

Yıllık Raporların Metin Analizi İle Değerlendirilmesi: BİST100 Endeksinde Bir Uygulama

Sedat ÇEREZ¹ - Abdullah Kürşat MERTER²-
Yavuz Selim BALCIOĞLU³ - Gökhan ÖZER⁴

Makale Gönderim Tarihi: 06 Ağustos 2023

Makale Kabul Tarihi: 11 Mart 2024

Öz

Bu çalışmada, Borsa İstanbul (BİST) 100 endeksindeki 60 finansal olmayan işletmenin faaliyet raporları incelenerek, raporların içeriğindeki konuların ve duygu tonunun analizi gerçekleştirilmiştir. Gizli Dirichlet Ayırımı yöntemiyle faaliyet raporları içerisinde dört temel konu üzerinde ayrışma sağlandığı ve raporların içerisinde yer alan açıklamaların yaklaşık %87'sinin olumlu, %13'ünün ise olumsuz bir duygu sergilediği tespit edilmekte ve bu da raporlarda genel olarak iyimser bir görünüm içerisinde olduğunu göstermektedir. Bu araştırma, kurumsal raporlardan anlamlı içgörüler elde etmek için bir araç olarak metin analizinin potansiyelinin altını çizmekte ve kurumsal performans ve stratejiyi anlamak için yeni bir yaklaşım sunmaktadır.

¹ Araş. Gör., sedatcerrez@gtu.edu.tr, ORCID: 0000-0002-6443-6319, Gebze Teknik Üniversitesi, İşletme Fakültesi, İşletme Bölümü (Sorumlu Yazar)

² Araş. Gör., akmerter@gtu.edu.tr, ORCID: 0000-0001-6874-1890, Gebze Teknik Üniversitesi, İşletme Fakültesi, İşletme Bölümü

³ Dr. Öğretim Gör., ysbalcioglu@gtu.edu.tr, ORCID: 0000-0001-7138-2972, Gebze Teknik Üniversitesi, İşletme Fakültesi, Yönetim Bilişim Sistemleri Bölümü

⁴ Prof. Dr. ozer@gtu.edu.tr, ORCID: 0000-0002-3255-998X, Gebze Teknik Üniversitesi, İşletme Fakültesi, İşletme Bölümü, Kırgızistan-Türkiye Manas Üniversitesi, İİBF, İşletme Bölümü

Anahtar kelimeler: *Faaliyet Raporları, Metin Analizi, Duyarlılık Analizi, Konu Modelleme, BIST100*

Jel Sınıflandırması: *G32, C81, C82, C53*

Evaluation Of Annual Reports By Text Annual Reports By Text Analysis: Evidence From BIST100 Index

Abstract

In this study, the annual reports of 60 non-financial firms in the Borsa Istanbul (BIST) 100 index are examined, and the topics and sentiment tone of the reports are analyzed. Using the Latent Dirichlet Allocation method, it is found that four main topics are differentiated in the annual reports, and approximately 87% of the disclosures in the reports exhibit a positive sentiment and 13% of the disclosures exhibit a negative sentiment, indicating that there is a generally optimistic tone in the reports. This research highlights the potential of text analysis as a tool for extracting meaningful insights from corporate reports and offers a new approach to understanding corporate performance and strategy.

Key words: *Annual Reports, Text Analysis, Sentiment Analysis, Topic Modeling, BIST100*

Jel Code: *G32, C81, C82, C53*

1. Giriş

Yıllık raporlar, bir şirketin mali durumunu, stratejilerini ve gelecekteki büyüme beklentilerini değerlendirmek için en değerli kaynaklardan biridir (Barth vd., 2001). Bu raporlar, gelir, gider ve karlılık dahil olmak üzere bir şirketin performansına ilişkin kapsamlı genel bir bakış sağlamakta (Yuthas, Rogers, & Dillard, 2002) ve şirketin faaliyetlerini etkileyebilecek önemli olayları veya gelişmeleri de sunmaktadır. Yatırımcılar ve paydaşlar bu raporları analiz ederek şirketin genel durumu hakkında bilgi sahibi olabilir ve yatırımları hakkında bilinçli kararlar verebilirler (Lewis & Young, 2019). Ayrıca, yıllık raporlar genellikle yönetimin şirketin performansına ilişkin tartışma ve analizini içermekte ve bu da sunulan finansal veriler için daha fazla bağlam ve açıklama sağlamaktadır. Bu niteliksel bilgiler, şirketin finansal başarı performansı ve gelecek beklentileri hakkında daha özgün verileri kapsamaktadır. Bunlara ek olarak faaliyet raporları, yönetici ücretleri ve yönetim kurulu yapısı gibi şirketin kurumsal yönetim uygulamaları hakkında önemli bilgileri de açıklamak-

tadır. Bu da yatırımcıların şirketin etik uygulamalara ve sorumlu karar alma sürecine bağlılığını değerlendirmesine olanak tanır.

Yatırımcılar ekonomik açıdan şirketin en önemli paydaşları arasında yer alsa da, şirket yöneticilerinin sorumlu olduğu farklı çıkarılara sahip çeşitli başka paydaş grupları da vardır. Paydaşlar arasında ticari ortaklar, çalışanlar, rakipler, kredi verenler ve düzenleyiciler yer almaktadır. Şirketlerin paydaşlarının çıkarılarını korumaları gerekmektedir. Bu paydaşlar faaliyet raporu vasıtasıyla yöneticilerin faaliyet döneminde gerçekleştirdiği eylemlere ve geleceğe yönelik stratejileri hakkında bilgilere erişebilmektedir. Dolayısıyla faaliyet raporunda sunulan metinler tüm paydaşlar açısından değerli bilgileri içermektedir (Yuthas vd., 2002).

Geleneksel olarak, bu raporlar yatırımcılar ve analistler tarafından manuel olarak analiz edilirken, metin madenciliği gibi analiz yöntemlerinin ortaya çıkışı, bu raporların yorumlama ve bunlardan içgörü çıkarımlarında değişikliğe neden olmuştur. (Loughran ve McDonald, 2011). Makine öğrenimindeki hızlı gelişmeler nedeniyle, bilgisayar destekli metin analizi yöntemleri de 2010'lu yıllarda önemli ölçüde gelişmiştir (Ranta vd., 2022). Sosyal medya verileri (Twitter, LinkedIn ve Facebook gibi), firmaya özgü haber girişleri, iş ilanları, şirket web sayfaları, finansal ve diğer medya yazıları, analistlerin raporları, araştırma notları ve çeşitli açıklamalar gibi çevrimiçi kaynakların tümü makine öğrenimi tabanlı metin analizi teknikleriyle incelenebilmektedir. Bu çalışmada, 2020 yılında Borsa İstanbul (BIST) 100 endeksinde listelenen 60 şirketin yıllık raporlarını analizinde bu teknolojik gelişmelerden yararlanılmıştır. Spesifik olarak, büyük metinsel veri kümelerinin analizinde sıklıkla kullanılan duyarlılık analizi, *Gizli Dirichlet Ayrımı* (GDA) ve Ngram gibi metin analiz yöntemleri kullanılmıştır (Dyer vd., 2017; Feuerriegel ve Pröllochs, 2021). Amaç, bu raporlarda aktarılan kalıpları, temaları, duyguları ayırt etmek ve bunların şirketlerin performansı ve stratejileri üzerindeki etkilerini anlamak, kurumsal faaliyet raporlarından daha derin içgörüler elde etmek için yenilikçi bir araç olarak metin analizinin gücünü ve potansiyelini belirlemektedir. Böylece çalışmanın kurumsal performans değerlendirmesi ve strateji analizine ilişkin mevcut literatüre katkıda bulunması beklenmektedir.

Çalışmanın ikinci bölümünde konu ile ilgili daha önce yapılmış çalışmalara yer verilmiştir. Üçüncü bölümde, veri seti ve kullanılan yöntem hakkında bilgilerin yer aldığı metodoloji yer almaktadır. Dördüncü bölümde elde edilen ampirik bulgular tartışılmış ve son bölümde ise çalışmanın sonuç kısmına yer verilmiştir.

2. LİTERATÜR TARAMASI

Şirketlerin yıllık raporlarının analizi, bir şirketin performansı, büyüme beklentileri ve stratejisi hakkında sundukları değerli içgörüler nedeniyle onlarca yıldır ilgi konusu olmuştur (Botosan, 1997; Dempsey vd., 2012; Arianpoor, ve Sahoor, 2022). Çalışmalar, bu raporların yatırımcıların kararlarına destek olmada ve yatırım portföylerine olan güvenlerini şekillendirmede oynadığı kritik rolü vurgulamıştır (Qiu vd., 2014). Ayrıca, yıllık raporlar yalnızca yasal bir gereklilik olarak değil, şeffaflığı, hesap verebilirliği ve bilinçli karar almayı teşvik eden araçlar olarak görülmektedir (Okere vd., 2018).

Metin ve duyarlılık (duygu) analiz tekniklerinin ortaya çıkışı, yıllık raporların değerlendirilmesine yeni bir boyut kazandırmıştır. Bu teknolojilerden yararlanan araştırmacılar, büyük hacimli metinlerdeki kalıpları ve eğilimleri ayırt edebilmekte ve içinde aktarılan duyguyu tanımlayabilmektedirler (Loughran ve McDonald, 2011). Bu çalışmada kullanılan Gizli Dirichlet Ayırımı gibi bu tür metodolojiler, kurumsal strateji ve performansın daha incelikli yorumlarını mümkün kılarak, çeşitli kurumsal açıklama biçimlerini analiz etmek için kullanılmaktadır (Blei vd., 2003).

Duygu analizi, finansal raporların duyarlılığını veya tonunu ölçmek için uygulanabilmekte ve elde edilen bulgular genellikle duyarlılığı finansal piyasa tepkileriyle ilişkilendirilmektedir (Tetlock, 2007). Örneğin, Loughran ve McDonald (2016) tarafından yapılan bir araştırma, daha olumlu bir tona sahip finansal raporların genellikle daha yüksek hisse senedi getirileri ile ilişkili olduğunu göstermiştir. Bu korelasyon, finans endüstrisinde öngörü amaçlı olarak duyarlılık analizinin kullanılmasına yönelik ilgiyi artırmıştır (Bollen vd., 2011). Endonezya Borsası'nda işlem gören inşaat sektöründeki 152 işletmenin sürdürülebilirlik raporları üzerinden duyarlılık analizi uygulayan Harymawan (2020), sürdürülebilirlik raporlarındaki açıklamaların yaklaşık %68-69'unun olumlu bir duyarlılığa sahip olduğunu tespit etmiş ve elde edilen bu sonuçların yüksek düzeyde kurumsal hesap verilebilirliğe işaret ettiğini ifade etmiştir. Mućko'nun (2021) bulgularına göre, kurumsal sosyal sorumluluk (KSS) açıklamalarının duyarlılığı ile şirketin büyüklüğü arasında bir ilişki bulunmaktadır. Bununla birlikte, bu korelasyonun yalnızca pozitif bir ton söz konusu olduğunda pozitif olduğuna dikkat etmek önemlidir. Li vd. (2019), yıllık rapordan bir risk açıklama endeksi (RDI) türetmek için metin analizinden yararlanmış ve ardından bu endeksin firmalar içindeki yatırım verimliliği üzerindeki etkisini araştırmıştır. Çalışmanın bulgularında, MD&A departmanlarında risk açıklama sıklığı ile kurumsal yatırım verimliliği arasında pozitif bir ilişki olduğunu tespit edilmiştir.

Türkiye gibi gelişmekte olan piyasalardaki faaliyet raporlarının analizinde bu gibi makine öğrenimi tekniklerini uygulayan kapsamlı çalışmalara rastlanmamıştır. Bu çalışmada, Borsa İstanbul (BİST) 100 endeksinde faaliyet gösteren işletmelerin yıllık raporlarındaki kurumsal açıklamaların değerlendirilmesinde çeşitli metin analizi tekniklerinin kullanımıyla literatüre katkıda bulunmaktadır. Bu analitik tekniklerin potansiyelini vurgulayan araştırmamız, kurumsal performans ve strateji analizi alanındaki faydalarının giderek daha fazla tanınmasıyla aynı çizgidedir.

3. METHODOLOJİ

3.1. Örneklem Verisi

Bu çalışmanın kapsamı, Borsa İstanbul 100 endeksinde faaliyet gösteren işletmelerden oluşmaktadır. Analiz sonuçlarının tutarlı olabilmesi için gayrimenkul yatırım ortaklığı, holdingler ve bankalar gibi finansal işletmeler farklı düzenleme ve muhasebe uygulamalarına tabi olmalarından dolayı çalışmaya dahil edilmemiştir. Finansal işletmelerin örneklemden çıkarılmasıyla birlikte net olarak 60 işletme üzerinden analizler gerçekleştirilmiştir. İşletmelere ait yıllık raporlar Kamuyu Aydınlatma Platformu'ndan (KAP) elde edilmiştir. Elde edilen raporların analizi için Python yazılımı kullanılmıştır. Örneklem seçimine ait bilgiler tablo 1'de verilmiştir.

Tablo 1: Örneklem Seçimi

Açıklama	İşletme Sayısı
BİST100 Endeksinde Faaliyet Gösteren İşletmeler	100
BİST100 Endeksinde Mali İşletmeler	40
Net İşletme Sayısı	60

3.2. Konu Modelleme

Konu modellemesi, bir belge koleksiyonundan, koleksiyondaki bilgileri en iyi şekilde temsil eden bir kelime grubu (konu) bulma yöntemi olarak tanımlanabilir (Pröllochs ve Feuerriegel, 2020). Gizli değişkenlerin açığa çıkarılması, boyutun küçültülmesi ve verilerin gruplandırılması bu algoritmaların birincil odak noktalarıdır. Ayrıca metinsel materyallerde kelime kalıplarının tespit edilmesini sağlayan konu modellemesi, metin içerisindeki tema ve kavramlar hakkında önemli bilgiler verebilmektedir. Genellikle konular, insan tarafından seçilmiş örnekler olmak yerine tamamen doğal verilerden oluşturulduğu için tipik olarak denetimsiz makine öğrenimi teknikleri olarak

nitelendirilir. Bunun nedeni, konuların raporların içerisinde yer alan verilerden türetilmiş olmasıdır.

Belgelerin bir araya getirilerek kümeler oluşturulmasını ifade eden konu modellemesinde benzer içeriğe sahip belgeler birbirleriyle ilişkilendirilir ve gruplandırma sonucunda elde edilen bilgiler daha anlamlı hale gelir. Dolayısıyla metin verilerinin analizinde ve anlamlı bilgilerin çıkarılmasında önemli bir araç olarak kullanılmaktadır. Bu algoritmaların kullanımı, pazarlama, sosyal medya analizi, müşteri ilişkileri yönetimi gibi birçok alanda uygulanabilir. “Çeşitli sosyal medya sitelerinin gündeminde neler var?”, “Herhangi bir konu ile ilgili geçmiş ile şimdiki zaman arasında bir farklılık var mı?”, “Herhangi bir alışveriş sitesinde müşteriler firmanın hangi yönünü beğeniyor veya hangilerini beğenmiyor?” gibi sorular konu modellemenin çalışma alanlarından bazılarını temsil etmektedir (Ekinci & Omurca, 2017). Veri tabanlı kararların ve stratejik planlamaların temelini oluşturmak için de kullanılabilir. Ayrıca, konu modellemesi sayesinde büyük veri setlerindeki önemli kavramlar ve ilişkiler daha hızlı ve etkili bir şekilde tespit edilebilir.

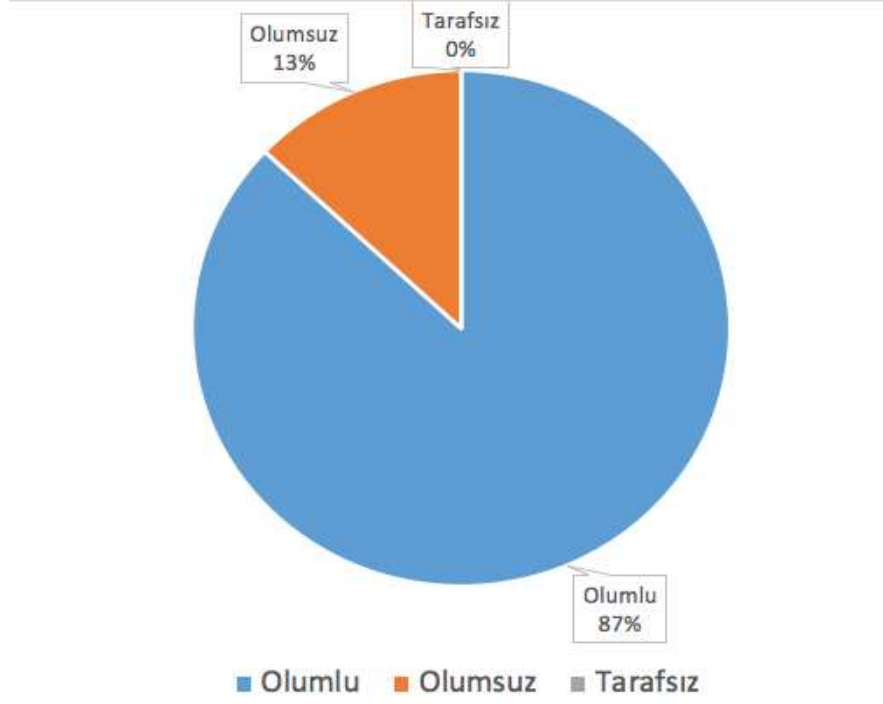
Konu modelleme teknikleri arasında sıkça kullanılan ve David Blei, Andrew Ng ve Micheal Jordan tarafından 2003 yılında tanıtılan GDA yönteminde, her belge, koleksiyon içerisinde bulunan konuların bir karışımı olarak görülür (Barde & Bainwad, 2017). Model, belgedeki her kelimenin konulardan birine atfedilebileceğini önermektedir (Ranta vd., 2022). Bundan yola çıkarak metin koleksiyonu içerisindeki konuları bulan tümevarımsal bir sürecin sonuçlarına dayalı olarak konulardaki kelime olasılıklarını çıkarır (Fiandrino ve Tonelli, 2021). Bu işlem uygulanırken her belgenin kendi dağılımına sahip olduğu ve her konunun kendi kelime dağılımına sahip olduğu ve bu kelimelerin rastgele yerleştirildiği varsayımlarına dayanır (Huang vd., 2017). Temel olarak belgelerin belirli konular sergilediği ve her konunun bir kelime dağılımından oluştuğu varsayılır (Barde & Bainwad, 2017). Bu varsayımlar altında GDA algoritması, metin belgelerini kelime torbası olarak kabul eder ve her belge için konu dağılımını ve her konunun kelime dağılımını tespit eder (Fiandrino ve Tonelli, 2021). Örneğin, bir konu "okul", "öğretmen" ve "derece" gibi çok sayıda kelime ve "ağaç" ve "kedi" gibi az sayıda kelime içeriyorsa, konunun muhtemelen eğitim olduğunu varsayılır.

4. BULGULAR

Araştırma konusu için seçilen toplam 60 işletme ve bu işletmelere ait yıllık raporların içerisindeki 2.347.594 kelime analize dahil edilmiştir. Bu raporların duygu analizi puanlarının dağılımı incelendiğinde %87 oranında olumlu, %13 oranında olumsuz içeriğe sahip olduğu tespit edilmiştir. Yüksek oranda

pozitif duygu tonu örneklemdaki işletmelerin faaliyet dönemini iyi değerlendirdiğini göstermektedir. Buna karşılık olarak elde edilen sonuçlar firmaların az da olsa faaliyet döneminde zorluklar ve olumsuz durumlarla karşılaştığını bildirmektedir. Duygu analizi sonucunda elde edilen bilgiler şekil 1’de sunulmuştur.

Şekil 1: Faaliyet raporlarının duyarlılık analizi



Duygu analizinin güvenilirliğini değerlendirmek için, aynı veri seti üzerinde birden fazla kez duygu analizi yaparak sonuçların tutarlılığını ölçülmüştür. Yapılan tekrarlı ölçümler sonucunda benzer sonuçlar elde edilmiştir.

Güvenilirlik Skorları

Test 1: %87 pozitif, %13 negatif (bulgularımızda belirtildiği gibi).

Test 2: %86 pozitif, %14 negatif.

Test 3: %87 pozitif, %13 negatif.

Raporlardaki kelimeler incelendiğinde "yönetim" kelimesi ilk sırada yer almaktadır. Yönetim kelimesini takiben sırasıyla "kurulu", "aralık", "şirket", "finansal", "yılında", "kurumsal" kelimeleri gelmektedir. Raporların içerisinde yer alan kelimeler şekil 2'de detaylı olarak gösterilmiştir.

Şekil 2: En yaygın ilk on kelime

Çalışmada raporlar içerisinde yer alan konuların belirlenmesi için konu modelleme tekniklerinden GDA kullanılmıştır. GDA, doğal dil işleme ve makine öğreniminde, özellikle büyük hacimli metin verilerini analiz etmek için güçlü ve yaygın olarak kullanılan bir istatistiksel modeldir. Bu da GDA'yı kitaplar, makaleler veya firmaların raporları gibi büyük metinsel verilerin tematik yapısını keşfetmek için önemli bir yöntem olarak göstermektedir. Analiz sonucunda yıllık raporlar içerisinde dört temel konu üzerinde ayrışma sağlanmıştır. Konularla bağlantılı olan 10 kelime ele alınmıştır. Tablo 2 yönetim ve finansal performans, tablo 3 ise kurumsal sorumluluk ve operasyonel içgörülere ilişkin kelimeleri göstermektedir.

Tablo 2: Yönetim ve Finansal Performans Kelimeleri

Konu 1 – Yönetim	Konu 2 - Finansal Performans
Kurul	Gelir
Yönetim	Kâr
Strateji	Finansal
Yönetişim	Büyüme
İcra	Yatırım
Komite	Pazar
Direktör	Sermaye
Politika	Kazanç
Denetim	Temettü
Liderlik	Hissedar

Tablo 3: Kurumsal Sorumluluk ve Operasyonel İçgörüler Kelimeleri

Konu 3- Kurumsal Sorumluluk	Konu 4 - Operasyonel İçgörüler
Sürdürülebilirlik	Operasyon
Çevre	Üretim
Toplum	Teknoloji
Sorumluluk	Yenilik
Sosyal	Verimlilik
Etik	Kalite
Çeşitlilik	Tedarik
Güvenlik	Süreç
Sağlık	Gelişim
Katılım	Kapasite

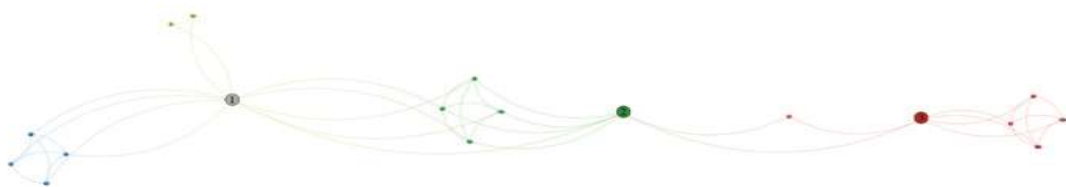
Elde edilen bulguların anlamlılık düzeylerini ölçmek için Content vector (C_V) uyum ölçüsü kullanılmıştır. Konulara ilişkin uyum skorları aşağıdaki gibidir. Sırasıyla konuların uyum skorları 0,51, 0,67, 0,63 ve 0,48'dir. Konu 2 (Finansal Performans) en yüksek uyum skoruna (0.67) sahiptir, bu da bu konunun kelimelerinin oldukça uyumlu olduğunu ve şirketlerin finansal yönleriyle ilgili açık, yorumlanabilir bir tema oluşturduğunu göstermektedir. Konu 1 (Yönetim ve Yönetişim) ve Konu 3 (Kurumsal Sorumluluk) orta düzeyde uyum skorlarına (sırasıyla 0.51 ve 0.63) sahiptir, bu da oldukça iyi bir tematik yapıya işaret etmektedir. Konu 4 (Operasyonel İçgörüler) ise en düşük uyum skoruna (0.48) sahiptir, bu da bu konunun kelimelerinin daha az uyumlu olduğunu ve tema olarak diğerlerine göre daha az net olduğunu işaret etmektedir.

Verilerin görsel biçimde sunulması tablo ve metin gibi diğer yöntemlere kıyasla elde edilen sonuçların anlaşılmasını basitleştirmektedir (Bashri & Kusumaningrum, 2017). Örneklemdaki veri miktarı çok büyük olmasa bile görselleştirme sayesinde elde edilen sonuçlar daha basit bir şekilde aktarılmakta ve belgeler içerisinde temel konuların belirlenmesi kolaylaşmaktadır. Kelime bulutu kelimelerin belgeler içerisinde görülme sıklığının görsel olarak temsilini sunmaktadır. Yazı boyutu, kelimenin görülme sıklığını belirlemektedir. Yazı boyutu ne kadar büyükse kelime sıklığı o kadar büyük olduğu varsayılır. Şekil 3'de analiz sürecinde yaygın olarak çıkan kelimelerin bulut görselleştirmesi sunulmuştur.

Şekil 3: Kelime bulutu görselleştirmesi

Kelime bulutuna göre görsel içerisinde en büyük kelime grubu “yönetim kurulu” olmuştur. Buda rapor içerisinde beklenen bir durumdur. Ayrıca firmaların planlarına yönelik açıklamaları baz alındığında sadece “uzun vadeli” ifadesi bulut görselleştirmesinin içerisine girmiştir.

Çalışmada kelimeler arası ilişki ağları da incelenmiştir. Şekil 4’de görüldüğü üzere, üç ana faktör üzerinde faaliyet raporlarının oluşturulduğu anlaşılmaktadır. Bunlar, yönetim kurulu (1), finansal durum (2), kurumsal yaklaşım (3) olarak tespit edilmiştir.

Şekil 4: Anahtar kelimeler arası ilişki ağları

Şekil 4 incelendiğinde üç ana faktör üzerinde en önemli faktör olan yönetim kurulu faktörünün görselde görüldüğü şekilde iki farklı kola ayrıldığı görülmekte ve 3. ana faktör olan kurumsal yaklaşımın yönetim kurulu ile bağlantısı olmadığı görülmektedir. Kurumsal yaklaşımın yönetim kurulu ile bağlantısı finansal durum üzerinden gerçekleşmektedir.

Rapor içerisindeki kelimeler tek tek ele alındığında yeterli bilgi elde edilemeyebilir. Dolayısıyla kelimeleri bileşik olarak ele almak daha doğru bilgiler verebilmektedir. Dolayısıyla bu çalışmada daha detaylı bir bilginin

elde edilmesi için art arda gelen ifadelerin görülme sıklıklarının tespit eden N-gram analizi de uygulanmıştır. Çalışmadaki 60 işletmenin faaliyet raporları içerisinde yer alan metinlerde art arda gelen kelimeler veya harf dizilerinin sıklıklarına bakarak elde edilen sonuçlar, ikili ve üçlü kelime dizileri şeklinde tablo 4’de verilmiştir.

Tablo 4: İkili ve Üçlü İfadelerin Görülme Sıklıkları

	Yönetim - Kurulu	Ortalama	Kurumsal - Yönetim	Ortalama	Yatırım - anlaşmaları	Ortalama
İkili ifade	8283	138,05	3385	56,41	746	12,46
	Yönetim - Kurulu - Üyesi	Ortalama	Karlılık - dönemine - ait	Ortalama	Finansal başarı üzerine	Ortalama
Üçlü ifade	1270	21,16	834	13,9	307	5,11

İkili ifadelere bakıldığında “Yönetim kurulu” metin dizisi, toplamda atmış rapor içerisinde 8283 kez, ortalama rapor başına ise 138 kez kullanılmıştır. Üçlü ifadelere bakıldığında “Yönetim Kurulu Üyesi” metin dizisi, toplamda atmış rapor içerisinde 1270 kez, ortalama rapor başına ise 21 kez kullanılmıştır. İkili ifadelere bakıldığında, “yatırım anlaşmaları” 746 kez tüm raporlarda kullanılırken, ortalama rapor başına 12 kez ifade edilmiştir. Buna karşın üçlü ifadelere bakıldığında ise “finansal başarı üzerine” metin dizisi toplamda 307 kez kullanılmış, ortalama rapor başına 5 kez ifade edilmiştir. Buradan anlaşıldığı üzere, firmalar raporlarında yatırım anlaşmalarını rapor başına, finansal başarı’dan daha fazla bahsetmişlerdir.

5. SONUÇ, TARTIŞMA VE ÖNERİLER

Bilimsel makalelerden firmaların yayınladığı raporlara kadar her gün üretilen metin verilerinin büyüklüğü anlamlı bilgilerin çıkarılmasında güçlükler oluşturmakta ve gelişmiş yöntemler gerektirmektedir. Makine öğrenimi teknikleri ise bu gibi büyük çaptaki verilerin analiz edilmesinde önemli bir rol oynamaktadır. Bu çalışmada, Borsa İstanbul (BIST) 100 endeksinde yer alan şirketlerin yıllık raporları metin analizi yöntemi ile incelenmiştir. Bulgular, bu tür tekniklerin finansal belgelerden değerli içgörüler elde etmedeki faydasını vurgulayan literatüre katkıda bulunmaktadır (Loughran ve McDonald, 2011; Loughran ve McDonald, 2016; Mućko, 2021; Li vd., 2019). Gizli Dirichlet Ayırımı yönteminin uygulanması, faaliyet raporlarındaki kalıpları ve temaların tespit edilmesinde, duygu analizinde ise raporlarda yer alan genel duygu tonu hakkında önemli bilgiler elde edilmesinde etkili olmaktadır.

Bu araştırmanın amacı, firmaların faaliyetleri ile ilgili çok sayıda raporu inceleyerek, yıllık raporların içerisinde yer alan temel konuların tespit edilmesi ve bu raporlar içerisindeki temel duyguların belirlenmesidir. Bu amaca ulaşmak için çalışmada, Borsa İstanbul (BIST) 100 endeksinde faaliyet gösteren firmaların İngilizce olarak yayınlanan altmış faaliyet raporunu ele alınmış ve konu modelleme tekniğini kullanarak duygu analizi, konu modelleme ve N-gram yöntemleri olmak üzere bir metin analizi gerçekleştirilmiştir. Elde edilen sonuçlara göre yıllık raporlardaki açıklamaların yaklaşık %87'sinin pozitif bir görünüm sergilemesiyle iyimser bir tonda olduğunu göstermektedir. Bu pozitif duygu durumu, örnekleme dahil edilen firmaların ilgili faaliyet döneminde olumlu bir algıya işaret etmektedir. Buna karşılık olarak raporların içerisinde yer alan açıklamaların %13'ü olumsuz bir duygu aktarmaktadır. Bu da raporların çoğunluk olarak olumlu olmasına rağmen, firmaların faaliyet döneminde zorlukların veya endişe verici durumların da mevcut olduğunu göstermektedir. GDA analizi sonucunda ise raporların içerisinde yönetim, finansal performans, kurumsal sorumluluk ve operasyonel içgörüler olmak üzere dört farklı konuda ayrışma sağladığı görülmektedir.

Bu çalışmanın sonuçlarından elde edildiği gibi makine öğrenimi destekli metin analizi yöntemleri, kurumsal raporlardan anlamlı içgörüler elde edilmesinde kullanışlı araçlar olarak potansiyelinin altını çizmektedir. Bu yöntemler sadece metin içerisinde duygu durumunu ve gizli konuların belirlenmesinin haricinde kurumsal performans ve stratejiyi anlamak için önemli bilgiler sunmaktadır. Sürekli gelişen ve büyük verinin mevcut olduğu ortamda, firma açıklamalarına metin analizi uygulayarak elde edilen içgörüler yöneticilerin, yatırımcıların ve diğer paydaşların karar alımlarında önemli destek sağlayabilir. Ancak, faaliyet raporlarında ortaya konan duyarlılık, bir şirketin performansını ve stratejisini anlamanın yalnızca bir yönüdür. Çalışmada öncelikle buna odaklanmış olsa da, gelecekteki araştırmalarda faaliyet raporlarının diğer unsurlarını daha derinlemesine inceleyebilir ve belirlenen duyarlılığın diğer finansal performans göstergeleri ve büyüme beklentileriyle nasıl ilişkili olduğu araştırılabilir.

Sonuç olarak, araştırmada, kurumsal faaliyet raporlarından anlamlı içgörüler elde etmek için metin analizinin faydalı bir araç olarak kullanılacağı anlaşılmaktadır. Bu tür analitik teknikler, kurumsal performans ve stratejiyi anlamak için yeni bir yaklaşım sunmakta ve nihayetinde paydaşların bilinçli kararlar almaları için değerli bir araç olarak hizmet etmektedir.

Kaynakça

- Arianpoor, A., & Sahoo, Z. (2022). The impact of business strategy and annual report readability on financial reporting quality. *Journal of Asia Business Studies*, 17(3), 598-616.
- Barde, B. V., & Bainwad, A. M. (2017). An overview of topic modeling methods and tools. In 2017 International Conference on Intelligent Computing and Control Systems (ICICCS) (745-750)
- Barth, M. E., Beaver, W. H., & Landsman, W. R. (2001). The relevance of the value relevance literature for financial accounting standard setting: another view. *Journal of accounting and economics*, 31(1-3), 77-104.
- Bashri, M. F., & Kusumaningrum, R. (2017). Sentiment analysis using latent dirichlet allocation and topic polarity wordcloud visualization. 5th International Conference on Information and Communication Technology, 1 – 5.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 993-1022.
- Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1-8.
- Botosan, C. A. (1997). Disclosure level and the cost of equity capital. *The Accounting Review*, 72(3), 323-349.
- Dempsey, S. J., Harrison, D. M., Luchtenberg, K. F., & Seiler, M. J. (2012). Financial opacity and firm performance: the readability of REIT annual reports. *The Journal of Real Estate Finance and Economics*, 45, 450-470.
- Dyer, T., Lang, M., & Stice-Lawrence, L. (2017). The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics*, 64(2-3), 221-245.
- Ekinci, E., & Omurca, S. İ. (2017). Ürün özelliklerinin konu modelleme yöntemi ile çıkartılması. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 9(1), 51-58.
- Feuerriegel, S., & Pröllochs, N. (2021). Investor reaction to financial disclosures across topics: An application of latent Dirichlet allocation. *Decision Sciences*, 52(3), 608-628.
- Fiandrino, S., & Tonelli, A. (2021). A text-mining analysis on the review of the non-financial reporting directive: bringing value creation for stakeholders into accounting. *Sustainability*, 1-18.
- Harymawan, I., Nasih, M., Ratri, M. C., Soeprajitno, R. R., & Shafie, R. (2020). Sentiment analysis trend on sustainability reporting in indonesia: evidence from construction industry. *Journal of Security And Sustainability Issues*, 9(3), 1017-1024.
- Huang, A. H., Lehavy, R., Zang, A. Y., & Zeng, R. (2017). Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach. *Management Science*, 2833-2855.
- Lewis, C., & Young, S. (2019). Fad or future? Automated analysis of financial text and its implications for corporate reporting. *Accounting and Business Research*, 49(5), 587-615.
- Li, Y., He, J., & Xiao, M. (2019). Risk disclosure in annual reports and corporate investment efficiency. *International Review of Economics & Finance*, 63, 138-151.
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35-65.

- Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4), 1187-1230.
- Mučko, P. (2021). Sentiment analysis of CSR disclosures in annual reports of EU companies. *Procedia Computer Science*, 192, 3351-3359.
- Pröllochs, N., & Feuerriegel, S. (2020). Business Analytics for strategic management: identifying and assessing corporate challenges via topic modeling. *Information & Management*, 57(1), 1-13.
- Ranta, M., Ylinen, M., & Järvenpää, M. (2022). Machine learning in management accounting research: literature review and pathways for the future. *European Accounting Review*, 1-30.
- Okere, W., Eluyela, F. D., Basse, I., & Ajetunmbi, O. (2018). Public sector accounting standards and quality of financial reporting: A case of Ogun state government administration in Nigeria. *Business and Management Research Journal*, 7(7), 76-81.
- Qiu, X. Y., Srinivasan, P., & Hu, Y. (2014). Supervised learning models to predict firm performance with annual reports: An empirical study. *Journal of the Association for Information Science and Technology*, 65(2), 400-413.
- Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3), 1139-1168.
- Yuthas, K., Rogers, R., & Dillard, J. F. (2002). Communicative action and corporate annual reports. *Journal of Business Ethics*, 141-157.