedir. Bu yatırımcıların hem kâr maksimizasyonunu hem de riski en aza indirecek şekilde yatırım yapması için piyasalardaki bilgi akışını olabilecek en iyi şekilde değerlendirmesi gerekmektedir. Çalışmadan elde edilen bulgular neticesinde sosyal medyanın hisse senetlerinin getiri oynaklıkları üzerinde etkisi olmasının tespit edilmesi yatırımcıların portföy tercihleri için sosyal medyayı kullanması gerektiğini vurgulamaktadır.

5. Sonuç ve Tartışma

Bu çalışmada, twitter verileri üzerinden yapılan duygu analizi ile Borsa İstanbul'da faaliyet gösteren 3 telefon üreticisinin hisse senedinin fiyat oynaklıkları üzerindeki etkisi BEKK-GARCH(1,1)-X yöntemi kullanılarak incelenmiştir. Bu çerçevede 2017 yılına (Şubat-Temmuz) ait Türkçe tweetler Twitter üzerinden çekilmiş ve ilgili formüller yardımıyla her hisse senedi için duygu puanları oluşturulmuştur. İlgili yazın taraması sonucunda bu çalışmanın Türkçe literatürdeki ilk çalışma olduğunu söylenebilir. Dolayısıyla yapılan analizler ve elde edilen bulguların literatüre katkısı oldukça önemlidir. Yapılan analiz sonuçlarına göre Turkcell ve Vestel için elde edilen duygu puanları Alcatel'in koşullu varyansını artırdığı tespit edilmiştir. Son olarak, Vestel'in duygu puanları kendi koşullu varyansını artırmaktadır. Elde edilen bu sonuç, Bollen vd. (2011), Bollen ve Mao (2011), Sabherwal vd. (2011), Vu vd. (2012), Sprenger vd. (2014) ve Oliveira vd. (2017) tarafından elde edilen sonuçlarla örtüşmektedir. Buna göre yatırımcıların, davranış finans perspektifinden sosyal medya araçlarını getiri oynaklıklarının tahmin edilmesinde kullanmasının, yatırımcı yararına olacağı ifade edilebilir. Çünkü hem bilgiye erişebilme kolaylığı hem de hızlı ve ücretsiz olması sosyal medya araçlarını cazip hale getirmektedir. Sosyal medya üzerinden yayılan bilgi aynı anda binlerce kişi tarafından takip edilerek bir yayılma etkisi yaratmakta ve bundan dolayı getiri oynaklığı oluşturmaktadır.

Gelecek çalışmalarda Türkçe tweetlerin farklı ağzlar ile telaffuzuna yönelik normalizasyon çalışmalarının yapılması ve analize dâhil edilecek Twitter verilerinin İngilizce için de yapılması planlanmaktadır. Ayrıca Facebook gibi yaygın kullanılan diğer sosyal medya ortamları ve diğer ülke borsalarının verileri de analize dâhil edilerek çalışmanın kapsamı genişletecektir.

Kaynaklar

- Aich, S. & H.C. Kim & M. Sain & B.B. Deo (2017), "Analyzing stock price changes using event related Twitter feeds", *19th International Conference on Advanced Communication Technology (ICACT)*, 484-487.
- Attigeri, G.V. & M.P. MM & R.M. Pai & A. Nayak (2015), "Stock market prediction: A big data approach", *TENCON 2015-2015 IEEE Region 10 Conference*, 1-5.

- Baker, H.K. & J.R. Nofsinger (eds.) (2010), *Behavioral finance: investors, corporations, and markets*, Vol. 6, John Wiley & Sons.
- Barberis, N. & R. Thaler (2003), "A survey of behavioral finance", *Handbook of the Economics of Finance*, 1, 1053-1128.
- Bing, L. & K.C. Chan & C. Ou (2014), "Public sentiment analysis in Twitter data for prediction of a company's stock price movements", *2014 IEEE 11th International Conference on e-Business Engineering*, 232-239.
- Bollen, J. & H. Mao & X. Zeng (2011), "Twitter mood predicts the stock market", *Journal of Computational Science*, 2(1), 1-8.
- Bollen, J. & H. Mao (2011), "Twitter mood as a stock market predictor", *Computer*, (10), 91-94.
- Cakra, Y.E. & B.D. Trisedya (2015), "Stock price prediction using linear regression based on sentiment analysis", *2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 147-154.
- Chan, W.S. & R. Frankel & S.P. Kothari (2004), "Testing behavioral finance theories using trends and consistency in financial performance", *Journal of Accounting and Economics*, 38, 3-50.
- Claster, W.B. & H. Dinh & M. Cooper (2010), "Naive Bayes and Unsupervised Artificial Neural Nets for Cancun Tourism Social Media Data Analysis", *The Nature and Biologically Inspired Computing (NaBIC), 2010 Second World Congress*, Kitakyushu, Japan.
- Çoban, Ö. & B. Özyer & G.T. Özyer (2015), "Sentiment analysis for Turkish Twitter feeds", *23rd Signal Processing and Communications Applications Conference (SIU)*, 2388-2391.
- Damasio, A.R. (1994), *Descartes' error: Emotion, reason, and the human brain*, New York: Grosset/Putnam.
- Engle, R. & K. Kroner (1995), "Multivariate simultaneous generalized ARCH", *Econometric Theory*, 11, 122-150.
- Hu, Z. & J. Jiao & J. Zhu (2014), *Using Tweets to Predict the Stock Market*, <<http://cs229.stanford.edu/proj2013/HuJiaoZhu-Using%20Tweets%20to%20Predict%20the%20Stock%20Market.pdf>>, 23.08.2019.
- Jin, X. & D. Shen & W. Zhang (2016), "Has microblogging changed stock market behavior? Evidence from China", *Physica A: Statistical Mechanics and its Applications*, 452, 151-156.
- Kaminski, J. & P. Gloor (2014), "Nowcasting the bitcoin market with twitter signals", *arXivpreprintarXiv:1406.7577*, Cornell University.
- Kul, S. (2018), "Türkçe tweetlerin analiz edilebilmesi için hadoop/hive kullanan melez bir sistemin geliştirilmesi", *Yayınlanmamış Doktora Tezi*, Atatürk Üniversitesi Sosyal Bilimler Enstitüsü.
- Liu, L. & J. Wu & P. Li & Q. Li (2015), "A social-media-based approach to predicting stock comovement", *Expert Systems with Applications*, 42(8), 3893-3901.
- Malkiel, B.G. & E.F. Fama (1970), "Efficient capital markets: A review of theory and empirical work", *The Journal of Finance*, 25(2), 383-417.
- Mao, Y. & W. Wei & B. Wang & B. Liu (2012), "Correlating S&P 500 stocks with Twitter data", *Proceedings of the first ACM international workshop on hot topics on interdisciplinary social networks research*, 69-72.

- Mittal, A. & A. Goel (2012), *Stock prediction using twitter sentiment analysis*, Stanford University, CS229, <<http://cs229.stanford.edu/proj2011/GoelMittal-StockMarketPredictionUsingTwitterSentimentAnalysis.pdf>>, 23.08.2019.
- MKK (2020), *İstatistikler*, <<https://www.mkk.com.tr/tr/content/Bilgi-Merkezi/Sistem-Istatistikleri>>, 23.08.2019.
- Nisar, T.M. & M. Yeung (2018), "Twitter as a tool for forecasting stock market movements: A short-window event study", *The journal of finance and data science*, 4(2), 101-119.
- Oliveira, N. & P. Cortez & N. Areal (2017), "The impact of microblogging data for stock market prediction: Using Twitter to predict returns, volatility, trading volume and survey sentiment indices", *Expert Systems with Applications*, 73, 125-144.
- Olsen, R. (1998), "Behavioral Finance and Its Implications for Stock-Price Volatility", *Financial Analysts Journal*, 54(2), 10-18,
- Pagolu, V.S. & K.N.R. Challa & G. Panda & B. Majhi (2016), "Sentiment Analysis of Twitter Data for Predicting Stock Market Movements", *International conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*.
- Qasem, M. & R. Thulasiram & P. Thulasiram (2015), "Twitter sentiment classification using machine learning techniques for stock markets", *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 834-840.
- Rao, T. & S. Srivastava (2012), "Analyzing stock market movements using Twitter sentiment analysis", *Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012)*.
- Ricciardi, V. & H.K. Simon (2000), "What is behavioral finance?", *Business, Education & Technology Journal*, 2(2), 1-9.
- Sabherwal, S. & S.K. Sarkar & Y. Zhang (2011), "Do internet stock message boards influence trading? Evidence from heavily discussed stocks with no fundamental news", *Journal of Business Finance & Accounting*, 38(9-10), 1209-1237.
- Shefrin, H. (2002), "Beyond greed and fear: Understanding behavioral finance and the psychology of investing", *Oxford University Press on Demand*.
- Shiller, R.J. (2003), "From efficient markets theory to behavioral finance", *Journal of Economic Perspectives*, 17(1), 83-104.
- Si, J. & A. Mukherjee & B. Liu & S.J. Pan & Q. Li & H. Li (2014), "Exploiting social relations and sentiment for stock prediction", *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1139-1145.
- Skuza, M. & A. Romanowski (2015), "Sentiment analysis of Twitter data within big data distributed environment for stock prediction", *2015 Federated Conference on Computer Science and Information Systems (FedCSIS)*, 1349-1354.
- Sprenger, T.O. & A. Tumaşjan & P.G. Sandner & I.M. Welpel (2014), "Tweets and trades: The information content of stock microblogs", *European Financial Management*, 20(5), 926-957.
- Şimşek, M.U. & S. Özdemir (2012), "Analysis of the Relation Between Turkish Twitter Messages and Stock Market Index", *The 6th International Conference: Application of Information and Communication Technologies (AICT)*, Tbilisi, Georgia.
- Thaler, R. (1991), *Quasi Rational Economics*, New York: Russell Sage Press.
- Thaler, R. (1993), *Advances in Behavioral Finance*, New York: Russell Sage Press.

- Türkmenoglu, C. & A.C. Tantug (2014), "Sentiment Analysis in Turkish Media", *The International Conference on Machine Learning*, China.
- Tversky, A. & D. Kahneman (1979), "Prospect theory: An analysis of decision under risk", *Econometrica*, 47(2), 263-291.
- Vu, T.T. & S. Chang & Q.T. Ha & N. Collier (2012), "An experiment in integrating sentiment features for tech stock prediction in twitter", *Proceedings of the workshop on information extraction and entity analytics on social media data*, 23-38.
- Wei, W. & Y. Mao & B. Wang (2016), "Twitter volume spikes and stock options pricing", *Computer Communications*, 73, 271-281.
- Wood, A.S. (1995), "Behavioral Finance and Decision Theory in Investment Management", *Proceedings of the AIMR Seminar Improving the Investment Decision-Making Process: Behavioral Finance and Decision Theory*, April 4 1995, Marina Del Rey, California. Association for Investment Management and Research.
- Ye, S. & S.F. Wu (2010), "Measuring message propagation and social influence on twitter.com", *International Conference on Social Informatics*, Springer, Berlin, Heidelberg, 216-231.
- Zhang, X. & H. Fuehres & P.A. Gloor (2011), "Predicting stock market indicators through twitter 'I hope it is not as bad as I fear'", *Procedia-Social and Behavioral Sciences*, 26, 55-62.