



*Istanbul Finance Congress (IFC), November 2-3, 2017, Istanbul, Turkey*

## USAGE OF DATA MINING FOR EVALUATION OF BORSA İSTANBUL REGISTERED COMPANIES' FINANCIAL STRUCTURE

DOI:10.17261/Pressacademia.2017.742

PAP-IFC- V.6-2017(6)-p.34-38

Mehmet Ozkan<sup>1</sup>, Levent Boran<sup>2</sup>

<sup>1</sup>Marmara University, Istanbul, Turkey. [mozkan@marmara.edu.tr](mailto:mozkan@marmara.edu.tr)

<sup>2</sup>Nişantaşı University, Istanbul, Turkey. [levent.boran@nisantasi.edu.tr](mailto:levent.boran@nisantasi.edu.tr)

### To cite this document

Ozkan, M., Boran, L. (2017). Usage of data mining for evaluation of Borsa İstanbul registered companies' financial structure. PressAcademia Procedia (PAP), V.6., p.34-38.

Permanant link to this document: <http://doi.org/10.17261/Pressacademia.2017.742>

Copyright: Published by PressAcademia and limited licenced re-use rights only.

### ABSTRACT

**Objective-** Financial analysis are mostly done for evaluation of companies' financing and investment needs with traditional analysis methods such as vertical analysis, horizontal analysis and ratio analysis. Although these traditional methods support the analyst for single company evaluation, they are inefficient while questioning many companies. Therefore, decision makers face time-consuming problem when they evaluate hundreds of companies, which are necessary for profit maximization, cash flow maximization and risk mitigation etc. It is aimed to define a new tool for financial analysis in this study.

**Methodology-** BIST Manufacturing Sector registered 190 companies for year 2015 and 173 companies for year 2016 are analyzed. Some liquidity ratios, fiscal ratios, operational ratios and profitability ratios are calculated and outlier companies are decided. Data Mining is the one of the most important data processing tool. It can be used for clustering the data, classification the data and defining variables that have similar behaviors. It is tried to define a new financial analysis technique with combination of ratio analysis and data mining. In this study, outlier detection and some clustering algorithms are applied on BIST Manufacturing Sector registered companies.

**Findings-** BIST Manufacturing Sector registered 121 of 190 companies for year 2015 and 127 of 173 companies for year 2016 are decided as outlier companies. These outlier companies might be evaluated for sectorel researches or fraud detection etc. Companies are divided two clusters with and without outlier companies for year 2015. In addition, companies are divided four clusters with outlier companies and two clusters without outlier companies for year 2016. Differences between the number of clusters and cluster characteristics are related to economical conditions.

**Conclusion-** In conclusion, Data Mining Techniques can be used as financial analysis method, especially when we need to analyze many companies' financial situation at the same time. It is considered that sector characteristics, global and local developments would indicate meaningful correlations with outlier companies. Besides that, it is determined that universal thresholds values for financial ratios (e.g. current ratio 2) are different for our country. These values are calculated for our country and evaluated with sectorel, global and local factors.

**Keywords:** Financial statement analysis, ratio analysis, data mining, clustering analysis, EM algorithm.

**JEL Codes:** D53, G10, G17, O16

## VERİ MADENCİLİĞİNİN BIST'E KAYITLI İŞLETMELERİN MALİ YAPILARININ DEĞERLENDİRİLMESİNDE KULLANILMASI

### ÖZET

**Amaç-** Finansal analiz çalışmaları, genellikle işletmelerin finansman ihtiyaçları ve yatırım yapabilmelerinin değerlendirilmesini amaçlamakta olup, günümüzde kullanılan geleneksel finansal analiz yöntemleri arasında çok sıklıkla kullanılan dikey analiz, yatay analiz ve oran analizi yöntemlerinin farklı zayıf ve güçlü yönleri bulunmaktadır. Bu yöntemler; muhataplara işletmelerin finansal durumlarıyla ilgili işletme bazlı olarak bilgi sağlamakla birlikte, çok fazla işletmenin bir arada değerlendirilip karşılaştırılmasında yetersiz kalmaktadırlar. Çalışmamız karar vericiler için bu eksikliği giderici bir araç tanımlamayı amaçlamaktadır.

**Yöntem-** Borsa İstanbul İmalat Sektörüne kayıtlı 2015 yılından 190 adet şirket, 2016 yılından 173 adet şirket çalışmaya konu edilmiştir. Şirketlerin finansal tabloları üzerinden bazı likidite oranları, mali yapı oranları, faaliyet oranları ve karlılık oranları hesaplanmış olup, bu

oranlar üzerinden aykırı şirketler belirlenmiştir. Veri Madenciliği Teknikleri, son dönemde ortaya konulmuş en önemli veri değerlendirme araçlarından biri olup; verinin önceden belirlenmiş sınıflara göre oluşturulması, verinin dağılımına göre kümelere ayrıştırılması ve sonuca aynı yönde etki eden değişkenlerin belirlenmesi gibi amaçlara hizmet etmektedirler. Çalışmamızda Borsa İstanbul İmalat Sektörü'ne kayıtlı şirketlerin finansal oranlarına aykırılık analiz ve kümeleme analizi teknikleri uygulanmıştır.

**Bulgular-** Borsa İstanbul İmalat Sektörü'ne kayıtlı 2015 yılı için 190 adet şirket, 2016 yılına için 173 adet şirket çalışmaya konu edilmiş olup, 2015 yılında 121 adet şirket aykırı, 2016 yılında 127 adet şirket aykırı olarak belirlenmiştir. Aykırı şirketler bazı sektörel araştırmalara, hile araştırmalarına konu edilebilir. 2015 yılı için hesaplanan finansal oranlara kümeleme analizi uygulandığında aykırı şirketlerin elenmesi ve elenmemesi durumlarında ikişer küme oluştuğu, 2016 yılı için ise aykırılıklar elenmemiş finansal oranlarla dört, elenmiş finansal oranlarla iki adet küme oluştuğu görülmüştür. Aradaki farklılaşma ve küme karakteristiklerinin günün ekonomik durumlarıyla ilişkili olduğunu düşünmekteyiz.

**Sonuç-** Sonuç olarak, Veri Madenciliği Tekniklerinin özellikle çok fazla sayıda şirketin finansal durumunun karşılaştırıldığı analizlerde araç olarak kullanılabilirliği görüldü. Aykırı şirketler arasında küresel, yerel gelişmeler ve sektör özellikleriyle anlamlı bağlantılar bulunabileceği düşünülmektedir. Ayrıca kümeleme analizinin uygulanması yoluyla finansal oranlar için tanımlanmış eşik değerlerin (örneğin cari oran için 2) ülkemiz için farklı olduğu belirlenerek ve bu eşik değerlere sektörel, küresel ve yerel faktörlerin etkisi yorumlanmıştır.

**Anahtar Kelimeler:** Finansal tablo analizi, oran analizi, veri madenciliği, kümeleme analizi, EM algoritması.

**JEL Kodları:** D53, G10, G17, O16

## 1.GİRİŞ

Finansal Tablolar; şirketler üzerinde verilecek yatırım, borç verme, vergilendirme gibi kararlarda işletme yöneticileri, hissedarlar, tahvil sahipleri, analistler, tedarikçiler, ödünç para veren kuruluşlar, çalışanlar, işçi sendikaları, kanun koyucu otoriteler ve kamu tarafından kullanılan önemli araçlardır. Karar verme sürecinin önemli bir parçası belirsizliğin ortadan kaldırılması olup, bu noktada Finansal Tablolar üzerinde tanımlı araçların kullanılması yoluyla ilgili taraflara faydalı bilgiler sağlarlar. Finansal Tabloları ilgili taraflar farklı amaçlarla incelemekte olup, örneğin kredi sağlayıcılar borç ödeme kapasitesiyle, yatırımcılar karlılığı ile yöneticiler ise operasyonel verimliliğiyle ve gelir gider dengeleriyle daha fazla ilgilenmektedirler. Yatırımcı için işletmelerin performansını ayırt etmek kötü bir yatırıma değerinden fazlasını vermemek için önemlidir. [1] Şirketler, faaliyet gösterdikleri sektöre ve şirketin içinde bulunduğu duruma göre farklı finansman yöntemlerini kullanmaktadırlar. Her bir finansman yönteminin kendine göre avantaj ve dezavantajları bulunmakta olup, en uygunun seçimi şirketler için tanımlanmış bir problemdir. [3] Dolayısıyla tüm tarafların farklı bakış açılarının bulunması kaçınılmazdır.

Günümüzde, çeşitli kanallardan toplanan çok fazla veri biriktirilmekte olup, verinin anlamlandırılmadan sadece biriktiriliyor olması gereksiz bir faaliyet hükmündedir. Verinin işlenmesi ve tecrübe ile yönlendirilerek bilgiye dönüştürülmesi yoluyla faydalanılabilir forma dönüştürülmesi gerekmektedir ve bu durumda Veri Madenciliği Teknikleri kullanım alanı bulmaktadır. Veri Madenciliği Teknikleri, son dönemde ortaya konulmuş en önemli veri değerlendirme araçlarından biri olup; verinin önceden belirlenmiş sınıflara göre oluşturulması, verinin dağılımına göre kümelere ayrıştırılması ve sonuca aynı yönde etki eden değişkenlerin belirlenmesi gibi amaçlara hizmet etmektedirler.

Veri Madenciliği uygulamaları; bu alanda bilgi birikimi ve alan uzmanlığı gerektirmektedir. Veri Madenciliği, standart teknikler kullanılarak veya herhangi bir teknik kullanmaksızın kolayca analiz edilebilecek küçük veri setleri için maliyeti yüksek olması nedeniyle tercih edilmemektedir. Doğru sonuçlar; mümkün olduğunca kusursuz bir veri üzerine uygun algoritmaların uygulanması ve elde edilen sonuçların kusursuz bir şekilde yorumlanmasına bağlıdır. Veri Madenciliği uygulamaların iki önemli özelliği çok yoğun ve hızla akan bir veri seti üzerinde çalışılması ve sonucun önceden tahmin edilebilir olmamasıdır.

Buradan hareketle bu çalışmamızda, Veri Madenciliği Teknikleri finansal oranlardan bazılarını uygulanarak Veri Madenciliği Teknikleri yeni bir finansal analiz yöntemi olup olamayacağı irdelenmektedir. Fazla sayıda işletmenin birbirleriyle karşılaştırılması şeklinde yapılacak çalışmalarda yoğun veriyi işleyebilen Veri Madenciliği Tekniklerinin kullanılması sonuca ulaşmayı kolaylaştırmıştır. Çalışmamızda Veri Madenciliği Tekniklerinin bir finansal analiz aracı olarak; Borsa İstanbul'a kayıtlı şirketlerin finansal durumlarının değerlendirilmesinde kullanılması ile işletme ve finans yazınına hem uygulamacılar hem de teorisyenler açısından katkı sağlayacağı öngörülmektedir.

## 2.LİTERATÜR İNCELEMESİ

Ekonomik değeri olan herhangi bir maddeyi yeryüzüne çıkarıp, o maddeyi paraya dönüştürme işlemine "madencilik" denir. Elmas, bor, kömür gibi, farklı madenlerde olduğu gibi; bu maddelerin madenlerden çıkarılması da farklı şekillerde gerçekleşmektedir.[2] Veri Madenciliği de, tıpkı madencilik tanımında olduğu gibi çok miktardaki veri arasından değerli olan bilgiyi çıkartmayı amaçlamakta olup, günümüzde sağlanan araçlarla veriden bilginin elde edilmesinin ilerisine gidilerek gerekli aksiyonların tanımlanması otomatize edilmiştir. Ayrıca verinin içerisinden değerli bilginin çıkartılması farklı analiz teknikleri ve bu teknikler kapsamında tanımlanmış algoritmaların uygun olanının seçilmesi ve Veri Madenciliği prensipleri çerçevesinde uygulanması yoluyla yapılmaktadır. Veri Madenciliği sadece veriden anlam çıkartmayı amaçlayan bir araç olarak yorumlanmamalıdır. Çünkü Veri Madenciliği'nin ayırt edici karakteristiği diğer yöntemlerin bir modele dayanmasının aksine, kendisinin veriye dayanıyor olmasıdır. İstatistikte, araştırmacı sık sık daha küçük boyutta verinin içinde problemin çözümünü arayarak güvenli tahminler yapmakta iken; Veri Madenciliği'nde, büyük miktarda veri ile küçük bir model oluşturup verinin en iyi şekilde tanımlanması amaçlanmaktadır[4]

Her geçen gün bilgisayarlar tarafından kaydedilen veri miktarı artmakta olup, artan veri ekonomistler, istatistikçiler, tahminleme yapanlar, iletişim mühendisleri, brokerler, yatırımcılar gibi çok değişik kitleler tarafından bu veriler kullanılmaktadır. Veri miktarının her 20 yılda bir katlandığı düşünülmekte olup, verinin birikme hızı giderek artmaktadır. [11] Bu durum yoğun veri içerisinden bilgiyi çıkartan araçların günümüz iş hayatında daha önemli hale gelmesine neden olmuştur. Veri Madenciliği'nin üniversiteler ve araştırma laboratuvarlarının dışına

çıkması, iş hayatının rekabet koşulları ve müşteri beklentileri gibi gereksinimler gereği bu kadar yaygınlaşmasının önemli nedenleri bulunmakta olup, bu nedenlerin bir kısmı şunlardır: [10]

- Modern hesaplama araçları ve metodları ile hesaplama kapasitesi ve veri kaydetme kapasitesinin çok fazla artması,
- Farklı entegre Veri Madenciliği ve istatistiksel yazılım paketlerinin sayılarının her geçen gün artması ve bahsi geçen algoritmaların kullanıcı dostu yazılımlar sayesinde birbirine bağlanabilmesinin sonucu olarak çıktı kalitesinin artması,
- Karar verme alanındaki gelişmelerin Veri Madenciliği üzerine olumlu etkileri ve bunun sonucu olarak üretim sürecinde periyodik bilgi çıktılarının otomatik bir şekilde son kullanıcıya ulaştırılması sayesinde otomatik olay tetiklemeleri.

Veri Madenciliği; istatistik, informatik ve matematikçi bilim insanlarının kullanım alanlarında birlikte çalışmalarında disiplinler arası bir araştırma alanı olup, temel araştırma araçları istatistiksel yöntemler, yapay zeka gibi daha çok kullanım bulan görsel tekniklerdir. [7] Veri Madenciliği'nin ilk kullanımı 1990'lı yıllara kadar gitmekte olup, 1998 yılı Mayıs ayında CIO Magazine Dergisi Bank of America'nın müşterilerle ürünleri eşleştirmek için Veri Madenciliği'ni kullandığını ve müşteri temsilcilerini yeni ürünler hakkında bilgilendirmek üzere bu çıktılar doğrultusunda yönlendirdiğini yazmıştır. Eski Amerika Birleşik Devletleri başkanı Bill Clinton 2002 yılında ülke güvenliği alanında sakıncalı kişilerin tespitinde yine Veri Madenciliği'nin kullanıldığını söylemiştir. Amerika Birleşik Devletleri'nde çocuklar arasında görülen ortalama yılda 3000 beyin tümörü vakası Veri Madenciliği Teknikleriyle incelenmiş ve tümörlerin sınıflandırması bu şekilde yapılmıştır. Spor alanında da Amerika Birleşik Devletleri profesyonel basketbol liginde yer alan Boston Celtics, New York Knicks gibi takımların şut, pas, top kapma gibi istatistikleri olarak Veri Madenciliği Teknikleriyle değerlendirdikleri ve taktiklerini bu çerçevede oluşturdukları bilinmektedir.[5]

Veri Madenciliğinin önemini artıran noktalardan biri de elde edilecek bilginin "Önceden Bilinmiyor" olmasıdır. Veri madenciliği sonunda ulaşılabilecek bilginin önceden bilinmiyor olmasından kasıt, elde edilecek sonucun tahmin edilememesi anlamını taşımaktadır. Zaten öngörülebilir, beklenen sonuçlar için Veri Madenciliği kullanmak ekonomik olmayacaktır. Ayrıca Veri Madenciliği öngörülürken, öngörülen ya da başka yöntemlerle çıkarılmış sonuçları ispatlamak amacıyla yönelik olarak kullanılacak bir araç değildir. Veri Madenciliği daha önce hiç akla gelmemiş, düşünülmemiş sonuçları ortaya çıkarmasıyla diğer yöntemlerden farklılık göstermektedir. Günümüz rekabet koşulları göz önüne alındığında, işletmelerin başarısı için daha önce hiç kimsenin düşünmediğini düşünmek ve uygulamak kritik bir noktadır. [9] Verinin anlamlandırılması süreci için problemin tanımlanması, verinin hazırlanması, analizin yapılması ve sonucun değerlendirilmesi adımları tanımları tanımlanmıştır. [6]

Veri Madenciliği Teknikleri, kullanım amaçlarına göre, isimlendirilmiş olup; istatistiksel yöntemlerden türetilmiş algoritmalarla desteklenmiştir. Kullanılan verinin niteliğine ve beklenen sonuçlara göre kullanılacak algoritma seçimi yapılması gerekmekte olup, doğru algoritma seçiminde veri bilimcisinin deneyimi oldukça önemlidir. Veri Madenciliği Tekniklerinin kullanım amaçlarına göre sınıflandırılması aşağıda verilmiştir: [8]

- **Sınıflandırma:** Bu tip problemlerde, veritabanındaki her bir kayıt önceden tanımlanmış sonlu belirli bir sınıf veya kategori etiketi altında toplanır. Örneğin bir banka, her bir müşterisini yeni bir kredi kartıyla ilgilenip ilgilenmeyeceğine göre sınıflandırmak isteyebilir.
- **Öngörme:** Bu tip problemlerde, sürekli değişkenlerin bilinmeyen değeri tahmin edilir. Örneğin, vergi otoriteleri, ev sahiplerinin gelirlerini tahmin etmekle ilgilenirler. Öngörme probleminin mümkün olan çıktı sayısı tanımlı gereği sonsuzdur.
- **Tahmin:** Gelecekteki değerin tahmin edilmesidir. Bu konudaki tipik örnekler hisse senedi fiyatlarının tahmini, ticari mal fiyatlarının tahmini ve hava kirliliği endeksinin gelecekteki değerinin tahmini şeklinde sıralanabilir.
- **Kümeleme:** Kümeleme, segmentasyon olarak bilinmekte olup, heterojen bir popülasyonun daha fazla homojen alt grup veya kümelere bölünmesi işidir. Bazı kaynaklarda, veri içerisinde benzerlik adasını bulmak şeklinde de isimlendirilmektedir. En tipik örnekleri müşteri ve pazarların birbirine benzer veya farklı özellikler gösterenlerinin belirlenmesidir. Çok büyük veri içerisinde kümeleme verinin küçük gruplara ayrılmasına basitleştirilmesini sağlayan ve veriye ilişkin hipotezin ortaya konulmasını kolaylaştıran bir araçtır.
- **Benzerlik Analizi:** Benzerlik analizinin diğer isimleri pazar sepeti analizi ve ilişki analizidir. Buradaki amaç birlikte giden şeyleri bulmaktır. Bu şeyler, ürünler, işlemler, operasyon sıraları veya veritabanına yazılan objeler olabilirler. En tipik örneği bir süpermarkette birlikte alınan ürünlerin tespiti için yapılan pazar sepeti analizidir. Örneğin tavuk ve barbekü sosunun birlikte satın alınıp alınmadığı ile ilgileniyor olabiliriz.

### 3.VERİ, YÖNTEM, BULGULAR, ANALİZ

Uygulamamızda Borsa İstanbul'a kayıtlı 2015 yılı için 190 adet, 2016 yılı için 173 adet İmalat Sanayi şirketinin mali tabloları kullanılmış olup, inceleme iki aşamada gerçekleştirilmiştir. İlk aşamada işletmelerin belirlenen finansal oran bakımından aykırı olanlar ayıklanmış, ikinci aşamada ise kalan işletmeler arasında homojen kümeler oluşturulmaya çalışılmıştır. Aykırı firmaların ayıklanmasının nedeni mümkün olduğunca homojen kümeler oluşturma çabası olup, kümeleme analizi yapılmasının nedeni ise mevcutta tanımlanmış finansal oran bazlı sınır değerlerin (örneğin cari oran 2) piyasa gerçeklerini yeterince yansıtmamasıdır. Ayrıca kontrol noktası olarak, kümeleme analizi tüm şirketlerin finansal oranlarına (aykırı şirketler dahil) uygulanmıştır.

Değerlendirmeye öncelikle mali yapı oranlarıyla başlanmış olup, oluşan kümelerin her bir finansal oran için çok geniş değer aralıklarına dağıldığı görülmüştür. Ayrıca finansal açıdan değerlendirildiğinde bilançonun sadece pasif tarafının değerlendirilmesi ve gelir tablosunun göz önüne alınmaması yanıltıcı ve eksik olacaktır. Bu nedenle de bilançonun her iki tarafını (aktif / pasif) ve gelir tablosunu içerecek şekilde aşağıda yer alan oranlar belirlenmiştir:

#### Likidite Oranları

- Cari Oran = Dönen Varlıklar / Kısa Vadeli Yabancı Kaynaklar
- Asit Test Oran = (Dönen Varlıklar - Stoklar) / Kısa Vadeli Yabancı Kaynaklar
- Hazır Değerler Oranı = Hazır Değerler / Kısa Vadeli Yabancı Kaynaklar

- Net İşletme Sermayesi Oranı = KVK / Dönen Varlıklar

#### Mali Yapı Oranları

- Borçları Özkaynaklara Oranı = ( KVK + UVVK ) / Özkaynaklar
- Finansal Kaldıraç = ( KVK + UVVK ) / Kaynak Toplamı
- Borçlanma Katsayısı =Yabancı Kaynaklar / Öz Kaynaklar

#### Faaliyet Oranları

- Alacakların Devir Hızı = Net Satışlar / Ortalama Ticari Alacaklar
- Nakit Devir Hızı = Net Satışlar / Ortalama Hazır Değerler
- Aktif Devir Hızı = Net Satışlar / Ortalama Varlıklar
- Alacak Tahsil Süresi= 365 / Alacak Devir Hızı
- Stok Devir Hızı=SMM./Toplam Stoklar

#### Karlılık Oranları

- Brüt Kar Oranı = Brüt Satış Karı / Net Satışlar
- Faaliyet Karı Oranı = Faaliyet Karı / Net Satışlar
- Faaliyet Kaldırıcı = Faiz ve Vergi Öncesi Kar / Satışlar
- Finansal Karlılık = Dönem Karı / Öz Kaynaklar
- İktisadi Karlılık = (VÖK + Faiz Giderleri) / Kaynak Toplamı =Pasif Toplamı

Aykırı firmaların açıklanmasında ortalama  $\pm$  standart sapma değer aralığı incelenebileceği gibi ortalama  $\pm 2$ \*standart sapma değerine de bakılabilir. Çalışmamızda her iki alternatif değerlendirilmiş olup, mümkün olduğunca homojen kümeler elde etme kaygısıyla daha dar bir aralığı işaret eden ortalama  $\pm$  standart sapma değer aralığı dikkate alınmıştır. Bu şekilde yapılan elemelerde 2015 yılı için 190 şirketten 121 adedi, 2016 yılı için 173 şirketten 127 adedi aykırı olarak belirlenmiştir. Aykırı şirketler bu çalışma kapsamında değerlendirilmemiş olup, hile araştırmaları veya sektörel araştırmalara ayrıca konu edilebilir.

Küme analizi algoritmalarından EM algoritması veri setimize uygun olduğu için tercih edilmiştir. EM Algoritması aracılığı ile benzerlik adalarının belirlenmesinin yanı sıra, her bir küme için birer ortalama ve standart sapma değeri elde edilmiştir. Elde edilen ortalama ve standart sapma değerleri kullanılarak küme sınırlarının belirlenmesi mümkün olmuştur. Algoritma yukarıda belirtilen finansal oranlara uygulandığında 2015 yılı için hem aykırı şirketler ayıklandığı ve ayıklanmadığı her iki durumda da iki kümenin oluştuğu; 2016 yılı için aykırı şirketler ayıklanmadığı durumda dört, ayıklandığı durumda ise iki kümenin oluştuğu görülmüştür. 2016 yılında gerek aykırı şirket sayısının fazla olması gerekse daha fazla benzerlik adacıklarının oluşması 2016 yılının ekonomik gelişmelerine bağlanabilir.

#### 4.SONUÇ, TARTIŞMA

Veri miktarının fazla olması çalışmamızda Veri Madenciliği Tekniklerini kullanmamızı gerekli kılmış olup, 2015 ve 2016 yılları için sırasıyla 121 ve 127 adet Finansal Tablo değerlendirilmiş olup, her iki yıl arasında yerel ve global ekonomik gelişmelerle bağlantılı olarak farklılaşmalar olduğu görülmüştür. 2016 yılında 2015 yılından farklı olarak daha fazla benzerlik adaları tespit edilmiş olup, ağırlıklı olarak şirketlerin belirli bir kümede toplandığı görülmüştür. 2017 yılı yılsonu finansal tabloları henüz hazır olmadığı ve ara dönem finansal tabloları yetersiz kalacağı için çalışmamıza dahil edilememiştir.

Herhangi bir aykırılık analizi yapılmadan algoritmanın belirlenen finansal oranlara uygulanması durumunda 2015 yılı için şirketlerin %87'si tek bir kümede, geri kalan %13'ü diğer bir kümede yer almıştır. 2016 yılı için bir kümede şirketlerin %65'i yer alırken, diğer kümelerdeki şirketlerin dağılımı %17, %13 ve %5 şeklindedir. Bu veri TÜİK'den alınan ihracat endeksi ve takvim etkisinden arındırılmış sanayi ciro endeksindeki düşüş ile anlamlandırılmıştır. İhracat endeksi 2010 yılından itibaren 2014 yılına kadar Aralık ayları itibarıyla 103.9, 108.3, 109.8, 108.2, 103.1 iken 2015 ve 2016 yılları için 91.9, 91.5 olmuş ve Takvim etkisinden arındırılmış sanayi ciro endeksindeki artış 2010 yılından 2014 yılına kadar sırasıyla %18.6, %28.7, %7.8, %9.9, %12.5 iken 2015 yılında %8.2 ve 2016 yıllarında %6.9 olmuştur. İmalat yapan analizimize konu olan şirketlerin satışları ihracat yoluyla yurt dışına ve yerli tüketiciye olacağı için bu iki endekste düşüşün 2016 yılında şirketlerin finansal durumunda farklılaşmaya yol açtığı düşünülmektedir.

Aykırılık analizi sonrası şirketlerin finansal tabloları değerlendirildiğinde 2015 yılı için şirketlerin iki küme oluşturduğu ve %51'inin bir kümede, %49'unun diğer kümede yer aldığı; 2016 yılı içinde şirketlerin iki küme oluşturdukları ve %54'ünün bir kümede, %46'sının diğer kümede yer aldığı görülmüştür. Şirketlerin kümelerine homojen olarak dağılmasının nedeni Borsa İstanbul'a kayıtlı işletmelerin belirli regülasyonlara tabi olmaları, profesyonel yöneticiler tarafından yönetilmeleri ve sık sık denetimden geçmelerinin olduğu düşünülmektedir. İşletmelerin finansal göstergelerinin farklılaşmasının altında küresel ve yerel ekonomik gelişmelerin yanı sıra sektör ve alt sektör bazında oluşan farkların yattığı düşünülmektedir.

Elde edilen kümeler yoluyla finansal oranların yeni eşik değerleri belirlenmiş olup; bu değerlendirme sonucunda da daha likit olan şirketler daha az borçlandıkları, daha likit olan şirketlerin daha karlı olduğu, faaliyet oranları açısından yıllar arasında çok büyük farklılaşma olmadığı genel anlamda tespit edilmiştir. 2016 yılında finansal oranlar bakımından uç değerlere giden kümeler olduğu diğer bir tespittir.

Veri madenciliğinin temel kullanım amacı, önceden bilinmeyen yoğun verilerden bilginin çıkartılması olup, çalışmamızda da işletme sayısının ve kullanılan oran sayısının yani veri matrisinde yer alan satır/sütun sayısının fazla olması nedeniyle Veri Madenciliği analiz tekniklerinden birisi olan kümeleme analizi tercih edilmiştir. Kullanılan yöntem sayesinde çok hızlı bir şekilde çok fazla finansal oran incelenmiş olup; Veri Madenciliği Teknikleri geleneksel finansal tablo analiz yöntemleri arasında yer alan yatay analiz, dikey analiz ve oran analizi yöntemlerine alternatif bir teknik olarak ortaya konulmuştur. Çalışmamızın finans alanını ile bilişim ve istatistik alanlarından türemiş Veri Madenciliği'ni bir araya getirmesi yoluyla yazına katkıda bulunduğunu düşünmekteyiz.

**KAYNAKLAR**

- Barnes, Paul. Stock Market Efficiency. Insider Dealing and Market Abuse. Fanham. GB Gower, 2009, p.7
- Gürsoy, U. Tuğba Şimşek. Veri Madenciliği ve Bilgi Keşfi. 1. Baskı. Ankara. Pegem Akademi. 2009. s.1
- Halka Arz ve Borsada İşlem Görme. Borsa İstanbul. 2016.  
[http://www.borsaistanbul.com/data/kilavuzlar/Halka\\_arz\\_ve\\_borsada\\_islem\\_gorme.pdf](http://www.borsaistanbul.com/data/kilavuzlar/Halka_arz_ve_borsada_islem_gorme.pdf) (Erişim Tarihi 7 Ekim 2016)
- Krzysztof, Cios & Others. Data Mining: A Knowledge Discovery Approach. USA: Springer Science+Business Media. LLC. 2007, p.6
- Larose, Daniel T. Discovering in Data: An Introduction to Data Mining. USA. Wiley&Sons, Inc. 2005. p.1-2-3
- Myatt, Glenn J., and Johnson, Wayne P. Making Sense of Data I : A Practical Guide to Exploratory Data Analysis and Data Mining, USA. Wiley. 2014. p.19
- Peter,Chamoni. [www.enzyklopaedie-der-wirtschaftsinformatik.de/...-/Data](http://www.enzyklopaedie-der-wirtschaftsinformatik.de/...-/Data) Eylül 2013 (Erişim Tarihi 7 Ekim 2016)
- Refaat, Mamdout. Data Preperation forData Mining Using SAS. San Francisco: Morgan Kaufmann Publishers. 2007. p.7-8
- Silahtaroglu, Gökhan. Veri Madenciliği. 1. Basım. İstanbul.: Papatya Yayıncılık Eğitim. 2008. s.10
- Tuffery, Stephane. Data Mining and Statistics for Decision Making. United Kingdom. John Wiley & Sons Ltd. 2008. p.12
- Witten, Ian H. Frank, Eibe. Hall, Mark A. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann Series. USA. 2011. p.5