



## VERİ MADENCİLİĞİ TEKNİKLERİ İLE HİSSE SENETLERİ ARASINDAKİ FİYAT ETKİLEŞİMLERİNİN BELİRLENMESİ

Özkan ÜNSAL\*

Süleyman Demirel Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Isparta, Türkiye

### Anahtar Kelimeler

*BİST,  
Veri Madenciliği,  
Kümeleme,  
Birliktelik Kuralları.*

### Öz

Yatırımcıların halka açık şirketlere ait hisse senetleri üzerinde alım/satım işlemlerini gerçekleştirdiği kurumsal piyasalara genel olarak borsa ismi verilmektedir. Ülkemizde Borsa İstanbul(BİST) adıyla faaliyet gösteren borsa piyasası günlük yaklaşık 2,5 milyar lotluk işlem hacmi ile yatırımcıların ilgisini çekmektedir. Yatırımcıların hisse senedi seçiminde teknik analiz yöntemleri giderek artan bir önem kazanmaktadır. Gelişen yazılım teknolojileri ve veri madenciliği teknikleri sayesinde hisse senetlerinin geçmiş verileri üzerinde başarılı teknik analizler uygulanabilmektedir. Bu çalışmada, veri madenciliği teknikleri kullanılarak günlük fiyat değişimine göre birbirleri ile en çok hareket eden hisse senetlerine ait ilişki kurallarının ortaya çıkarılması amaçlanmıştır. Çalışma için BİST'te işlem gören 408 hisse senedine ait 2019 yılı boyunca 249 işlem günü için kapanış fiyatı, gün sonu fiyat değişimi ve günlük hacim verileri elde edilmiştir. Veriler üzerinde her bir işlem günü için K-Means yöntemi ile kümeleme yapılmış ardından elde edilen kümelere Birliktelik Kuralı yöntemlerinden Apriori algoritması uygulanarak yıl boyunca birbirleri ile en çok hareket eden hisse senetlerine ait ilişki kuralları elde edilmiştir.

## DETERMINATION OF PRICE INTERACTIONS BETWEEN STOCKS WITH DATA MINING TECHNIQUES

### Keywords

*BIST,  
Data Mining,  
Clustering,  
Association Rules.*

### Abstract

Institutional markets where investors trade on stocks of publicly traded companies are generally called stock markets. The stock market, operating under the name of Borsa Istanbul (BIST) in our country, attracts the attention of investors with a daily trading volume of approximately 2.5 billion lots. Technical analysis methods are becoming increasingly important in investors' selection of stocks. Thanks to the developing software technologies and data mining techniques, successful technical analysis can be applied on the historical data of stocks. In this study, by using data mining techniques, it is aimed to reveal the relationship rules of stocks that move with each other the most according to daily price changes. For the study, closing prices, end-of-day price changes and daily volume data were obtained for 249 trading days of 408 stocks traded on BIST during the year 2019. On the data, clustering was made with K-Means method for each trading day, and then, Apriori algorithm, one of the Association Rule methods, was applied to the obtained clusters, and the relationship rules for the stocks that moved with each other the most throughout the year were obtained.

### Alıntı / Cite

Ünsal, Ö., (2020). Veri Madenciliği Teknikleri ile Hisse Senetleri Arasındaki Fiyat Etkileşimlerinin Belirlenmesi, Mühendislik Bilimleri ve Tasarım Dergisi, 9(1), 106-112.

### Yazar Kimliği / Author ID (ORCID Number)

Ö. Ünsal, 0000-0002-6741-7826

### Makale Süreci / Article Process

Başvuru Tarihi / Submission Date	01.12.2020
Revizyon Tarihi / Revision Date	09.12.2020
Kabul Tarihi / Accepted Date	12.12.2020
Yayın Tarihi / Published Date	29.12.2020

\*İlgili yazar / Corresponding author: [ozkanunsal@gmail.com](mailto:ozkanunsal@gmail.com), +90-554-797-4296

## 1. Giriş (Introduction)

Yatırımcıların halka açık şirketlere ait hisse senetleri üzerinde alım satım işlemlerini gerçekleştirdiği kurumsal piyasalara genel olarak borsa ismi verilmektedir. Borsa aynı zamanda bir ülkenin önemli ve giderek büyümekte olan şirketlerine, bireysel veya kurumsal yatırımcılar tarafından yatırım yapılabilmesini sağlayan bir platformdur. Ülkemizde borsa faaliyetleri ilk defa 1985 yılında başlamış, 2013 yılından itibaren Borsa İstanbul (BİST) adı ile devam etmektedir. BİST'te işlem gören 408 hisse senedi günlük yaklaşık 2,5 milyar lotluk işlem hacmi ile yatırımcıların önemli düzeyde ilgisini çekmektedir. Yatırımcılar hisse senedi tercihinde bulunurken şirketin faaliyet alanı, dönemlik bilanço verileri, kâr payı dağıtım politikası ve ileriye dönük beklentiler gibi çeşitli kriterlere bakmaktadırlar. Bunların yanı sıra geçmiş hacim ve fiyat değişimi gibi verilere bakarak ileriye dönük fiyat tahmini yapmaya çalışan teknik analiz sonuçlarını da takip etmektedirler. Bu büyük veri içerisinde en doğru seçimi yapmak yatırımcılar açısından oldukça zor bir karar olmaktadır.

Veri madenciliği süreci; veri ambarlarında yer alan çok çeşitli ve çok miktarda veriye dayanarak keşfedilmemiş bilgileri ortaya çıkarmak ve bunları karar destek mekanizmaları ile eylem planlarının gerçekleştirilmesi için kullanılması şeklinde tanımlanabilir (Dener vd., 2009). Veritabanı temelli bilgisayar sistemlerinin yaygınlaşması ile birlikte her türlü veri elektronik ortama aktarılmakta ve bunun sonucunda büyük veri yığınları ortaya çıkmaktadır. Veri tabanlarında saklanan veri, bir dağa benzetilirse, bu veri dağı tek başına değersizdir ve kullanıcı için çok fazla bir anlam ifade etmez. Ancak bu veri dağı, belirli bir amaç doğrultusunda sistematik olarak işlenir ve analiz edilirse, değersiz görülen veri yığnında, amaca yönelik sorulara cevap verebilecek çok değerli bilgilere ulaşılabilir (Özekes, 2003). Veri madenciliği tekniklerinin tıp, eğitim, finans ve meteoroloji gibi birçok alanda bulunan veriler üzerinde uygulanmasıyla gelecekte olabilecek durumların tahminleri ve geçmiş verilerin sınıflandırmaları yapılabilmektedir. Bu sayede insanların ya da makinelerin istenilen alana yönelik sonuç, yorum çıkartmaları ve ileriye yönelik karar vermelerine önemli bir destek sağlanmaktadır.

Yapılan çalışmada, veri madenciliği teknikleri kullanılarak hisse senetlerinin fiyat hareketlerinde birbirlerine olan etkilerinin ortaya çıkarılması amaçlanmıştır. Çalışma için BİST'te işlem gören 408 hisse senedi ve BIST endeksine ait 2019 yılı boyunca kapanış fiyatı, gün sonu fiyat değişimi ve günlük hacim verileri elde edilmiştir. Veriler üzerinde öncelikle K-Means yöntemi ile kümeleme işlemi yapılmış ardından elde edilen kümelere Birliktelik Kuralı yöntemlerinden Apriori algoritması uygulanarak yıl boyunca birbirleri ile en çok hareket eden hisse senetlerine ait ilişki kuralları elde edilmiştir. Elde edilen kurallar incelendiğinde çalışmanın hisse senetleri arasındaki fiyat etkileşimlerini ortaya çıkarmada kullanılabilecek bir yöntem olduğu değerlendirilmektedir.

Çalışmanın bundan sonraki bölümünde hisse senedi fiyat oluşumlarına yönelik gerçekleştirilen çalışmaların kaynak araştırması yapılmış ve bu çalışmalarda kullanılan yöntemlerden bahsedilmiştir. Üçüncü bölümde hisse senedi piyasalarının genel yapısından bahsedilerek, veri madenciliğinde yer alan K-Means kümeleme tekniği ile birliktelik kuralı yöntemlerinden Apriori algoritması anlatılmıştır. Dördüncü bölümde ise veri madenciliği kullanılarak hisse senetleri arası fiyat etkileşimlerinin bulunma aşamaları ve bunun sonucunda birbirleri ile en çok hareket eden hisse senetlerine ait elde edilen ilişki kuralları verilmiştir. Sonuç bölümünde yapılan çalışma genel olarak değerlendirilerek elde edilen sonuçlar yorumlanmıştır.

## 2. Kaynak Araştırması (Literature Survey)

Veri madenciliği yöntemleri ile hisselerin geçmiş verileri kullanılarak fiyat oluşumlarına etki eden faktörler analiz edilebilmektedir. Literatür incelendiğinde hisse senedi piyasalarında fiyat ilişkilerinin ortaya çıkarılmasına yönelik veri madenciliği yöntemlerinin kullanıldığı birçok çalışma olduğu görülmüştür.

Ergür (2014), BİST üzerinde yer alan Bankacılık, Telekomünikasyon ve Bilgi Teknolojileri sektörlerine ait hisselerle yönelik fiyat değişim tahmininde kullanmak üzere ilişkide buldukları hisse senetlerini bulabilmek için yaptığı çalışmada, geçmiş fiyat vektörleri arasındaki zaman kaymalı Pearson benzerliğinden yararlanmıştır.

Gazel ve Akel (2018), farklı sektör endeksleri içinde yer alan hisse senetlerinin, geçmiş fiyat verilerine göre kümeleme analizine tabi tutulduğunda aynı sektör içerisinde yer alıp almadığını tespit etmek amacıyla yaptıkları çalışmada veriler üzerinde Hiyerarşik Yığınsal Kümeleme analizi uygulamışlardır.

Tekin (2018), çalışmasında, BİST100 endeksinde yer alan şirketlere ait 2015 yılı fiyat ve mali tablo verileri üzerinde Ward, K-Means ve İki Adımlı Kümeleme Analizi yöntemlerini uygulayarak yatırım yapılabilecek en uygun portföyü oluşturmaya çalışmıştır.

Bozma ve Kul (2020), bir sosyal medya platformu olan Twitter üzerindeki paylaşımların hisse senetleri fiyatlarına ve hisse senetlerinin birbirlerine olan etkisini araştırmıştır. Çalışmada, BİST'te yer alan 3 telefon şirketine ait hisse senetleri hakkında yapılan paylaşımların Naive Bayes yöntemi ile duygu analizi yapılmış ve analiz sonuçları incelendiğinde iki telefon şirketine ait duygu puanlarının üçüncü şirketin koşullu varyansını anlamlı olarak arttırdığı tespit edilmiştir.

### 3. Materyal ve Yöntem (Material and Method)

Çalışmanın bu bölümünde hisse senedi piyasalarının genel yapısı ile veri madenciliğinde yer alan kümeleme tekniği ve birliktelik kuralı analizinden bahsedilecektir.

#### 3.1. Hisse Senedi Piyasası (Stock Market)

Borsa, her ölçekten tüm yatırımcılara dinamik yatırımlarla anaparalarını arttırma şansı tanıdığı ve kârlılık paylaşımı adına temel bir ortam oluşturduğundan ekonomide önemli bir rol oynamaktadır (Soni, 2011). Borsa; tıpkı döviz, emtia, faiz gibi bir yatırım aracıdır. Ülkemizde borsa faaliyetleri ilk defa 1985 yılında başlamış, 2013 yılından itibaren Borsa İstanbul(BİST) adı ile devam etmektedir. BİST üzerinde 2019 yılı boyunca 408 şirket hisse senedine ait yaklaşık olarak günlük 2,5 milyar lotluk işlem hacmi gerçekleşmiştir. BİST'te işlem gören şirketler faaliyet alanlarına göre sanayi, bankacılık, madencilik, teknoloji gibi farklı gruplara ayrılmıştır. Bir başka gruplama ise şirketlerin halka arz edilen hisselerinin piyasa değerine göre; Yıldız, Ana, Gelişen İşletmeler ve Yakın İzleme isimli pazar gruplarına dâhil edilmesidir.

BİST üzerindeki hisse senedi alım/satım işlemleri alışı karşı satış şeklinde takas mantığına dayanmaktadır. Hisse senedi piyasasında işlemler elektronik alım/satım sistemi aracılığıyla fiyat ve zaman önceliği kuralına göre "Sürekli İşlem", "Piyasa Yapıcı Sürekli İşlem" ve "Tek Fiyat" yöntemlerinde otomatik olarak gerçekleşmektedir (Borsa İstanbul, 2020). Bir hisse senedini almak ya da satmak isteyen yatırımcılar lot cinsinden işlem miktarı ve fiyat bilgisi ile alım/satım emri girmektedir. Aynı fiyattan girilen emirler ise giriş zamanına göre sıraya alınır. Alım/satım emirlerinin gün boyu eşleştirilmesi sonucunda her hisselerin kapanış fiyatı ve günlük işlem hacmi ortaya çıkmaktadır. Herhangi bir ceza ya da kısıta tabi olmayan hisselerin bir işlem gününde fiyatları en fazla  $\pm\%20$ 'lik bir marjda hareket etme limitine sahiptir(çalışmanın yapıldığı tarih aralığında). Bu limit kuralı hisselerin fiyat istikrarı ve olağanüstü durumlarda yatırımcıların korunması için uygulanmaktadır.

#### 3.2. Veri Madenciliği (Data Mining)

Yatırımcılar hisse senetleri üzerinde alım/satım işlemi yaparak kâr elde etmeyi amaçlamaktadır. Yatırımcıların hisse senedi seçiminde birçok kriter bulunmaktadır. Bu kriterler genel olarak temel ve teknik analiz altında iki farklı yönetime göre belirlenmektedir. Temel analiz şirketlerin finansal tabloları, faaliyet alanları, geleceğe yönelik potansiyel iş beklentileri gibi verilere dayanır. Teknik analiz ise hisse senedinin gelecek fiyatının kestirimi için geçmiş fiyat ve hacim verileri ile ilgilidir. Yazılım teknolojilerindeki gelişmeler ve veri madenciliği teknikleri sayesinde hisse senedi fiyatlarının oluşma nedenlerinin tespiti ve geleceğe yönelik tahmininde teknik analiz yöntemleri giderek daha çok kullanılmaktadır.

Veri madenciliği büyük veri yığınlarından anlamlı ve gelecekteki süreçlerde karar vermeye yardımcı olacak nitelikte bilgilerin elde edilme sürecidir (Ünsal, 2011). Veri madenciliği insan ve bilgisayarların ortak çabası ile ortaya çıkmıştır. İnsanlar problemleri tarif ederler, veritabanlarını tasarlarlar ve amaçları belirler. Bilgisayarlar ise veriler arasında bu amaçlara uygun verileri eleyerek desenleri ortaya çıkarırlar (Weiss ve Indurkha, 1998). Veri madenciliği tekniklerinin tıp, eğitim, finans ve meteoroloji gibi birçok alanda geçmişte bulunan veriler üzerinde uygulanması ile mevcut verilerin analizi ve ileriye yönelik tahminler yapılabilir.

Bu çalışmada hisse senetlerinin günlük kapanış fiyatına göre yüzdelerle değişim ve hacim bilgisi kullanılarak birlikte hareket eden hisse senetleri ortaya çıkarılmaya çalışılmıştır. Bunun için ise veriler üzerinde veri madenciliği tekniklerinden K-Means kümeleme yöntemi ve Apriori algoritması ile Birliktelik Kuralı analizi uygulanmıştır.

##### 3.2.1. K-Means kümeleme tekniği (K-Means clustering method)

Kümeleme, n adet gruplanmamış bireyin p adet özelliğinden yararlanarak, bunları benzerliklerine göre sınıflara ayırma işlemlerini kapsar. Kümeleme ile veriler gruplara ayrılırken seçilen kriterlere göre hesaplanan mesafe matrislerinden yararlanır. Her bir veri kendisine en yakın mesafedeki veriler ile bir kümeyi oluşturur.

K-Means kümeleme tekniğinde “K” küme sayısını, “Means” ise kümeyi oluşturan elemanların ağırlıklı ortalamasını ifade eder. Kümeleme işlemine başlamadan önce k küme sayısı araştırmacı tarafından verilir. K-Means algoritması kümelerin sürekli olarak güncellendiği bir yöntemdir. Başlangıçta belirlenen k adet küme sayısı için rastgele küme merkezleri oluşturulur. Elemanların belirlenen özellikleri kullanılarak her bir küme merkezine olan uzaklıklarının hesaplanmasıyla mesafe matrisi oluşturulur. Mesafe matrisine göre her bir eleman kendisine en yakın küme merkezine atanır. Başlangıç atamasından sonra her bir küme merkezi üye özelliklerinin ağırlıklı ortalamasının hesaplanması ile güncellenir ve elemanlar yeni merkezlere göre yine kendilerine en yakın olanına atanır. Bu işlem hiçbir elemanın küme merkezi değişmeyinceye kadar döngüsel olarak devam eder ve nihai kümeler oluşturulur. Bu çalışmada, hisse senedi verilerinin günlük kümelenmesinde bölümleyici kümeleme teknikleri altında yer alan ve bilimsel, endüstriyel uygulamalarda çok tercih edilen K-Means algoritması kullanılmıştır.

### 3.2.2. Apriori birliktelik kuralı algoritması (Apriori association rule algorithm)

Birliktelik kuralları analizi veri madenciliği uygulamalarında, olayların meydana gelmesinde beraber hareket eden unsurların ortaya çıkarılmasında kullanılır. Pazarlama alanında birlikte satın alınan ürünlerin belirlenerek satış stratejisi oluşturulmasında çok tercih edilen bir yöntemdir.

Birliktelik kuralları üretilirken, belirlenen destek(support) ve güven(confidence) ölçütleri dikkate alınır. Bir kuralın destek ölçütü, o kuralın geçtiği olay sayısının veri setindeki tüm olayların sayısına oranıdır. Güven ölçütü ise kuralda birlikte hareket ettiği belirtilen öğelerin birbirlerine oranıdır. Kurallar oluşturulurken araştırmacı tarafından verilen minimum destek ve güven ölçütü üzerinde yer alanlar dikkate alınmaktadır. A ve B gibi iki elemanın beraber hareket etme kuralı  $A \Rightarrow B$  şeklinde ifade edilir. Oluşan kuralların kalitesini ve anlamlılığını gösteren en önemli parametre ise güven ölçütüdür. Güven ve destek ölçütünün dışında, conviction, leverage ve lift parametreleri de bulunmaktadır. Birliktelik kuralları analizine yönelik birçok algoritma geliştirilmiştir.

Birliktelik kuralı madenciliği, tüm sık geçen öğelerin bulunması ve sık geçen bu öğelerden güçlü birliktelik kurallarının üretilmesi olmak üzere iki aşamalıdır. Birliktelik kuralının ilk aşaması için kullanılan Apriori Algoritması, sık geçen öğeler madenciliğinde kullanılan en popüler ve klasik algoritmadır. Bu algoritmada özellikler ve veri, Boolean ilişki kuralları ile değerlendirilir (Özçakır ve Çamurcu, 2007). Bu algoritma temelinde tekrarlayan bir niteliğe sahiptir (Han ve Kamber, 2006). Algoritma, veri setinde yer alan tüm elemanların destek ölçütlerini hesaplar ve belirlenen minimum destek ölçütü altında kalan elemanlar elenmiş olur. Sonraki adımlarda destek ölçütüne uyan elemanlar ile ikili, üçlü, vb. şekilde kümeler oluşturularak her bir kümenin destek ölçütü hesaplanır ve yine minimum destek ölçütüne uyanlar ile yola devam edilir. Son olarak oluşturulan kuralların güven ölçütleri hesaplanarak belirlenen minimum güven ölçütü değerinin üzerinde olan kurallar ortaya çıkarılmış olur. Bu çalışmada kuralların çıkarılmasında en çok tercih edilen algoritma olan Apriori algoritması kullanılmıştır.

## 4. Veri Madenciliği ile Fiyat Etkileşimlerinin Belirlenmesi (Determining Price Interactions with Data Mining)

Çalışmanın bu bölümünde, günlük fiyat değişimine göre birbirleri ile en çok hareket eden hisse senetlerine ait ilişki kurallarının ortaya çıkarılması için geliştirilen yöntemin uygulama aşamaları anlatılacaktır.

### 4.1. Veri Seti (Dataset)

Çalışma için 2019 yılına ait BIST kapanış verileri, tr.investing.com internet sitesi üzerinden elde edilmiştir. İşlem gören 408 hisse senedine ait 2 Ocak 2019 ile 31 Aralık 2019 tarihleri arasındaki 249 işlem günü için 100.755 adet kapanış verisi Tablo 1’de gösterilen formatta Mysql veritabanı üzerine alınmıştır. Hisse senetlerine ek olarak BIST endeks verisi de günlük olarak elde edilmiştir. Tabloda gösterilen hacim alanı hisselerin gün boyu gerçekleşen lot(adet) cinsinden işlem miktarını, günlük değişim ise hisselerin gün sonundaki kapanış fiyatının bir önceki gün kapanış fiyatına oranıdır.

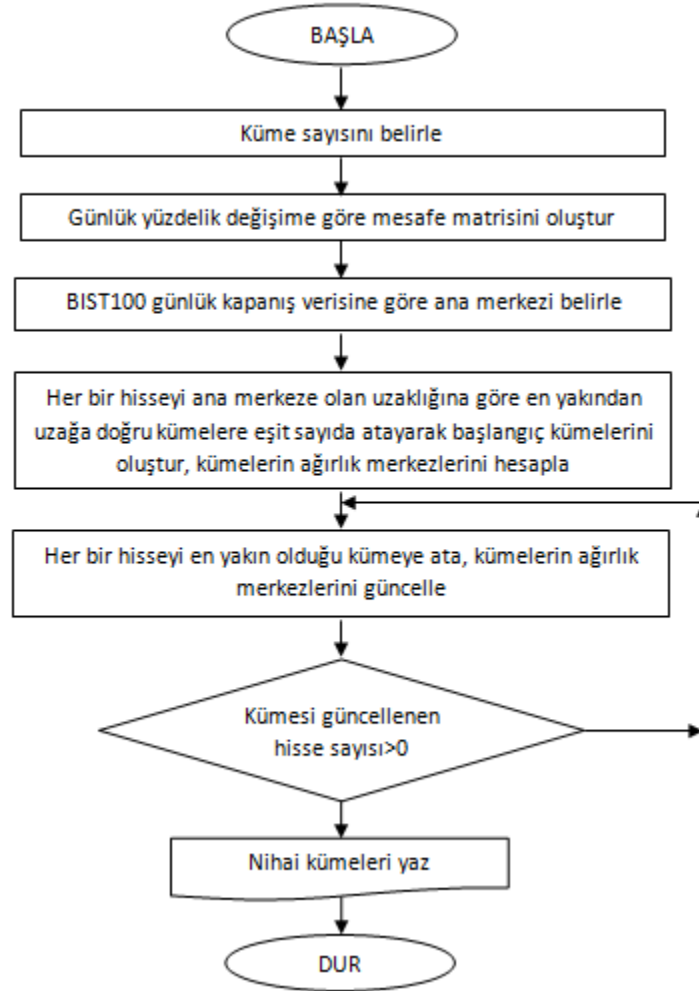
**Tablo 1.** Bir hisseye ait günlük kapanış veri örneği (Example of daily closing data for a share)

Tarih	Hisse	Kapanış Fiyatı	Hacim	Günlük Değişim(%)
02.01.2019	ACSEL	3,21	177.750	-1,53

Çalışmanın yapıldığı 2019 yılı boyunca hisselerin önceki kapanış fiyatına göre en fazla aşağı ya da yukarı yönlü %20 fiyat değişim marjı bulunmaktadır. Elde edilen verilerde yıl boyu kesintisiz işlem gören hisse senedi sayısı 380 adettir. İşlem yasağı, hissenin yılın sonraki dönemlerinde borsada işlem görmeye başlaması ya da borsadan çıkarılması gibi sebepler nedeniyle 28 hisseye ait tüm işlem günleri için kapanış verileri alınamamıştır. Çalışmada hacim bazında düşük işlem gören sığ piyasaya sahip hisse senetleri analiz dışında tutulmuştur, bunun için yıllık ortalama işlem hacmine göre en yüksek 250 hisse senedi dikkate alınmıştır.

#### 4.2. Hisse Senetlerinin K-Means İle Günlük Olarak Kümeleneşmesi (Clustering Stocks Daily with K-Means)

Elde edilen hisse senedi kapanış verilerine göre her bir işlem günü için C# dili ile geliştirilen yazılım aracılığıyla K-Means yöntemi kullanılarak kümeleme işlemi yapılmıştır. Kümeleme işleminin başlangıcında K-Means algoritmasından farklı olarak başlangıç küme merkezleri rastgele belirlenmemiştir. O gün için BIST100 endeks verisi ana küme merkezi olarak seçilmiş, hisse senetlerinin BIST100 verisine olan uzaklık matrisine göre belirlenen küme sayısına eşit olacak şekilde en yakından uzağa atanması yapılmıştır. İlk atamalar bittikten sonra K-Means algoritmasındaki gibi her bir kümenin ağırlık merkezi hesaplanarak hisselerin en yakın kümeye atanması işlemine geçilmiştir. Hisselerin kümelerinde değişim olmayıncaya kadar bu işlem devam ettirilmiş ve o gün için kümeler ortaya çıkarılmıştır. Bu kümeleme işlemi 249 işlem günü için tekrarlanmıştır. Anlatılan kümeleme işlemine ait akış diyagramı Şekil 1’de, bir işlem günü sonu oluşan örnek kümeleme sonucu ise Şekil 2’deki gibi gösterilmiştir.



Şekil 1. Kümeleme işlemi akış diyagramı (Clustering process flowchart)

```
[ 'IEYHO', 'NIBAS', 'OZGYO', 'PEKGY' ],

[ 'ADESE', 'AEFES', 'AFYON', 'AGYO', 'AKBNK', 'AKENR', 'AKGUV', 'AKSEN', 'ALARK', 'ALCTL', 'ALYAG', 'ANACM',
  'ANELE', 'ARCLK', 'ARSAN', 'ASELS', 'ASUZU', 'AVGYO', 'AYEN', 'BANVT', 'BEYAZ', 'BIMAS', 'BJKAS', 'BMELK',
  'BNTAS', 'BOLUC', 'BOSSA', 'BRKSN', 'BRMEN', 'BSOKE', 'CEMTS', 'CIMSAS', 'CRDFA', 'CRFSA', 'DENGE', 'DESPC',
  'DEVA', 'DGATE', 'DIRIT', 'DMSAS', 'DOAS', 'DOHOL', 'DZGYO', 'ECILC', 'EDIP', 'EKGYO', 'EMKEL', 'ENJSA',
  'ERGL', 'ERSU', 'ETYAT', 'EUHOL', 'EUKYO', 'EUYO', 'FLAP', 'FONET', 'GARAN', 'GARFA', 'GEREL', 'GLRYH', 'GLYHO',
  'GOODY', 'GOZDE', 'GRNYO', 'GSDHO', 'GUBRF', 'HALKB', 'HDFGS', 'HEKTS', 'HLGYO', 'HURGZ', 'ICBCT', 'IHEVA',
  'IHYAY', 'INDES', 'INFO', 'IPEKE', 'ISCTR', 'ISDMR', 'ISGSY', 'ISGYO', 'IZMDC', 'IZTAR', 'KAREL', 'KARSN',
  'KATMR', 'KERTV', 'KFEIN', 'KLMSN', 'KORDS', 'KOZAA', 'KOZAL', 'KRDMB', 'KRDMD', 'KUYAS', 'MAKTK', 'MARKA',
  'MARTI', 'METUR', 'MGROS', 'MRGYO', 'NETAS', 'NTHOL', 'OSTIM', 'OYAYO', 'OZKGY', 'PARSN', 'PEGYO', 'PETKM',
  'PGSUS', 'PNSUT', 'PRKME', 'RHEAG', 'RTALB', 'RYSAS', 'SAHOL', 'SANFM', 'SASA', 'SEKUR', 'SILVR', 'SISE',
  'SKBNK', 'SNGYO', 'TCELL', 'TEKTU', 'THYAO', 'TKFEN', 'TMSN', 'TOASO', 'TRCAS', 'TRKCM', 'TSKB', 'TSPOR',
  'TTKOM', 'TUPRS', 'ULKER', 'ULUUN', 'USAK', 'VAKBN', 'VAKFN', 'VAKKO', 'VERUS', 'VESTL', 'VKFYO', 'VKGYO',
  'VKING', 'YATAS', 'YKGYO', 'YUNSA', 'ZOREN' ],

[ 'ADNAC', 'AGHOL', 'AKFGY', 'AKSGY', 'ALBRK', 'ATLAS', 'AVTUR', 'BAGFS', 'BERA', 'BTCIM', 'BURCE', 'CUSAN',
  'DAGI', 'DESA', 'DGGYO', 'DOBUR', 'DOGUB', 'DYOBY', 'ENKAI', 'ESCOM', 'FENER', 'GENTS', 'GSDDE', 'GSRAY',
  'HUBVC', 'IDGYO', 'IHGZT', 'IHLAS', 'IHLGM', 'ISFIN', 'ISMEN', 'ISYAT', 'ITTFH', 'KCHOL', 'KLGYO', 'KLNMA',
  'KRDMA', 'KRGYO', 'KRSTL', 'MEPET', 'MERKO', 'METRO', 'MIPAZ', 'MNDRS', 'MRDIN', 'MSGYO', 'NUGYO', 'ODAS',
  'ORGE', 'OYLUM', 'OZBAL', 'PENGD', 'PINSU', 'POLHO', 'RAYSG', 'RYGYO', 'SAMAT', 'SANKO', 'SAYAS', 'SELEC',
  'SKTAS', 'SNKRN', 'SODA', 'SOKM', 'TATGD', 'TAVHL', 'TDGYO', 'TIRE', 'TKNSA', 'TRGYO', 'TSGYO', 'TUCLK', 'TUKAS',
  'UNYEC', 'UTPYA', 'YAYLA', 'YGYO', 'YKBNK', 'YYAPI' ],

[ 'AVOD', 'CEMAS', 'DAGHL', 'DGKLB', 'FORMT', 'GLBMD', 'GUSGR', 'MTRYO', 'PKART', 'PRKAB', 'ULAS', 'VANGD' ]
```

Şekil 2. Bir işlem gününe ait örnek kümeleme sonuçları (Sample clustering results for one trading day)

### 4.3. Kümeler Üzerinde Apriori Algoritmasının Uygulanması (Application of Apriori Algorithm on Sets)

249 işlem günü için küme sayısı 4 olacak şekilde kümeleme işlemi yapıldıktan sonra toplam 982 adet küme Şekil 2'deki formatta oluşturulmuştur. Veri seti, hisse senetleri sütun başlığı ve hissenin sıradaki kümede yer aldığı 1 yer almadığında ise 0 değeri hisse verisi olacak şekilde yatay bir tabloya dönüştürülmüştür. Veri setine, minimum destek seviyesi %15 ve minimum güven seviyesi %60 parametre değerleri kullanılarak Python programlama dili ve Weka veri madenciliği yazılımı aracılığı ile Apriori algoritması uygulanmıştır. Uygulama sonucunda güvenilirlik oranı en yüksek 10 birliktelik kuralı sonucu Tablo 2'deki gibi gösterilmiştir.

Tablo 2. Güven oranı en yüksek 10 birliktelik kuralı (10 association rules with the highest trust rate)

No	Kural	Conf.	Lift.	Lev.	Conv.
1	EUYO, DYOBY (188) => NETAS (172)	0.91	3.61	0.13 (124)	8.25
2	EUYO, NETAS (208) => DYOBY (172)	0.83	3.26	0.12 (119)	4.2
3	GRNYO (249) => HALKB (182)	0.73	2.88	0.12 (118)	2.73
4	AGYO (249) => GOZDE (179)	0.72	2.84	0.12 (115)	2.62
5	INDES (249) => TMSN (176)	0.71	2.79	0.11 (112)	2.51
6	EUYO (249) => NETAS, DYOBY (172)	0.69	3.44	0.12 (122)	2.55
7	INDES (249) => HALKB (170)	0.68	2.69	0.11 (106)	2.32
8	KRDMD (249) => MARTI (164)	0.66	2.6	0.1 (100)	2.16
9	INDES (249) => IHGZT (164)	0.66	2.6	0.1 (100)	2.16
10	VERUS(249)=>GRNYO(160)	0.64	2.53	0.1 (96)	2.07

Tablo 2'de verilen 1 numaralı birliktelik kuralı yorumlandığında; 249 işlem gününde EUYO, DYOBY hisseleri 188 kez beraber aynı kümede yer almış ve NETAS hissesi de bu iki hissenin beraber görüldüğü kümelerde 172 kez yer almıştır. Bu kuralın güven ölçütü ise 172/188 işlemi sonucu %91 olarak hesaplanmıştır. Kurallar incelendiğinde, önerilen yöntemin günlük fiyat oluşumlarında beraber hareket eden hisse gruplarının ortaya çıkarılmasında kullanılabileceği değerlendirilmektedir.

### 5. Sonuç ve Tartışma (Result and Discussion)

Bu çalışmada, veri madenciliği teknikleri kullanılarak BİST'te işlem gören hisse senetlerinin fiyat hareketlerinde birbirlerine olan etkilerinin ortaya çıkarılması amaçlanmıştır. Çalışma için 2019 yılında 249 işlem günü için 408 hisse senedi ve BİST endeksine ait günlük kapanış verileri elde edilmiştir. Yıllık ortalama işlem hacmi en yüksek olan 250 hisse senedi her bir işlem günü için K-Means yöntemi ile kümelendi ve oluşan kümelere Apriori algoritması kullanılarak birliktelik kuralı analizi yapılmıştır. Analiz sonucunda güvenilirlik yüzdesi en iyi 10 kural sonuçlarında verilmiştir. Sonuçlar yorumlandığında geliştirilen yöntemin, günlük fiyat değişimine göre birbirleri ile en çok hareket eden hisse senetlerine ait ilişki kurallarının ortaya çıkarılmasında kullanılabildiği değerlendirilmektedir.

### Çıkar Çatışması (Conflict of Interest)

Yazar tarafından herhangi bir çıkar çatışması beyan edilmemiştir. No conflict of interest was declared by the author.

### Kaynaklar (References)

- Borsa İstanbul, 2020. Pay Piyasası. Çevrimiçi (Erişim, 2 Mart 2020): <https://www.borsaistanbul.com/urunler-ve-piyasalar/piyasalar/pay-piyasasi>
- Bozma, G., Kul, S., 2020. Twitter ile Hisse Senetleri Oynaklığı Tahmin Edilebilir mi?. *Sosyoekonomi*, 28(45), 315-326.
- Dener, M., Dörterler, M., Orman, A., 2009. Açık kaynak kodlu veri madenciliği programları: WEKA'da örnek uygulama. *Akademik Bilisim'09 - XI. Akademik Bilisim Konferansı, Şanlıurfa*, 787-796.
- Ergür, B., 2014. Borsa İstanbul (BİST) hisse fiyat değişim yönünün ilişkisel borsa ağı kullanılarak tahmin edilmesi. *Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Türkiye*.
- Gazel, S., Akel, V., 2008. Borsa İstanbul'da sektör sınıflandırmasının kümeleme analizi ile belirlenmesi. *Muhasebe ve Finansman Dergisi*, 77, 147-164.
- Han, J., Kamber, M., 2006. *Data Mining Concepts and Techniques*, Morgan Kauffmann Publishers Inc., USA, 1-35.
- Özçakır, F.C., Çamurcu, A.Y., 2007. Birliktelik kuralı yöntemi için bir veri madenciliği yazılımı tasarımı ve uygulaması. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 6(12), 21-37.
- Özekes, S., 2003. Veri madenciliği modelleri ve uygulama alanları. *İstanbul Ticaret Üniversitesi Dergisi*, 3, 65-82.
- Soni, S., 2011. Applications of ANNs in stock market prediction: A Survey. *International Journal of Computer Science & Engineering Technology*, 2(3), 71-83.
- Tekin, B., 2018. Ward, K-Ortalamalar ve iki adımlı kümeleme analizi yöntemleri ile finansal göstergeler temelinde hisse senedi tercihi. *Balıkesir University The Journal of Social Sciences Institute*, 21(40), 401-436.
- Ünsal, Ö., 2011. Mesleki alan seçimlerinin makine öğrenmesi algoritması kullanılarak belirlenmesi. *Yüksek Lisans Tezi, Gazi Üniversitesi, Türkiye*.
- Weiss, S.M., Indurkha, N., 1998. *Predictive data mining: a practical guide*. Morgan Kaufmann Publications, USA.