

COVID-19 Salgını Sürecinde Duygu Skorlarının ve Teknik İndikatörlerin Kullanılmasıyla BİST 100 Endeksi Trend Tahmini

Araştırma Makalesi/Research Article

 Meltem ALACA,  Aysun GÜRAN*

Bilgisayar Mühendisliği Bölümü, Doğuş Üniversitesi, İstanbul, Türkiye

20182105003@dogus.edu.tr, adogrusoz@dogus.edu.tr

(Geliş/Received:19.04.2022; Kabul/Accepted:19.08.2022)

DOI: 10.17671/gazibtd.1106017

Özet— Duygu analizi, metinsel verilerin yansıttığı olumlu, olumsuz ya da nötr duyguların tespit edilme aşamalarını içermektedir. Farklı konular hakkında yayınlanan metinsel verilerin yansıttığı olumlu ve/veya olumsuz duygular, insanların, küçük veya büyük ölçekli şirketlerin karar alma süreçlerini etkileyebilmektedir. Metinsel kaynakların yansıttığı duygular, duygu skorları ile sayısal hale getirilebilmekte ve bu skorlar zaman serilerinin ileriye yönelik kestirimlerinin yapılması aşamasında kullanılabilirler. Bilindiği gibi 1 Aralık 2019 tarihinde Çin'in Vuhan kentinde ortaya çıkan ve tüm dünyayı etkisi altına alan koronavirüs küresel bir pandemiye sebep olmuştur. Bu durum uluslararası ve yerel sermaye piyasaları üzerinde keskin düşümlere neden olmuştur. Bu çalışmada pandemi döneminde toplanan haber metinlerinden elde edilen duygu skorlarının Bist100 endeksi trendinin belirlenmesinde etkili olup olmadığının analizi yapılmıştır. Analizlerde borsa endeksleri kestirimlerinde büyük önem arz eden teknik indikatörlerden de yararlanılmıştır. Böylece pandemi döneminde borsa endeksinin artma veya azalma yönün belirlenmesinde duygu skorlarının ve teknik indikatörlerin etkisi görülebilmektedir. Yapılan analizler sonucunda, haber metinlerinden elde edilen duygu skorlarının borsa endeksi trend tahminlemesi üzerinde bazı periyotlar için etkili olduğu gözlemlenmiştir.

Anahtar Kelimeler— duygu skoru, borsa teknik indikatörleri, covid-19, Bist100, duygu analizi

Predicting BIST 100 Index Movement by using Sentiment Scores and Technical Indicators during the COVID-19 Pandemic

Abstract— Sentiment analysis includes the stages of identifying the positive, negative or neutral emotions contained in the text data. Positive and/or negative emotions reflected by the text data can affect the decision-making processes of people, small or large-scale companies. Emotions reflected by documents can be vectorized with sentiment scores and these scores could be useful to forecast time series models. As it is known, the coronavirus, which emerged in Wuhan, China on December 1, 2019 caused a global pandemic all around the world. The Covid-19 pandemic has brought sharp and sudden declines in global and domestic stock markets. Within the scope of our study, it was analyzed whether the sentiment scores obtained from the Covid-19 related news documents were effective in forecasting the trend of the Bist100 index. Technical indicators that have great importance in estimating stock market indices, were also used in the analysis. Hence, the effect of sentiment scores and technical indicators in determining the trend of the stock market index during the pandemic could be observed. As a result of the study, it was observed that the sentiment scores were effective to predict the price trend of stock market index for some periods.

Keywords— sentiment scores, technical indicators, covid-19, Bist100, sentiment analysis

1. GİRİŞ (INTRODUCTION)

Sermaye piyasalarında işlem gören endekslerin getiri yönlerinin tahmin edilebilmesi daha etkili yatırım stratejilerinin geliştirilebilmesine olanak sağlamaktadır. Fakat dinamik yapısından ötürü borsa endekslerinin değişim yönünün pek çok faktöre bağlı olması, endekslerin getiri değerlerinin ya da hareket yönlerinin tahminini zorlaştırmaktadır. Beklenmedik bir şekilde ortaya çıkan olaylar borsa endekslerinde dalgalanmalar yaratabilecek faktörler arasındadır [1]. Bilindiği gibi 1 Aralık 2019 tarihinde Çin'in Wuhan kentinde ortaya çıkan ve tüm dünyayı etkisi altına alan koronavirüs küresel bir pandemiye neden olmuştur. Pandemi nedeniyle ortaya çıkan belirsiz ortam ulusal ve uluslararası mali piyasalar üzerindeki dalgalanmaları arttırmıştır. ABD, İngiltere, İspanya ve İtalya gibi gelişmiş ekonomiler bu ani değişimden en çok etkilenen ülkeler arasına girmiştir [2]. ABD borsalarında işlem gören Dow Jones endeksi, 2020 yılı Mart ayında %12,94 düşüşle 1987'de yaşanan kara pazartesiden bu yana en büyük kaybı yaşamıştır. Ülkemizde de 11 Mart 2020 tarihinde görülen ilk vaka ile iktisadi faaliyetler sınırlanmaya başlamıştır. Mart ayı boyunca Borsa İstanbul'da hacmi en yüksek 100 adet hisse senedini kapsayan borsa endeksinde (BİST-100) %16,75 oranında değer kaybı yaşanmıştır. Pandemi koşullarının Borsa İstanbul'da yer alan sektörel endeksler üzerindeki etkisini inceleyen bir çalışma, salgın ile ilgili olaylar karşısında duyarlılığı en yüksek olan sektörlerin teknoloji, ulaştırma, turizm, KOBİ ve bilişim sektörleri olduğunu belirtmiştir [1]. Çalışma [3], salgın döneminde ulaştırma ve turizm sektörlerinin de olumsuz etkilendiğini ve bu sektörlerin hisse senedi getirilerinin düştüğünü ifade etmiştir.

Etkin piyasa hipotezi Farma tarafından 1970 yılında geliştirilen ve finans alanında en çok araştırılan hipotezlerden biri olmuştur [4]. Bu hipoteze göre yatırımcılar rasyoneldir. Piyasalardaki fiyatlar mevcut bilgiyi tam olarak yansıtır. Öte yandan davranışsal finans yaklaşımına göre menkul kıymetlerin fiyatlarının piyasa aktörlerince "algılanan" fiyatı yansıtmaktadır [5]. Yani yatırımcılar her zaman rasyonel olamayabilir. Bu iki yaklaşımın yakınlaştırılmasıyla ortaya çıkan Adaptif Piyasa Hipotezi'ne göre piyasaların etkinliği olaylar ve yapısal değişiklikler nedeniyle değişebilmektedir [6]. Bu yaklaşımlar doğrultusunda, son zamanlarda beklenmedik bir şekilde gelişen olayların, finansal haber dokümanlarının/makalelerin, sosyal medya platformlarında paylaşılan kısa iletilerin yansıtmış olduğu duygu durumları ile sermaye piyasaları arasındaki ilişki analizi önemli araştırma konuları arasına girmiştir.

Çetin çalışmasında [7], beklenmedik bir şekilde ortaya çıkan pandemi döneminde salgın oranının yüksek olduğu bir dönemde, Borsa İstanbul BİST-100 endeksinin kapanış, açılış, en yüksek ve en düşük fiyatları ile COVID-19 vaka sayılarının pozitif yönde ilişkili olduğunu belirtmiştir. Atan ve Çınar çalışmasında [8], finansal piyasalar ile ilgili haber metinlerinin yansıttığı olumlu, olumsuz duygu değerleri ile finansal değerler arasında anlamlı bir korelasyonun

olduğunu vurgulamışlardır. Ateş ve Güran çalışmaları ile [9], sosyal medyada paylaşılan kısa iletiler ile BİST30 endeksi arasında anlamlı korelasyon ilişkileri olduğunu belirtmiş ve ülkemiz adına önemli olayların olduğu bir dönemde duygu değerleri ve endeks arasında Granger nedensellik ilişkine rastlandığını ifade etmişlerdir. Pek çok çalışmanın deneysel olarak elde etmiş olduğu bulgular neticesinde günümüzde haber kaynaklarının beklenmedik olaylar ile ilgili yansıttığı olumlu ve/veya olumsuz algıların finansal piyasaları etkileyebileceği görülmektedir [10-15]. Bu çalışmalarda, olumlu ve olumsuz haber sayılarının kullanılması ile elde edilen duygu değerlerinin, gerek borsa endekslerinin gelecekteki değerlerinin ön görülmesinde; gerekse endekslerin pozitif/negatif yönlü hareketlerinin tahmin edilmesinde etkili olduğu belirtilmiştir.

Duygu değerleri haber metinlerinde yansıtılan olumlu, olumsuz genel kanının sayısallaştırılmasını sağlayan bir ölçümdür. Duygu değerlerinin endeksler üzerindeki etkinliğini araştıran çalışmaların çoğu yabancı sermaye piyasaları hakkındadır. Bu çalışmada 20.12.2019 – 16.02.2021 tarihleri arasında koronavirüs hakkında yayınlanan haber dokümanları VADER kütüphanesi yardımıyla olumlu, olumsuz ve nötr olarak etiketlendikten sonra üç farklı duygu değeri elde edilmiştir. Duygu değerleri elde edildikten sonra aynı tarih aralığındaki BİST100 endeksleri ele alınmış ve endeks değerlerinin günlük bazda artış veya azalış durumları, değer artışı olursa +1, tersi durumda 0 etiketleri ile belirtilerek analiz edilen durum bir sınıflandırma problemine dönüştürülmüştür. Sınıflandırma probleminde endeks trend tahmini için haber metinlerinden elde edilen duygu değerlerinin yanı sıra en sık kullanılan indikatörler arasında yer alan sekiz farklı adet teknik indikatör de kullanılmıştır. Teknik indikatörler yatırım araçlarının alımı/satımı konusunda karar verilmesine veya gelecekteki fiyatlarının/fiyat yönlerinin tahmin edilebilmesine katkı sunmaktadırlar. Belirtilen özelliklerin kullanımıyla farklı zaman periyotlarının da baz alınmasıyla sınıflandırma algoritması olarak Rassal Orman (RO) topluluk öğrenmesi algoritması kullanılmış ve elde edilen doğruluk değerleri paylaşılmıştır. Analizler sonucunda haber verilerinden elde edilen duygu değerlerinin ve borsa teknik indikatörlerinin trend tahminlemesine önemli ölçüde katkı sağladığı gözlemlenmiştir.

Çalışmamızın geri kalanı şu şekildedir. Bölüm 2 ile literatür araştırması ele alınmış; Bölüm 3 ile kullanılan veri kümesi tanıtılmış; Bölüm 4 ile sınıflandırma problemine ele alınan özellikleri ve RO algoritması sunulmuş; Bölüm 5 ile model parametreleri ve deneysel sonuçlar paylaşılmış ve son olarak sonuçlar başlığı elde edilen bulgular yorumlanmıştır.

2. LİTERATÜR TARAMASI (LITERATURE REVIEW)

Bu bölümünde borsada işlem gören hisse senetleri getirilerinin tahmin edilmesine ya da getirilerin değişim yönünün tespit edilmesine yönelik yapılan çalışmalar incelenmiştir.

Pen ve ark. tarafından hazırlanan çalışmada hisse senedi fiyatlarının pozitif ya da negatif yönlü değişim durumu bir sınıflandırma problemine dönüştürülmüştür [10]. İlgili problemde derin ağlar kullanılmıştır. Algoritmaya girdi olarak geçmiş hisse senedi fiyatlarının yanı sıra haberlerden elde edilen BoK (Bag of Keywords), PS (Polarity Score), CT (Category Tag) değerleri de kullanılmıştır. Analizler sonucunda haberlerden elde edilen değerlerin tahminlemede başarımı yükselttiği görüldüğü en düşük hata oranının %43.13'lük bir değer ile kapanış fiyatının yanı sıra BoK, PS ve CT değerlerinin beraber kullanılmasıyla elde edildiği görülmüştür.

Huynh ve ark. tarafından gerçekleştirilen analizde S&P 500 borsa endeksi üzerinde çalışılmıştır. Yapılan çalışmada finansal haberlerden elde edilen vektör değerleri ve geçmiş hisse senedi fiyatları bir araya getirilerek hisse senedi fiyatları yön tahminlemesi yapılmıştır [15]. Uzun kısa süreli bellek (LSTM), kapılı tekrarlayan hücre (Gated Recurrent Unit) ve makale tarafından önerilen BGRU (Bidirectional Gated Recurrent Unit) modelleri ilgili veriler üzerinde uygulanmıştır. Elde edilen sonuçlara göre BGRU modelinin diğer modellerden daha yüksek başarımlı değerleri ürettiği belirtilmiştir.

Bouktif ve ark. çalışmasında AMZN NASDAQ üzerinde hisse senedi artma azalma yönü tahmin edilmiştir [11]. Yapılan çalışmada DVM, lojistik regresyon (LR), ileri beslemeli sinir ağları (Feed Forward Neural Networks), RO ve ekstrem gradyan artırma (Extreme Gradient Boosting) modelleri kullanılmıştır. Algoritmaya girdi olarak ise hisse senedi fiyatlarının açılış, kapanış gibi özelliklerinin yanı sıra Twitter'dan ilgili hisse senedi ile toplanmış kısa iletilerden elde edilen duygu polarite özellikleri, duygu skorları, N-gramlar ve gecikmeli özellik değerleri kullanılmıştır. Elde edilen sonuçlarda en iyi sonucun yaklaşık %63 doğruluk oranıyla RO algoritması ile alındığı görülmüştür.

Mohan ve ark. çalışmasında S&P500 endeksine ait beş yıllık günlük endeks değerleri ve şirketler hakkındaki 265000 adet finansal makaleden elde edilen pozitif ve negatif polarite değerleri kullanılarak endeks tahminlemesi yapılmıştır [13]. Çalışmada, pozitif ve negatif polarite değerlerinin modellere katıldığı durumun, sadece hisse senedi fiyatlarının değerlendirildiği durumdan daha iyi sonuçlar verdiği belirtilmiştir. Model olarak Facebook Prophet, ARIMA ve tekrarlayan sinir ağları (Recurrent Neural Networks) kullanılmıştır. Sonuç olarak duygu değerleri ile kapanış fiyatlarının yönü arasında korelasyon olduğu görülmüştür.

Mehtap ve Şen çalışmasında, ulusal Hindistan borsasının NIFTY 50 endeksi üzerinde çalışılmıştır [19]. 2015 – 2018 tarihleri arasındaki günlük borsa verileri Yahoo finance web sitesi üzerinden elde edilmiştir. Kapanış fiyatı tahminlemesi ve trend tahminlemesi için sekiz adet regresyon, sekiz adet sınıflandırma modeli kullanılmıştır. Modellere girdi olarak açılış, kapanış gibi elde edilen borsa verileri kullanılmıştır. Sınıflandırma modelleri içerisinde en başarılı sonuçları %74.91 hassaslık oranı ile RO

algoritmasının, regresyon modelleri içerisinde ise uygulamalı evrişimsel sinir ağı (Applied Convolutional Neural Network) algoritmasının sergilediği görülmüştür.

Picasso ve ark. tarafından sunulan çalışmada NASDAQ100 endeksindeki en değerli 20 şirketin borsa verileri üzerinde çalışılmıştır. Çalışmanın amacı piyasa eğilimini tahminlemektir [12]. Piyasa eğilimini tahminlemede algoritma özellik girdisi olarak Basit Hareketli Ortalama (Simple Moving Average), Üstel Hareketli Ortalama (Exponential Moving Average), Hareketli Ortalama Yakınlaşma Uzaklaşma Göstergesi, Göreceli Güç Endeksi, Bollinger Bantları, Stokastik Osilatör, Gerçek Aralık (True Range), Ortalama Gerçek Aralık, Williams %R Göstergesi ve CR teknik indikatörlerinin yanında haberlerden elde edilen duygu değerleri de işleme katılmıştır. Model olarak RO, DVM ve YSA kullanılmıştır. En başarılı sonuçların yaklaşık %70 doğruluk oranı ile haberlerden elde edilen duygu değerleri ve teknik indikatörlerin bir arada kullanılmasıyla elde edildiği görülmüştür.

Li ve ark. tarafından gerçekleştirilen çalışmada Hong Kong borsasında beş yıllık bir veri kümesi üzerinde, borsa trend tahminlemesi için çalışılmıştır [14]. Özellik değeri olarak duygu skorları ve Hareketli Ortalama (Moving Average), Hareketli Ortalama Yakınlaşma Uzaklaşma Göstergesi (Moving Average Convergence Divergence), Göreceli Güç Endeksi (Relative Strength Index) ve Para Akış Göstergesi (Money Flow Index) borsa teknik indikatörleri kullanılmıştır. Çalışmada uzun kısa süreli bellek (Long Short Term Memory), çoklu çekirdek öğrenme (Multiple Kernel Learning) ve SVM modelleri karşılaştırılmıştır. Yapılan çalışmada uzun kısa süreli bellek algoritmasının yaklaşık %83 doğruluk oranı ile daha başarılı bir başarımlı değeri ürettiği belirtilmiştir.

Patel ve ark. tarafından sunulan çalışmada Hindistan borsa hisse senetlerinin negatif/pozitif yönlü hareketleri bir sınıflandırma problemine dönüştürülerek tahmin edilmeye çalışılmıştır [16]. Çalışmada eğitim veri kümesi olarak on yıllık borsa verisi analiz edilmiş ve yapay sinir ağları (YSA), destek vektör makineleri (DVM), RO ve Naive-Bayes (NB) sınıflandırma algoritmaları kullanılmıştır. Sınıflandırma probleminde dikkate alınan özellik değerleri popüler borsa teknik indikatörleridir. Kullanılan teknik indikatörler, “Basit hareketli ortalamalar (Simple Moving Average)”, “Hareketli ortalamalar (Weighted Moving Average)”, “Stokastik K% (Stochastic K%)”, “Stokastik D% (Stochastic D%)”, “Momentum”, “Göreceli güç endeksi (Relative Strength Index)”, “William R% (William's R%)”, Toplam Dağıtım Osilatörü (Accumulation/Distribution Oscillator)”, “Mal Kanal Endeksi (Commodity Channel Index)” göstergeleridir. Elde edilen sonuçlarda RO algoritmasının %91.48 doğruluk değeriyle en iyi başarımlı değerine sahip olduğu belirtilmiştir.

Khaidem ve ark., Apple, Microsoft ve Samsung hisse senetleri üzerinde çalışmıştır [17]. Model olarak RO algoritması kullanılmıştır. Modele girdi olarak Göreceli

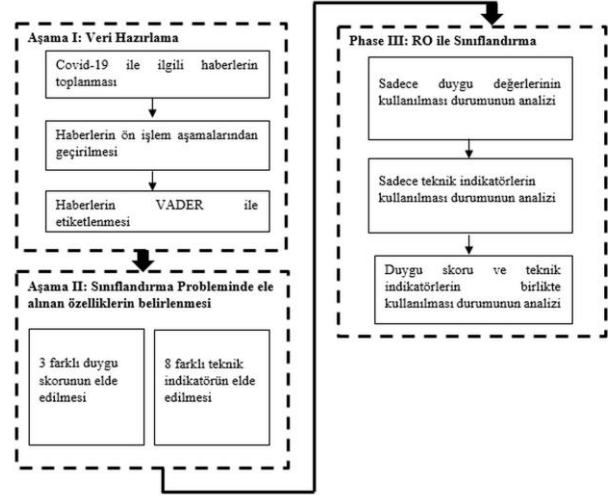
Güç Endeksi, Stokastik Osilatör, Hareketli Ortalama Yakınlaşma Uzaklaşma Göstergesi, Williams %R Göstergesi, Fiyat Değişim Oranı (Price Rate of Change), Denge İşlem Miktarı (On Balance Volume) borsa teknik indikatörleri sunulmuştur. Hisse senetleri kapanış fiyatı artma azalma yönü tahminlemesi için alınan sonuçlarda RO algoritması ile %85 ve %95 aralığında başarımlar elde edilmiştir. RO algoritması ile daha önce kullanılan [18] referansı ile belirtilen DVM algoritmasından daha etkili başarımlara ulaşıldığı belirtilmiştir. Bununla beraber en yüksek başarımlar %92.08 ile 10 günlük kestirimler için Apple hisse senedi üzerinde elde edilmiştir.

Di., Apple, Amazon, Microsoft hisse senetleri üzerinde çalışmıştır. Sınıflandırma algoritması olarak destek vektör makinesi algoritması kullanılmıştır [18]. Sınıflandırma probleminde ilgili hisse senetlerinin artma azalma yönü tahminlemesi yapılmıştır. Sınıflandırma aşamasında borsa değerlerinin yanı sıra Momentum, Hareketli Ortalama Yakınlaşma Uzaklaşma Göstergesi, Göreceli Güç Endeksi, Williams %R Göstergesi, Mal Kanalı Endeksi (CCI), Fiyat Değişim Oranı, Ortalama Yönel Gösterge (ADX), Üçlü Üstel Hareketli Ortalama, Denge İşlem Miktarı, Zaman Serisi Tahminlemesi, Ortalama Gerçek Aralık, Para Akış Göstergesi borsa teknik indikatörleri kullanılmıştır. Çalışmada teknik indikatörlerin trend tahminlemedeki başarılı sonuçları paylaşılmıştır. Belirtilen sonuçlara göre en yüksek doğruluk değerinin %77.125 ile Microsoft hisse senedi için 7 günlük kestirimlerde elde edildiği vurgulanmıştır.

Kullanımları 1933'lü yıllara uzanan teknik indikatörler hisse senetlerinin gelecekteki fiyat eğilimlerini belirlemek için geçmiş fiyatları, işlem hacimlerini ve diğer mevcut verileri kullanmaktadır [32]. Biz bu çalışmada özellikle duygu analizinin de katılarak gerçekleştirildiği hisse senedi eğilimlerini tahminleyen [12], [14], [16], [18] çalışmalarında kullanılan yüksek başarımlarına ulaşılmasını sağlayan "Ağırlıklı Hareketli Ortalama", "Göreceli Güç Endeksi", "Hareketli Ortalama Yakınlaşma Uzaklaşma Göstergesi", "Williams %R Göstergesi", "Stokastik %K", "Stokastik %D", "Chaikin Toplama-Dağıtım Osilatörü", "Mal Kanalı Endeksi" göstergelerini kullandık. Bu teknik indikatörlerin yanına duygu skor değerlerini de ekleyerek Bist100 enkesinin trend tahminlemesi üzerine bir çalışma gerçekleştirdik. Böylece finansal metinlerin yansıttığı duygu durumlarının Türkiye piyasaları üzerinde etkili olabileceğini deneysel bir şekilde göstermiş olduk.

3. ÖNERİLEN METODOLOJİ (RECOMMENDED METHODOLOGY)

Bu çalışmada Şekil 1 ile belirtilen aşamalar takip edilmiştir. Bundan sonraki alt bölümlerde her bir aşama detaylıca anlatılmıştır.

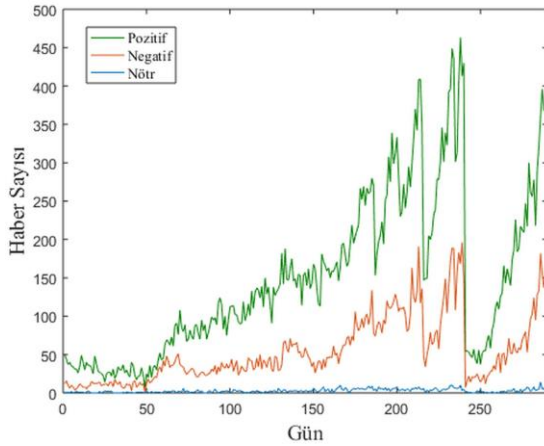


Şekil 1. Önerilen Metod Çerçevesi (Suggested Method Framework)

3.1. Kullanılan Veri Kümesi (Dataset)

Bu çalışmada 20.12.2019-16.02.2021 tarihleri arasında farklı haber kaynakları tarafından Covid-19 ile ilgili olarak yayınlanmış olan ("korona", "koronavirüs", "coronavirus", "corona", "kovid19", "covid19", "kovid-19", "covid-19" vb. kelimeleri içeren) haber metinleri Java programlama dili ile yazılan bir yazılım ile toplanmıştır. Veri kümemiz bu tarihlardaki 290 işlem gününde paylaşılan 59313 haber dokümanını içermektedir. Veri kümesi üzerinde duygu analizi aşamalarını uygulamadan önce haber metinleri durak kelimelerinden arındırılmıştır. Ön işlem aşamasının ardından duygu analizinin gerçekleştirilmesi için sözlük tabanlı bir duygu analizi aracı olan VADER yazılımı kullanılmıştır. Bu araç ile -4 ile +4 arasında bir ölçekte kelimelerin duygu yoğunluk toplamlarına göre olumlu, olumsuz ve nötr olarak etiketlenebilmektedir. VADER yazılımından faydalanmak için Türkçe haber metinleri Google çeviri ara yüzü (Google translate api) kullanılarak İngilizce'ye çevrilmiştir. 20.12.2019-16.02.2021 tarih aralığında toplam 290 işlem günü mevcuttur. Bu günlerde duygu analizi gerçekleştirilen toplam haber sayısı 59313'tür. VADER yazılımı ile bu haber dokümanlarından 42892 tanesi olumlu, 15621 tanesi olumsuz ve 800 tanesi nötr olarak etiketlenmiştir. 290 iş günü boyunca yayınlanan olumlu, olumsuz ve nötr haber sayılarının yayılımı Şekil 2 ile belirtilmiştir.

Günlük bazda yayınlanan olumlu ve olumsuz haber sayılarının kullanılmasıyla Bölüm 4.1'de belirtilen S_1 , S_2 ve S_3 olarak adlandırılan 3 farklı duygu değeri oluşturulmuş ve bu değerlerin her biri sınıflandırma probleminde 8 adet teknik indikatörün yanında ayrı ayrı birer özellik olarak kullanılmıştır.



Şekil 2. 290 iş gününde paylaşılan olumlu, olumsuz ve nötr haber dağılımı
(Distribution of positive, negative and neutral news shared in 290 working days)

Duygu değerleri dışında özellik değeri olarak kullanılan sekiz adet teknik indikatör “ağırlıklı hareketli ortalama”, “göreceli güç indeksi”, “hareketli ortalama yaklaşma uzaklaşma göstergesi”, “Williams %R göstergesi”, “stokastik %K”, “stokastik %D”, “Chaikin toplama dağıtım Osilatörü”, “mal kanal indeksi” göstergeleridir.

290 işlem günü için Bist100 endeksi kapanış fiyatlarının belirli bir güne göre artma ya da azalma durumunu bir sınıflandırma problemi olarak değerlendirebilmek adına BİST-100 endeksi kapanış değerleri günlük olarak baz alınmıştır. 290 işlem günü için günlere ait sınıf etiketleri, endekste belirli periyotta göre artış meydana geldiğinde +1, azalış meydana geldiğinde ise 0 etiketlerinin yaratılması ile belirlenmiştir. Böylece sınıflandırma probleminde ele alınacak özellikler ve değişecek periyotlara göre belirlenen sınıf etiketleri elde edilebilmiştir.

3.2. Modelde Kullanılan Özellikler (Features Used in the Model)

Bu bölümde sınıflandırma problemi olarak ele aldığımız endeks eğilimini öngörme problemimizde kullanılacak olan özellikler (duygu değerleri ve sıklıkla kullanılan 8 adet teknik indikatör) anlatılacaktır. Bilindiği gibi teknik indikatörler, yatırım araçlarının geçmiş ve güncel fiyat ve/veya hacim faaliyetlerine bağlı olarak yatırım araçlarının alımı/satımı konusunda karar verilmesine veya gelecekteki fiyatlarının ve/veya fiyat yönlerinin tahmin edilebilmesine olanak sağlarlar. Çalışmamızdaki amacımız pandemi döneminde yayımlanan olumlu ya da olumsuz haber sayılarından elde edilen duygu skor değerlerinin ve popüler teknik indikatörlerin BİST-100 endeksi kapanış değerlerinin pozitif ya da negatif yönlü değişim durumlarının tahmini üzerindeki başarımlarını incelemektir. Çalışmamızda kullanılan duygu skorlarına ve 8 adet teknik indikatöre ait olan açıklamalar sırasıyla aşağıdaki alt başlıklarda verilmektedir:

3.2.1. Duygu Skorları (Sentiment Scores)

Duygu skorlarının oluşturulması için [9] çalışması baz alınmış ve VADER yazılımı ile olumlu ve olumsuz olarak etiketlenen haber sayıları kullanılmıştır. Bu kapsamda S_1, S_2, S_3 olmak üzere üç adet duygu skoru elde edilmiştir.

$$S_1 = \frac{Nolumlu_t - Nolumsuz_t}{Nolumlu_t + Nolumsuz_t} \quad (1)$$

$$S_2 = \frac{Nolumlu_{t+1}}{Nolumsuz_{t+1}} \quad (2)$$

$$S_3 = \frac{Nolumlu_t}{Nolumlu_t + Nolumsuz_t} - \frac{Nolumlu_{t-1}}{Nolumlu_{t-1} + Nolumsuz_{t-1}} \quad (3)$$

Bu denklemlerde $Nolumlu_t$ ve $Nolumsuz_t$ incelenen günlerde paylaşılan olumlu ve olumsuz haber sayılarını belirtmektedir. $Nolumlu_{t-1}, Nolumsuz_{t-1}$ ise (t-1). günde paylaşılan olumlu, olumsuz haber sayısını ifade etmektedir.

3.2.2. Ağırlıklı Hareketli Ortalama (Weighted Moving Average – WMA)

Ağırlıklı ortalama, verilen periyotta sondaki kapanış fiyatlarına daha çok ağırlık verilerek hesaplanan ortalamadır. Yakın zamandaki fiyattan, geçmiş verilere doğru önemlilik doğrusal olarak azalmaktadır. Pek çok çalışmada kısa dönem tahminlemelerinde 10 günlük periyot kullanıldığı için bu çalışmada da 10 günlük periyot baz alınmıştır [16], [20]. Ağırlıklı hareketli ortalama mevcut trendi teyit etmektedir. C_n , endeksin n. günündeki kapanış fiyatı olmak üzere ağırlıklı hareketli ortalama formülü denklem 4 ile belirtildiği gibidir.

$$WMA = \frac{(1) \times C_1 + (2) \times C_2 + (3) \times C_3 + \dots + (n) \times C_n}{1+2+3+\dots+n} \quad (4)$$

3.2.3. Göreceli Güç Endeksi (Relative Strength Index – RSI)

Bu indikatör endeksin bünyesi içindeki gücünü kendisine göre karşılaştırır [21]. Endeksin geçmişte yükseldiği günlerle, azaldığı günlerin karşılaştırılmasına yardımcı olur. RSI hesaplanırken genellikle kısa dönem analizleri için on dört günlük periyot kullanılmaktadır [22-24]. RSI, istenen periyotlardaki artış ve azalış yaşanan günleri ve bu günlere ait ortalama değerleri olarak bu değerlerin birbirlerine oranını bulmaktadır. C_i ilgili gündeki kapanış fiyatı; n periyot olmak üzere, RSI hesabı denklem 5-8 ile belirtilen ara aşamalara bağlı olarak denklem 9 ile belirtilmiştir [16-10]:

$$PD_i = \begin{cases} C_i - C_{i-1}, & C_i > C_{i-1} \\ 0, & C_i \leq C_{i-1} \end{cases} \quad (5)$$

$$ND_i = \begin{cases} 0, & C_i > C_{i-1} \\ |C_i - C_{i-1}|, & C_i \leq C_{i-1} \end{cases} \quad (6)$$

$$OPD_i = \frac{\sum_{i=1}^n PD_i}{n} \quad (7)$$

$$OND_i = \frac{\sum_{i=1}^n ND_i}{n} \quad (8)$$

$$RSI_i = 100 - \left[\frac{100}{1 + \left(\frac{OPD_i}{OND_i} \right)} \right] \quad (9)$$

3.2.4. Hareketli Ortalama Yaklaşma Uzaklaşma Göstergesi (Moving Average Convergence Divergence – MACD)

Hareketli ortalama yaklaşma uzaklaşma göstergesi, belirlenen periyotlardaki üssel hareketli ortalamaların farkı alınarak elde edilir. Bu gösterge endekste oluşabilecek trendin düşüş mü yükseliş mi olduğunu anlamamıza yardımcı olur. Genellikle, on iki ve yirmi altı günlük periyotlar için üstel hareketli ortalama sonuçlarının çıkarılmasıyla bulunur [22], [25]. On iki gün, hızlı hareketli ortalama için seçilen periyot iken, yirmi altı gün yavaş hareketli ortalama için seçilmiştir. MACD serisi belirtilen üstel hareketli ortalamaların birbirinden çıkarılmasıyla elde edilmektedir. Bu gösterge için ayrıca MACD serisinin 9 dokuz günlük üstel hareketli ortalamasının alınmasıyla MACD sinyal değeri hesaplanmaktadır. EMA, üstel hareketli ortalama; n, periyot ve C_n ilgili periyottaki kapanış fiyatı olmak üzere gösterge formülü denklem 10 – denklem 12 ile belirtilmiştir.

$$EMA_n = \left(\frac{2}{n+1} C_n \right) + \left(\frac{n-1}{n+1} EMA_{n-1} \right) \quad (10)$$

$$MACD = EMA_{12} - EMA_{26} \quad (11)$$

$$Sinyal = S = EMA_9(MACD) \quad (12)$$

3.2.5. Williams %R Göstergesi (Williams %R)

Williams %R göstergesi, kapanış fiyatının saptanan periyot içerisindeki en düşük ve yüksek fiyatlara göre yerini tespit edilebilmektedir. Bu göstergede hesaplamalar genellikle 14 gün üstünden yapılmaktadır [18], [22], [26], [27]. $L_{n(t)}$, son n periyodun en düşük değeri (t. günde görüldüğü varsayılınsın), $H_{n(t)}$ son n periyodun en yüksek değeri (t. günde görüldüğü varsayılınsın) olmak üzere gösterge formülü denklem 13 ile belirtilmiştir.

$$Williams \%R = \frac{H_{n(t)} - \text{güncel kapanış fiyatı}}{H_{n(t)} - L_{n(t)}} \times 100 \quad (13)$$

3.2.6. Stokastik Osilatör (Stochastic Oscillator – STOCH)

Stokastik Osilatör, kısa dönem için işlem yapanların en çok kullandığı alım satım göstergelerindendir [22], [28]. Bu gösterge hesabı %K ve %D değerlerine sahiptir. [26], [17] çalışmalarında belirtildiği gibi %K için on dört günlük, %D için ise üç günlük periyot ile çalışılmıştır. Çalışmamız kapsamında uygulanan modelde bu indikatörden gelen %K

ve % D değerleri iki ayrı teknik indikatör olarak kullanılmıştır.

$L_{n(t)}$, son n periyodun en düşük değeri (t. günde görüldüğü varsayılınsın) ve $H_{n(t)}$ son n periyodun en yüksek değeri (t. günde görüldüğü varsayılınsın) olmak üzere göstergenin kapsadığı %K ve %D formülleri denklem 14 ve denklem 15 ile belirtilmiştir.

$$\%K = \frac{\text{güncel kapanış fiyatı} - L_{n(t)}}{H_{n(t)} - L_{n(t)}} \times 100 \quad (14)$$

%D eğrisi, yavaşlatılmış %K'nın 3 günlük hareketli ortalamasının alınmasıyla elde edilmektedir.

$$\%D = \frac{\sum_{t=1}^n \%K_t}{n} \quad (15)$$

3.2.7. Chaikin Toplama-Dağıtım Osilatörü (Chaikin Accumulation/Distribution Oscillator – ADOSC)

Chaikin toplama-dağıtım osilatörü fiyat ve işlem hacminin analiz edilmesiyle yönsel bilgi vermeyi amaçlar ve toplama dağıtım endeksinin üç güne ait üstel hareketli ortalaması ile on güne ait üstel hareketli ortalamasının farkı alınarak elde edilir [22], [29].

CO, Chaikin Osilatörü; ADL, toplama dağıtım çizgisi (Accumulation-Distribution Line); $ADL(EMA_3)$, ADL'nin 3 günlük üstel hareketli ortalaması; $ADL(EMA_{10})$, ADL'nin 10 günlük üstel hareketli ortalaması; MFM, para akışı çarpanı (Money Flow Multiplier); MFV, para akışı hacmi (Money Flow Volume); n periyot; V_n , n. gündeki hacim; L_n , n. gündeki düşük fiyat; H_n , n. gündeki yüksek fiyatı olmak üzere indikatör formülü denklem 16 – denklem 21 ile belirtilmiştir:

$$CO = ADL(EMA_3) - ADL(EMA_{10}) \quad (16)$$

$$ADL(EMA_n) = [ADL_n \times k] + [ADL(EMA_{n-1}) \times (1 - k)] \quad (17)$$

$$k = \frac{2}{n+1} \quad (18)$$

$$ADL = ADL_{n-1} + MFV_n \quad (19)$$

$$MFV_n = MFM \times V_n \quad (20)$$

$$MFM = \frac{(C_n - L_n) - (H_n - C_n)}{(H_n - L_n)} \quad (21)$$

3.2.8. Mal Kanal Endeksi (Commodity Channel Index – CCI)

Mal kanal endeksi, fiyattaki döngüsel değişimleri belirtir [18]. İlgili indikatör için kullanılan periyot [30] çalışmasında olduğu gibi on dört gün alınmıştır. Gösterge formülü denklem 22- 25 ile verilmektedir [22].

$$CCI_i = \frac{A_i - B_i}{D \times 0.015} \quad (22)$$

$$A_i = \frac{H_i + L_i + C_i}{3} \quad (23)$$

$$B_i = \frac{\sum_{i=1}^n \frac{H_{(i)} + L_{(i)} + C_{(i)}}{3}}{n} \quad (24)$$

$$D_i = \frac{\sum_{i=1}^n |A_i - B_i|}{n} \quad (25)$$

4. MODEL PARAMETRELERİ VE DENEYSSEL SONUÇLAR (MODEL PARAMETERS AND EXPERIMENTAL RESULTS)

Çalışmamızda BİST100 endeksi kapanış değerlerinin artış/azalış yönü (trend) tahminlemesi için Jupyter Notebook geliştirme ortamında Python dili kullanılmıştır. Yararlanılan kütüphaneler numpy, pandas, scikit-learn, talib kütüphaneleridir. Endeksin artış/azalış yönü tahminlemesi için ilgili günün sınıf etiketleri, endeks değeri bir önceki periyoda göre arttığında 1 ve diğer durumda 0 değerine sahip olacak şekilde oluşturulmuştur. Böylece trend tahminlemesi problemi ikili sınıflandırma problemine dönüştürülebilmıştır. Sınıflandırma probleminde ele alınacak özneliker haber metinlerinden elde edilen duygu değerleri ve 8 adet teknik indikatördür. Tüm özellikler min-max normalizasyona göre 0-1 aralığında değer alacak şekilde normalize edilmiştir. Özneliklerin de oluşturulmasıyla beraber modelimizde RO sınıflandırma algoritması kullanılmıştır. Sınıflandırma algoritması olarak RO'nun seçilme sebebi, pek çok çalışma tarafından bu alanda iyi başarımlar elde edildiğinin belirtilmesidir [11], [16], [17], [19]. RO algoritması, eğitim aşamasında çok sayıda karar ağacının oluşturulmasıyla çalışan bir topluluk öğrenmesi yöntemidir. RO'da karar ağaçlarının inşa edilmesinde de yeniden örnekleme metodu ve bölme kriteri olarak ise Gini safsızlık (Gini Impurity) yöntemi temel alınmaktadır [17]. Çalışmamızda RO algoritmasında ağaç sayısı olarak [17] çalışmada kullanıldığı gibi 5 ağaç kullanılmıştır. Algoritmada endeksin artma azalma yönü tahminlemesi için mevcut gün ile belirli periyottaki önceki gün karşılaştırmasında [31] çalışmasında olduğu gibi 20 günlük periyoda kadar tahminlemeler yapılmıştır. Ele alınan her periyot için eğitim veri kümesi %75-%25 olarak bölünmüş ve bu kümenin baz alınmasıyla algoritma beş kez çalıştırılmıştır. Son olarak elde edilen doğruluk (accuracy) ve F-skor değerlerinin standart sapma ve ortalama değerleri hesaplanarak Tablo 1 ve Tablo 2 ile sonuçlar paylaşılmıştır.

Bir sınıflandırma algoritmasının tahmin değerlerini ve örneklemlerin gerçek sınıf etiketleri hakkında bilgi veren aşağıda belirtilen doğruluk matrisindeki TP (True Positive), TN (True neagtive), FP (False Positive) ve FN (False Negative) değerleri kullanılarak doğruluk değeri ve F-skor değeri denklem 26 - denklem 29 ile belirtildiği gibi hesaplanmaktadır.

	Tahmini Durum		
	+1	-1	
Gerçek	+1	TP	FN
Durum	-1	FP	TN

Şekil 3. Karmaşıklık Matrisi (Confusion Matrix)

$$\text{Doğruluk değeri} = (TP+TN)/(TP+FN+FP+TN) \quad (26)$$

$$P = TP/(TP+FP) \quad (27)$$

$$R = TP/(TP+FN) \quad (28)$$

$$F\text{-Skor} = 2PR / (P+R) \quad (29)$$

Tablo 1, farklı özellik kümelerinin kullanılmasıyla gerçekleştirilen deneylerin doğruluk değerlerini göstermektedir.

Tablo 1'e ait ilk üç kolon (Doğruluk S₁, Doğruluk S₂ ve Doğruluk S₃) özellik değeri olarak sadece S₁, S₂, S₃ duygu değerlerinin ayrı ayrı kullanılması ile elde edilen doğruluk değerlerini göstermektedir. Endeksin artış azalış yönünün belirlenmesini ele alan bu sınıflandırma probleminde sadece S₁, S₂, S₃ duygu değerlerinin bireysel olarak kullanılması 20 farklı periyot için sırasıyla ortalama (ilk üç kolon ortalamaları) 0,5771; 0,5659 ve 0,5620 doğruluk değerlerini üretmiştir. Bu sonuçtan yola çıkarak S₁ duygu değerinin daha iyi bir başarımlar değeri ürettiği söylenebilir.

Tablo 1'in "Doğruluk BTİÖK" isimli dördüncü kolonu 8 adet borsa teknik indikatör özellik kümesinin (BTİÖK) dikkate alınmasıyla elde edilen sonuçları paylaşmaktadır. Buradaki sonuçlardan da görüldüğü gibi, tüm periyotlar için borsa teknik indikatörlerinin kullanımı, sadece S₁, S₂, S₃ duygu değerlerinin kullanımına göre daha yüksek başarımlar değerlerinin elde edilmesini sağlamıştır. Ayrıca bu kolonun ortalaması alındığında elde edilen 0,8038 değeri S₁, S₂, S₃ duygu değerlerinin ortalama değerlerinden daha yüksek çıkmıştır.

Tablo 1'in beşinci (Doğruluk BTİÖK + S₁) kolonu, BTİÖK ve S₁'in aynı anda kullanılması ile elde edilen doğruluk değerlerini göstermektedir. Belirli periyotlar için (3,5,6,7,11,12,14,15,16,17,18,20) yani toplam 12 deneyde BTİÖK ve S₁'in birlikte kullanılması BTİÖK ve S₁'in bireysel kullanımlarına göre daha yüksek doğruluk değerlerinin üretilmesini sağlamıştır.

Benzer şekilde Tablo 1'in altıncı kolonu (Doğruluk BTİÖK + S₂) bazı periyotlar için (5,7,15,16,17,18,20), BTİÖK ve S₂'nin birlikte kullanılması BTİÖK ve S₂'nin bireysel kullanımlarına göre daha yüksek başarımlar değeri üretmiştir.

Tablo 1. S_1 , S_2 , S_3 , BTİÖK, BTİÖK+ S_1 , BTİÖK+ S_2 , BTİÖK + S_3 doğruluk değerleri
(S_1 , S_2 , S_3 , BTİÖK, BTİÖK+ S_1 , BTİÖK+ S_2 , BTİÖK + S_3 accuracy values)

Pe ri yo t	Doğruluk S_1	Doğruluk S_2	Doğruluk S_3	Doğruluk BTİÖK	Doğruluk BTİÖK + S_1	Doğruluk BTİÖK + S_2	Doğruluk BTİÖK + S_3
1	0,5126±0,0483	0,4952±0,0418	0,4638±0,0365	0,5579±0,0363	0,5557±0,0366	0,5499±0,0420	0,5422±0,0409
2	0,5357±0,0506	0,5138±0,0437	0,5153±0,0438	0,6142±0,0308	0,6067±0,0323	0,6100±0,0387	0,6085±0,0427
3	0,5801±0,0457	0,5538±0,0455	0,5291±0,0345	0,6709±0,0374	0,6755±0,0405	0,6602±0,0467	0,6656±0,0359
4	0,6068±0,0401	0,5894±0,0500	0,5707±0,0459	0,7248±0,0375	0,7188±0,0385	0,7147±0,0336	0,6959±0,0335
5	0,5978±0,0413	0,5906±0,0348	0,5456±0,0533	0,7412±0,0392	0,7521±0,0338	0,7595±0,0369	0,7532±0,0427
6	0,6018±0,0364	0,5761±0,0428	0,5256±0,0439	0,7935±0,0386	0,7938±0,0396	0,7795±0,0364	0,7862±0,0316
7	0,5992±0,0492	0,5694±0,0495	0,5612±0,0416	0,7981±0,0361	0,8022±0,0289	0,8023±0,0305	0,7886±0,0482
8	0,5571±0,0330	0,5559±0,0429	0,5315±0,0623	0,8216±0,0340	0,8042±0,0312	0,8047±0,0352	0,8201±0,0344
9	0,5696±0,0437	0,5657±0,0403	0,5681±0,0521	0,8232±0,0397	0,8161±0,0394	0,8177±0,0302	0,8314±0,0269
10	0,6087±0,0408	0,5844±0,0433	0,5537±0,0482	0,8368±0,0323	0,8306±0,0419	0,8292±0,0341	0,8251±0,0339
11	0,6017±0,0401	0,5778±0,0412	0,5283±0,0453	0,8495±0,0367	0,8502±0,0336	0,8375±0,0371	0,8315±0,0383
12	0,5901±0,0446	0,5911±0,0518	0,5635±0,0463	0,8699±0,0323	0,8703±0,0339	0,8556±0,0323	0,8587±0,0310
13	0,5710±0,0494	0,5597±0,0367	0,5814±0,0505	0,8766±0,0307	0,8530±0,0370	0,8645±0,0371	0,8546±0,0230
14	0,5836±0,0361	0,5699±0,0459	0,5542±0,0473	0,8627±0,0467	0,8628±0,0337	0,8647±0,0320	0,8546±0,0422
15	0,5735±0,0374	0,5614±0,0380	0,5802±0,0392	0,8702±0,0313	0,8736±0,0396	0,8733±0,0367	0,8620±0,0375
16	0,5774±0,0454	0,5744±0,0493	0,5900±0,0404	0,8704±0,0361	0,8717±0,0293	0,8788±0,0297	0,8698±0,0321
17	0,5458±0,0522	0,5882±0,0397	0,6384±0,0326	0,8817±0,0258	0,8852±0,0278	0,8913±0,0338	0,8740±0,0285
18	0,5559±0,0427	0,5595±0,0479	0,5811±0,0470	0,8613±0,0306	0,8764±0,0325	0,8680±0,0337	0,8651±0,0411
19	0,5868±0,0506	0,5612±0,0447	0,6351±0,0469	0,8756±0,0360	0,8718±0,0275	0,8729±0,0276	0,8689±0,0348
20	0,5865±0,0456	0,5810±0,0483	0,6233±0,0364	0,8767±0,0272	0,8782±0,0338	0,8810±0,0409	0,8749±0,0329

Son olarak Tablo 1'in yedinci kolonu (Doğruluk BTİÖK + S_3) bazı periyotlar için (5,9,18), BTİÖK ve S_3 'ün birlikte kullanılması BTİÖK ve S_3 'ün bireysel kullanımlarına göre daha yüksek başarımlar üretmiştir. Genel olarak bu deneylerde de en iyi doğruluk değerlerine S_1 duygu değerinin katkısıyla ulaşılmıştır.

Tablo 2 yukarıda da bahsi geçen deneylere ait F-Skor değerleri sonuçlarını yansıtmaktadır. Elde edilen F-ölçüm skorları sonuçlarına göre yine S_1 başarımlarının kullanımı S_2 ve S_3 değerlerine göre daha verimli sonuçlar üretmiştir. Koyu renkle gösterilen hücre değerlerine denk gelen periyotlarda borsa teknik indikatörleri ile birlikte duygu değerlerinin de kullanılması başarımların değeri daha yüksek sonuçların görülmesini sağlamıştır.

Tablo 1 ve Tablo 2'den edinilen sonuçlara göre BİST100 endeksinin artış ve azalış yönünün bir sınıflandırma problemi olarak ele alındığı bu çalışmada haber metinlerinden elde edilen duygu skorlarının problem üzerinde pozitif bir etkiye sahip olduğu görülmüştür.

5. DEĞERLENDİRME VE SONUÇ (EVALUATION AND CONCLUSION)

Günümüzde metinsel haber kaynaklarının farklı durumlar ile ilgili yarattığı olumlu ya da olumsuz algılar metin madenciliğinin araştırma alanlarından biri olan duygu analizi aşamalarıyla yapısal hale

getirilebilmektedir. “Olumlu” veya “olumsuz” etiketler ile etiketlenen haber sayılarına bağlı olan duygu skorları, haber kaynaklarının yarattığı olumlu/olumsuz algının pek çok durum üzerindeki etkisini analiz etmek için kullanılmaktadır. Özellikle koronavirüs salgını döneminde yayınlanan haber metinlerinin yansıttığı duygu durumlarının borsa üzerindeki bir etkiye sahip olup olmadığı merak konusu olmuştur. Bu çalışmada koronavirüs konulu haber metinlerinden elde edilen duygu skorlarının BİST100 endeksinin kapanış fiyatlarının trend tahmininde bir etkiye sahip olup olmadığı araştırılmıştır. Böylece haber metinlerinde yansıtılan olumlu ve olumsuz duygularının borsa gibi dinamik bir alanda geleceğe dair kestirimlerin yapılması için kullanılan popüler teknik indikatörler gibi rol alıp alamayacağı test edilebilmiştir. Bu amaçla çalışmamızda 20.12.2019 – 16.02.2021 tarihleri arasındaki bir dönemde derlenen koronavirüs konulu haber külliyatından elde edilen pozitif ve negatif haber sayılarının kullanılmasıyla üç farklı duygu skoru hesaplanmıştır. Ardından BİST 100 trendin belirlenmesi için kullanılacak olan popüler 8 farklı teknik indikatör hesabı gerçekleştirilmiştir. Böylelikle duygu değerlerinin ve teknik indikatörlerin RO algoritmasının kullanılmasıyla sınıflandırma problemine dönüştürülen BİST100 endeksi trend tahmini üzerindeki başarımların değeri kıyaslanabilmiştir.

Tablo 2. $S_1, S_2, S_3, BTiÖK, BTiÖK+S_1, BTiÖK+S_2, BTiÖK + S_3$ F-Skor değerleri $(S_1, S_2, S_3, BTiÖK, BTiÖK+S_1, BTiÖK+S_2, BTiÖK + S_3$ F-Score values)

Pe ri yo t	F-Skor S_1	F-Skor S_2	F-Skor S_3	F-Skor BTiÖK	F-Skor BTiÖK + S_1	F-Skor BTiÖK + S_2	F-Skor BTiÖK + S_3
1	0,5879±0,0491	0,5708±0,0466	0,5468±0,0353	0,6349±0,0360	0,6356±0,0411	0,6304±0,0436	0,6225±0,0443
2	0,6275±0,0471	0,6073±0,0446	0,6215±0,0423	0,6973±0,0326	0,6925±0,0359	0,6989±0,0373	0,6940±0,0432
3	0,6634±0,0500	0,6293±0,0370	0,6127±0,0399	0,7395±0,0366	0,7373±0,0396	0,7295±0,0440	0,7317±0,0332
4	0,7078±0,0359	0,6940±0,0473	0,6719±0,0423	0,8030±0,0316	0,7983±0,0328	0,7959±0,0288	0,7809±0,0285
5	0,7059±0,0346	0,6995±0,0311	0,6633±0,0491	0,8173±0,0311	0,8252±0,0289	0,8296±0,0302	0,8245±0,0342
6	0,7079±0,0354	0,6836±0,0457	0,6519±0,0413	0,8565±0,0299	0,8559±0,0300	0,8465±0,0264	0,8495±0,0254
7	0,7048±0,0515	0,6877±0,0472	0,6799±0,0391	0,8594±0,0280	0,8607±0,0246	0,8627±0,0226	0,8534±0,0367
8	0,6556±0,0350	0,6552±0,0410	0,6391±0,0591	0,8680±0,0282	0,8546±0,0259	0,8540±0,0266	0,8672±0,0306
9	0,6770±0,0393	0,6767±0,0388	0,6773±0,0485	0,8716±0,0309	0,8687±0,0317	0,8697±0,0218	0,8769±0,0218
10	0,7041±0,0387	0,6902±0,0402	0,6673±0,0455	0,8832±0,0246	0,8762±0,0340	0,8781±0,0278	0,8735±0,0249
11	0,6979±0,0372	0,6801±0,0473	0,6442±0,0425	0,8913±0,0289	0,8943±0,0250	0,8826±0,0288	0,8775±0,0324
12	0,6981±0,0462	0,7028±0,0491	0,6809±0,0401	0,9080±0,0255	0,9109±0,0248	0,8977±0,0249	0,8996±0,0228
13	0,6854±0,0451	0,6806±0,0342	0,6973±0,0487	0,9152±0,0219	0,8972±0,0297	0,9074±0,0283	0,8979±0,0192
14	0,6914±0,0333	0,6881±0,0386	0,6734±0,0503	0,9049±0,0331	0,9022±0,0267	0,9041±0,0246	0,8964±0,0309
15	0,6827±0,0401	0,6772±0,0320	0,6936±0,0353	0,9096±0,0242	0,9099±0,0317	0,9109±0,0289	0,9010±0,0294
16	0,6863±0,0470	0,6865±0,0450	0,7039±0,0338	0,9086±0,0275	0,9104±0,0223	0,9157±0,0228	0,9078±0,0237
17	0,6653±0,0503	0,7066±0,0317	0,7376±0,0265	0,9177±0,0184	0,9215±0,0198	0,9241±0,0232	0,9113±0,0217
18	0,6662±0,0366	0,6658±0,0475	0,6843±0,0446	0,9002±0,0240	0,9087±0,0270	0,9037±0,0251	0,9043±0,0303
19	0,6912±0,0497	0,6764±0,0405	0,7372±0,0405	0,9130±0,0263	0,9087±0,0203	0,9102±0,0230	0,9074±0,0263
20	0,6919±0,0385	0,6898±0,0445	0,7270±0,0306	0,9154±0,0200	0,9157±0,0254	0,9190±0,0293	0,9134±0,0252

Yapılan analizler sonucunda öncelikle popüler teknik indikatörlerinin borsa değeri artma azalma yönü tahminlemesi için oldukça etkili olduğu gözlemlenmiştir. Koronavirüs konulu haber verilerinden elde duygu skorlarının trend analizinde tek bir özellik olarak kullanılması iyi bir sınıflandırma doğruluk değeri ürettiği görülmüştür. Duygu değerleri ve teknik indikatörlerin birlikte kullanıldığı analizlerde bazı durumlarda, sadece teknik indikatörlerin kullanıldığı durumlara göre daha iyi başarımların elde edildiği görülmüştür. Böylece haber metinleri ile toplumda yaratılan olumlu, olumsuz algıların borsa üzerinde etkili olabileceği ve bu alanda gerçekleştirilecek olan çalışmalarda duygu skorlarının incelenebilecek bir özellik olabileceği görülmüştür. Bu çalışma ile elde edilen bulgular literatürdeki pek çok çalışma ile örtüşmektedir. Literatür taraması kısmında ele alınan yabancı sermaye piyasaları için gerçekleştirilen çalışmalarda ([16-19]), teknik indikatörlerin trend tahmininde kullanılabilecek etkili özellikler arasında olduğu bildirilmiştir. Paralel bir şekilde, bu çalışma sonucunda da kullanılan teknik indikatörlerin yüksek başarımların elde ettiği görülmüştür. Ayrıca literatürdeki pek çok çalışma, finansal makalelerden [13], [15], borsa ve finans ile ilgili haber metinlerinden [10], [12] ve popüler sosyal medya platformlarında paylaşılan kısa iletilerden elde edilen duygu skor değerlerinin [11], [14] borsa endekslerinin trend analizlerinde kullanılabileceğini belirtmiştir. Bu çalışmada da bazı periyotlar için haber metinlerinden elde edilen duygu değerlerinin BİST100 endeksi üzerinde etkili olduğu gözlemlenmiştir. Sonuç olarak finansal metin

kaynaklarının borsa analizlerinde kullanılabileceği yönünde bir bulguya erişilmiştir.

Gelecek çalışmalarda özellik seçim metodları kullanılarak sınıflandırma aşamasında kullanılan teknik indikatörler ve duygu skorlarının değişik kombinasyonlarla kullanılması planlanmaktadır. Ayrıca tahminleme aşamasında derin ağların (RNN, LSTM vb. gibi [33]) etkisi gözlemlenecek ve geleneksel makine öğrenmesi metodları ile kıyaslamalı analizler sunulacaktır.

KAYNAKLAR (REFERENCES)

- [1] G. G. Yağcılar, "Borsa İstanbul'da COVID-19 Etkisi: Kısa Dönemli Sektörel Piyasa Tepkilerinin Endeks Bazında Ölçülmesi", *Ekonomi Politika ve Finans Araştırmaları Dergisi*, 6(2), 439-463, 2021.
- [2] D. K. Pandey, V. Kumari, "Event study on the reaction of the developed and emerging stock markets to the 2019-nCoV outbreak", *International Review of Economics & Finance*, 71, 467-483, 2021.
- [3] Y. Peker, E. Demirhan, **Covid-19 küresel salgınının borsa İstanbul'daki sektörel etkileri**, Türkiye Ekonomi Politikaları Araştırma Vakfı (TEPAV), Türkiye, 2020.
- [4] E.F. Fama, "Efficient Capital Markets: A Review of Theory and Empirical Work.", *The Journal of Finance*, 25(2), 383-417, 1970.
- [5] K. Eyüboğlu, S. Eyüboğlu, "Borsa İstanbul Endekslerinde Adaptif Piyasa Hipotezinin Geçerliliğinin Test Edilmesi", *Yaşar Üniversitesi E-Dergisi*, 15(59), 642-654, 2020.

- [6] A. W. Lo, "The Adaptive Markets Hypothesis." *The Journal of Portfolio Management*, 30(5), 15-29, 2004.
- [7] A. C. Çetin, "Koronavirüs (Covid-19) Salgınının Türkiye'de Genel Ekonomik Faaliyetlere ve Hisse Senedi Borsa Endeksine Etkisi", *Mehmet Akif Ersoy Üniversitesi Uygulamalı Bilimler Dergisi*, 4(2), 341-362, 2020.
- [8] S. Atan, Y. Çınar, "Borsa İstanbul'da finansal haberler ile piyasa değeri ilişkisinin metin madenciliği ve duygu (sentiment) analizi ile incelenmesi", *Ankara Üniversitesi SBF Dergisi*, 74(1), 1-34, 2019.
- [9] A. Güran, E. Ateş, "Twitter iletileri duygu değerleri ve Bist 30 endeksi günlük değer değişimlerinin Pearson korelasyonu ve Granger nedensellik analizi", *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 36(3), 1687-1702, 2021.
- [10] Y. Peng, H. Jiang, "Leverage financial news to predict stock price movements using word embeddings and deep neural networks", *arXiv preprint arXiv:1506*, 2015.
- [11] S. Bouktif, A. Fiaz, M. Awad, "Stock market movement prediction using disparate text features with machine learning", **In 2019 Third International Conference on Intelligent Computing in Data Sciences (ICDS)**, IEEE, 1-6, October, 2019.
- [12] A. Picasso, S. Merello, Y. Ma, L. Oneto, E. Cambria, "Technical analysis and sentiment embeddings for market trend prediction", *Expert Systems with Applications*, 135, 60-70, 2019.
- [13] S. Mohan, S. Mullapudi, S. Sammeta, P. Vijayvergia, D. C. Anastasiu, "Stock price prediction using news sentiment analysis", **In 2019 IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService)**, IEEE, 205-208, April, 2019.
- [14] X. Li, P. Wu, W. Wang, "Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong", *Information Processing & Management*, 57(5), 102212, 2020.
- [15] H. D. Huynh, L. M. Dang, D. Duong, A new model for stock price movements prediction using deep neural network, **In Proceedings of the Eighth International Symposium on Information and Communication Technology**, 57-62, December, 2017.
- [16] J. Patel, S. Shah, P. Thakkar, K. Kotecha, "Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques", *Expert systems with applications*, 42(1), 259-268, 2015.
- [17] L. Khaidem, S. Saha, S. R. Dey, "Predicting the direction of stock market prices using random forest", *arXiv preprint arXiv:1605.00003*, 2016.
- [18] X. Di, **Stock trend prediction with technical indicators using SVM**, Independent Work Report, Standford: Leland Stanford Junior University, USA, 2014.
- [19] S. Mehtab, J. Sen, "Stock price prediction using convolutional neural networks on a multivariate timeseries", *arXiv preprint arXiv:2001.09769*, 2020.
- [20] Y. Kara, M. A. Boyacioglu, Ö. K. Baykan, "Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange", *Expert systems with Applications*, 38(5), 5311-5319, 2011.
- [21] Ç. Özari, K. Turan, E. Demir, "Teknik indikatörlerin etkinliği: Bist30 ve bist100 endeksleri üzerine bir uygulama", *Nevşehir Hacı Bektaş Veli Üniversitesi SBE Dergisi*, 6(1), 94-113, 2016.
- [22] C. Budak, **Teknik Analiz indikatörlerinin Performans karşılaştırması üzerine bir araştırma**, Doctoral dissertation, Marmara Üniversitesi, 2019.
- [23] S. Gumparthi, "Relative strength index for developing effective trading strategies in constructing optimal portfolio", *International Journal of Applied Engineering Research*, 12(19), 8926-8936, 2017.
- [24] E. A. S. Ahmed, S. B. Goyal, "**Impact of Technical Parameters for Short-and Long-term Analysis of Stock Behavior**", *Materials Today: Proceedings*, 2021.
- [25] P. Fernández-Blanco, D. J. Bodas-Sagi, F. J. Soltero, J. I. Hidalgo, Technical market indicators optimization using evolutionary algorithms, **In Proceedings of the 10th annual conference companion on Genetic and evolutionary computation**, 1851-1858, July, 2008.
- [26] M. R. Vargas, C. E. dos Anjos, G. L. Bichara, A. G. Evsukoff, Deep learning for stock market prediction using technical indicators and financial news articles, **In 2018 International Joint Conference on Neural Networks (IJCNN)**, IEEE, 1-8, July, 2018.
- [27] G. Yapıcı, **Teknik Analiz Yöntemi Ve Bist100'de İşlem Gören Ticari Bankacılık Payları Üzerine Bir Uygulama**, Yüksek Lisans Tezi, İstanbul Okan Üniversitesi, İşletme Anabilim dalı, İşletme Programı, 2020.
- [28] Y. Erdiñç, **Borsada Teknik Analiz El Kitabı, Siyasal Kitabevi**, Ankara, 1996.
- [29] M. N. Günak, **İleri Teknik Analiz Uygulamaları ve Bu Uygulamaların İMKB'de Test Edilmesi**, Yüksek Lisans Tezi, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, 2007.
- [30] V. Vaghela, R. Gor, N. Chavda, "Elliott Wave formation using combination of CCI and DEMA", *IOSR Journal of Mathematics (IOSR-JM)*, 2319-765X, 14-21, 2021.
- [31] P. Sadorsky, "A Random Forests Approach to Predicting Clean Energy Stock Prices", *Journal of Risk and Financial Management*, 14(2), 48, 2021.
- [32] A. Cowles, "Can stock market forecasters forecast?", *Econometrica Journal of the Econometric Society*, 309-324, 1933.
- [33] A. Ayaydın, M.A. Akçayol, "Deep Learning Based Forecasting of Delay on Flights.", *Bilişim Teknolojileri Dergisi*, 15(3), 239-249, 2022.