

# Veri Madenciliği Teknolojisini Kullanarak Fiyat Değişimlerinde Paralellik Gösteren Hisse Senetlerinin Bulunması ve Risk Azaltılması

**Hüseyin AKTÜRK, Serdar KORUKOĞLU**

Ege Üniversitesi Bilgisayar Mühendisliği Bölümü, İzmir  
huseyinakturk@gmail.com, serdar.korukoglu@ege.edu.tr

**Özet:** Hayatımızın her alanında veriler gittikçe artmaktadır. Veriler arttıkça sahip olduğumuz verilerin ne anlama geldiğini ortaya çıkarmak gittikçe zorlaşmaktadır. Bunun sonucunda ise, ister istemez birçok önemli çıkarsamalar gözden kaçmaktadır. Data, bilgi ve knowledge(çıkarsama) farklı birer olgudur. Knowledge herkes için farklı olabilir. İşte veriler üzerinde veri madenciliği teknolojisi uygulanarak, knowledge seviyesine yakın bir takım çıkarsamalar bulabilir, ilginç sonuçlar ortaya çıkartabiliriz. Borsa verileri üzerinde de veri madenciliği teknolojisi uygulanarak ilginç sonuçlar ortaya çıkartılabilir. Bu makalede birbirine fiyat değişimleri bakımından benzerlik gösteren hisse senetlerinin bulunması ile ilgili bir konuya değinilecektir. Bu sayede borsa üzerinden para kazanmak isteyen kişiler benzer hisse senetleri sayesinde mevcut riski azaltabilme şansını yakalayabilecektir.

**Anahtar Sözcükler:** IMKB, Veri Madenciliği, Borsa, Hisse Senetleri, Kural Tabanlı Veri Madenciliği Analizi

## Finding The Price Changing Parallelsim Of Stock Exchanges Using Data Mining Technology And Decreasing Risk

**Abstract:** Data is growing on every stage in our life. While data is growing, it is difficult to understand the meaning of data that we have. As a result of this, we are missing a lot of inference although we don't want to miss these inference. As known, data, information and knowledge are different facts. Knowledge can be different for everybody. So, it is possible to find some inference that is near to knowledge level and it is possible to find interesting results applying data mining technology on our data. And we can find interesting results on stock exchanges data using data mining technology. In this essay, a subject will be explained that is deal with finding the price changing parallelsim of stock exchanges. So, the people that want to earn money with stock exchanges can have a chance to decrease current risk.

**Keywords:** IMKB, Data Mining, Stock Market, Stock Exchanges, Rule Based Data Mining Analyses

### 1. Giriş

Genel olarak yatırımcıların menkul kıymet varlıklarına yaptıkları yatırımları karşılığında birincil amaçları yüksek gelir elde etmek olurken bunun paralelinde kaybetme riski de

minimuma indirmeye çalışılır. Bu durumda belirsizlik altında karar verme sorunu ortaya çıkar. Eğer yatırım hisse senetlerine yapılmışsa portföyde çeşitlendirmeler yapılarak kayıp riskini azaltmak söz konusu olur[1].

## **2. Veri Madenciliği Nedir?**

Veri madenciliği, sahip olduğumuz verilerden üstü kapalı, çok açık olmayan, daha öncesinden bilinmeyen fakat potansiyel olarak işe yarayacak bilginin çıkarılmasıdır. Bu çıkarımda kümeleme, veri özetleme, değişikliklerin analizi, sapmaların tespiti gibi teknik yaklaşımları içerir (William J. Frawley, Gregory Piatetsky – Shapiro, Cristopher J. Matheus).

Veri madenciliği sonucu ortaya çıkartılan bilgi elimizde var olan ancak direkt olarak bakıldığında hemen ayırt edilemeyen bir bilgidir. Ortaya çıkartılığında, veriyi kullanacak kullanıcılara bu bilgi yönetildiğinde şaşkınlık yaratacak kadar önemlidir. İşte veri madenciliğinin büyüğü de tam bu noktadadır. Esasında veri madenciliği gizli kalmış bir takım örüntülerin ortaya çıkartılmasıdır. Bu çıkarım işleminde pek çok yöntem kullanılmaktadır. Ve bu çıkarsama işleminde pek çok aşamadan geçilmektedir.

Veri madenciliği faydalı, yeni olan ve mantıklı bir takım bilgiler elde etmek için çok büyük hacimlerde veri içeren veritabanları üzerinde birçok basamaktan oluşan işlemleri yapmaktır. Veri madenciliğinin temel amacı, gizli kalmış olan verileri ortaya çıkarmaktır.

Veri madenciliği birçok aşamadan oluşmaktadır. Bu aşamalardan desen ortaya çıkarmada pek çok algoritma vardır. Bu algoritmalar veri madenciliğinin olmazsa olmazlarıdır. Bu algoritmalar sayesinde gizli kalmış olan veriler oraya çıkarılabilmektedir.

## **3. Veri Madenciliği Ne Yapar?**

Son zamanlarda veri madenciliği pekçok şirket için pek anlamlı olmasa da, artık günümüzde pek çok kurum veri madenciliğinin değerini anlamaya başlamıştır. Bu anlam sonucunda

kendi alanlarındaki verileri nasıl işleyebileceklerini, verilerden nasıl bir takım sonuçlar çıkartabileceği konusunda bir takım araştırmalara gitmektedir. Günümüzde, kurumsal kaynak planlama, müşteri ilişkileri yönetimi, sağlık ve daha pek çok alanda veri madenciliği aktif bir şekilde rol almaktadır.

İş hayatında, veri madenciliği yeni bir takım desenlerin ortaya çıkartılmasında oldukça önemlidir. Bu çıkarsamalar sonucu firmalar, üretimlerinden tutun insan kaynaklarını yönetimini, ürün sergileme yöntemlerini ve pazarlama stratejilerini bile değiştirebilmektedir. Rekabatchi bir dünyada ürününüzü, hizmetinizi müşteriye sunabilmek ve sunduktan sonra da satın alınmasını sağlamak için veri madenciliği çok önemli rol oynamaktadır.

Günümüzdeki veri madenciliğinin kullanılmış olduğu alanların bazılarını şu şekilde sıralayabiliriz:

- Reklam Hizmetleri
- CRM ( Müşteri İlişkileri Yönetimi )
- E – Ticaret
- Spor alanları
- Eğlence Sektörü
- Üretim Bilgi Sistemleri
- Pazarlama Stratejileri
- Telekom Sektörü
- Astronomi
- Sağlık Sektörü
- Arama Motorları
- Metin Madenciliği (Bazı haber siteleri, E-Mail, çeşitli dökümanlar)
- Web Sitesi Analizleri
- Terör ile Mücadele
- Vergi kaçakçılarının profiline çıkartılması



## **5.2. Veri Ön İşleme**

Veri madenciliğine başlamadan önce kullanmak istediğini verilerin temiz olması, hatasız olması çok önemlidir. Veri madenciliği sonradan yapılan bir işlemdir. Yani öncelikli olarak sürecin işlenmesi önemlidir. Daha sonrasında ise veri madenciliği işlemi yapılır. Bu nedenle, mevcut verilerinizde hata çıkması durumu gayet doğaldır. Bu hataların ayıklanması işleminde pek çok yöntem kullanılabilir.

Dataların temizlenmesi, hataların düzeltilmesi işleminde ilk önce başvurulacak olan yöntemlerin başında hatalı olan verilerin başka kaynaklardan yeniden okunması işlemi yapılır. Bu işlem esnasında farklı kaynaklardan gelebilecek olan doğru veriler bizim için artık temizlenmiş, hatası düzeltilmiş olan verilerdir.

Verilerin temizlenmesi esnasında eğer veri çok önemli bir bilgi taşıyor ise, bu bilginin oluşumu da başka sistemler yardımı ile yapılıyor ise verinin oluşumu için diğer sistemler sadece bizim sistemimize doğru veri sağlaması açısından tekrar çalıştırabilir. Tabi burda unutulmaması gereken maliyete göre hareket edilmesidir. Ve eğer bu sistemlerin yeniden çalıştırılması eğer insanların bir takım bilgileri girmesi isteniyor ise ve maliyeti karşılanabilecek düzeyde ise yapılmalıdır.

Veriler eğer hiçbir şekilde geri kazanılmıyor ise bu veriyi kullanmak çok doğru olmayacaktır. Çünkü veri madenciliği geçmişteki verilerin doğru olduğunu kabul etmesi önemli bir olgudur.

## **5.3 Desen Geliştirimi, Veri Madenciliği İşlemi**

Veri madenciliği işleminin asıl yapılmış olduğu işlem burada gerçekleşmektedir. Bu kısımda kullanılacak olan algoritmaların seçimi, verilerin daha hızlı nasıl analiz edileceği gibi bir takım değerlendirmeler yapılmaktadır. Algoritmalar kullanılır iken algoritmaya ait bir

takım katsayıların belirlenmesi işlemi yine burada yapılmaktadır. Bu tip bazı belirlemeler yapıldıktan sonra veri madenciliği işlemi başlatılır. Ayrıca veri madenciliğinin temel amacı ilginç olan bir takım desenleri ortaya çıkarmak olduğu için size göre ilginç olmayan bir takım desenlerin de burada elimine edilmesi gerekmektedir. Gerekirse algoritmaya ait bir takım sabitleri yeniden belirlemek ve süreci tekrar başlatmak bir çözüm olabilecektir.

## **5.4. Çıkarılmaların Gösterimi**

Eğer çıkarılmalar oldukça ilginç olsa bile eğer bunu karşınızdaki insanlara etkili bir şekilde anlatamıyor iseniz malesef çok fazla bir iş yapmış sayılmazsınız. Çünkü işin asıl önemli noktası sizin sonuçlarınız kullanacak olan kişilerin sizin yapmış olduğunuz çalışmanın sonuçlarını anlayabilmesidir. Bu aşamada mümkün olduğunda görselliğe önem vermek gerekmektedir. Çıkarılmalarınıza ait sonuçların her bir objesini ilişkilendirmeniz önemlidir. Bir takım görsellik işlemlerini ön planda tutan programların kullanılması gerekmektedir.

## **6. Hisse Senetlerinin Analizi**

Günümüzdeki hisse senetlerinin gittikçe artması ve bunun sonucunda bir takım analiz işlemlerinin yapılması önemli bir iş haline gelmiştir. Özellikle belirleyici birçok etkenin bulunmuş olduğu bu sektörde analiz işlemleri pek çok farklı şekilde yapılabilir. Buradaki amaç hisse senetlerinin alım satımını yapan insanların işlerini daha kolay bir şekilde yapmasını sağlamak, insanların analiz yapar iken onlara bir takım çıkarılmalar sunmaktır. Bu çalışma sonucunda ise insanlar eğer sizin onlara sunmuş olduğunuz bu işlemlerden fayda sağlamış ise yapmış olduğunuz çalışma başarılı olmuş demektir.

Sadece, hisse senedi alım satımı yapan kişiler değil aynı zamanda pek çok firmada kendi hisse senetleri olsun başka hisse senetleri ol-

sun bir takım analiz işlemleri yapmaktadırlar. Çünkü kendi hisse senetleri başka hisse senetleri tarafından da etkilenmektedir. Bu nedenle tüm hisse senetlerinin analiz edilmesi firmaların vazgeçilmez bir işidir.

Elbetteki sadece hisse senetlerinin fiyatlarının analiz edilmesi sonucu bütün işlemler yapılmış olmaz. Bazı bir takım risklerin incelenmesi, hisse senetlerinin fiyatlarını değiştiren diğer etkenlerin de incelenmesi gerekmektedir. Ancak bu işlemlerin hepsini tek bir projede incelemek mümkün olamaz. Bu nedenle siz insanlara ya da firmalara ne kadar bilgi sunabiliyor iseniz bu onlar için önemli bir değerlendirme kriteri olacaktır.

Bu makalede birbirlerine fiyat benzerliği bakımından benzerlik gösteren hisse senetlerinin incelenmesi bunun sonucunda bazı bir takım çıkarsamaların yapıldığı anlatılmaktadır. Burada temel değerlendirme noktası, referans olarak bir hisse senedinin seçilmesi işlemi olmaktadır. Bu referans hisse senedinin işleminin seçilmesi ile birlikte gerek uzun vadede gerekse kısa vadede referans hisse senedine fiyat değişimi bakımından benzerlik gösteren diğer hisse senetlerini de bulabilme imkanı oluşmaktadır.

Makale sadece referans bir hisse senedinin seçilmesi ve diğer hisse senetlerinin buna olan benzerliğini içermemektedir. Referans olarak endeksi seçebilme ya da referans olarak döviz kurlarından birini de seçebilme imkanı vermektedir. Bu sayede, endeksin fiyat değişimi ile diğer hisse senetlerini fiyat değişimlerini analiz etme imkanı da bulunmaktadır.

Sürekli olarak fiyat değişimlerinin benzerliğinden bahsettik. Bu fiyat değişimleri benzerliğini sorgulayacak olur isek, burada yapılmış olan çalışma referans menkul değerinizin verilen tarih aralığında, verilen periyotlarda fiyatının artması sonucu artan diğer hisse senetlerinin

bulunması, fiyatının azalması sonucu azalan diğer hisse senetlerinin bulunması işleminden ibarettir. Bu işlemler yapılırsa, benzerlik oranının verilmesi de zorunlu bir sabittir. Bu sayede verilen tarih aralığında alınan iki periyot arasında toplam periyoda göre yüzdelik olarak benzerliğin bulunması sağlanmaktadır.

Çalışmada kullanılan veri madenciliği algoritması her ne kadar kümeleme algoritması olarak gözüke de bu çalışmada kural tabanlı algoritmalar kullanılmıştır. Kural tabanlı algoritmanın temel girdisi olarak ise, diğer hisse senetleri olmuştur. Biraz açıklayacak olur isek, iki periyot arasında eğer referans hisse senedi düşmüş ise, bu iki periyot arasından fiyatı düşen diğer hisse senetleri bulunmuştur. Daha sonra diğer periyotlar aynı şekilde incelenmiş olup fiyat değişimleri bakımından benzer olan diğer hisse senetleri kümesi elde edilmiştir ve kural tabanlı algoritmaya her bir küme tek tek input olarak sağlanmıştır. Verilen yüzdelik değeri de toplam periyot aralıklarına göre oranlarında support değeri olarak düşünülmüştür.

## **7. Çalışmaya Ait Örnekler**

Bu makaleden önce uygulama gerçekleştirilmiştir. IMKB den 2005 yılından önceki 10 yıllık veriler akademik çalışma amaçlı alınmıştır[2]. Aynı zamanda uygulamada Türkiye Cumhuriyet Merkez Bankasına ait veriler de kullanılmıştır[3]. İlerki başlıklarda çalışmaya ait örnekler gösterilecektir.

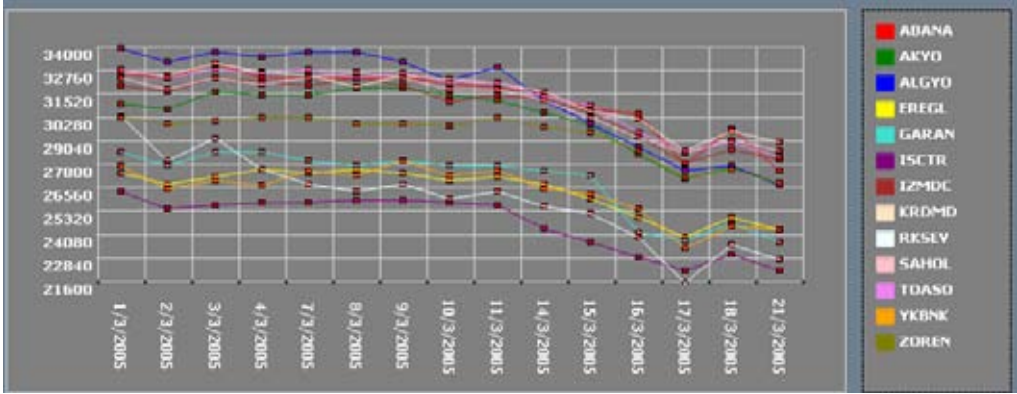
### **7.1. Referans Olarak Hisse Senedinin Seçilmesi**

Şekil 2 ye bakıldığında sağ tarafta yer almakta olan hisse senedi isimlerinden en tepedeki hisse senedi referans olarak seçilen hisse senedir. İki tarih aralığı 01/03/2005 ile 21/03/2005 tarih aralığıdır. Periyot olarak ise 1 gün seçilmiştir. Ve görüleceği üzere ABANA hisse senedinin değeri artar iken diğer hisse senetlerinin değerleri artmış, ABANA iki periyot

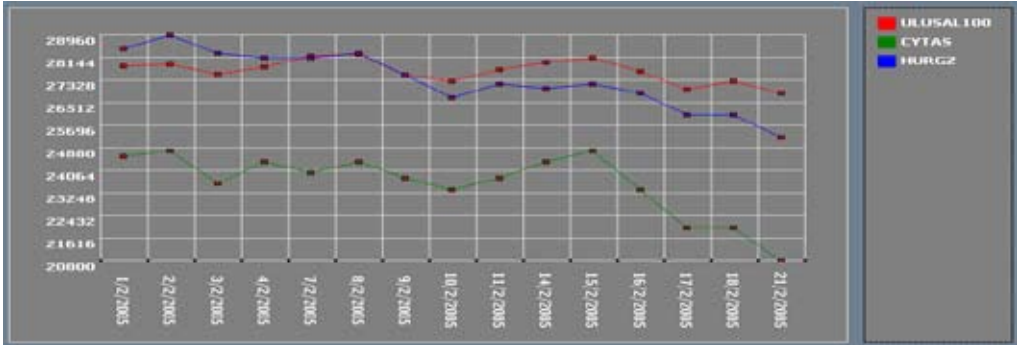
arasında sabit iken diğer hisse senetleri de sabit olmuş, ABANA iki periyot arası düşer iken diğer hisse senetleri de düşmüştür. Burada verilen benzerlik oranına göre bu işlem yapılmıştır. Yani fiyat artışı, azalış ya da değişmeme durumu her periyotta bütün hisse senetleri için benzer değildir.

## 7.2. Referans Olarak Endeksin Seçilmesi

Şekil 3 e bakıldığında ULUSAL100 endeksi referans olarak seçilmiştir. Verilen tarih aralığı 01/02/2005 tarihi ile 21/02/2005 tarih aralığıdır. Ve ULUSAL100 endeksine bağlı olarak fiyat değişimleri bakımından benzerlik gösteren diğer hisse senetlerini görmekteyiz.



Şekil 2 : Referans Hisse Senedi ve Diğer Hisse Senetleri



Şekil 3. Endeks ve Hisse Senetleri

## 8. Sonuç

Rekabet ortamlarının artması sonucu daha hızlı çözümler üretebilmek, daha hızlı işlemler yapabilmek için artık işlemlerimiz bilgisayarlar vasıtası ile yapılmaktadır. Bu işlemlerin bilgisayarlar da yapılması sonucu büyük bir

ivme ile saklanan veriler artmaktadır. Daha rekabetçi olabilmek, sahip olduğumuz potansiyelleri değerlendirebilmek, daha karlı çözümler sunabilmek için veri madenciliği artık vazgeçilmezler arasında yerini gün geçtikçe almaktadır. İş dünyasında ise borsa önemli bir yere sahiptir. Bu neden borsa verilerinin gü-

zel bir şekilde analiz edilebilmesi, bu işler ile uğraşan insanların daha karlı işlemler yapabilmesi, almış oldukları riski mümkün olduğunda minimuma indirgenebilmesi için borsa verileri üzerinden veri madenciliğinin işleminin yapılması önemli bir konumdadır. İnanıyorum ki, bu çalışma bundan sonraki pek çok çalışma için bir referans olacaktır ve faydalı bir takım işlere imzasını atacaktır.

## **9. Kaynaklar**

[1]. KORUKOĞLU, Serdar ve Ayşen, İMKB’de İşleme Gören Mali Kuruluş Hisselerinin Beta ve Alfa Risk Katsayılarına Dayalı Veri Madenciliği.

[2]. İstanbul Menkul Kıymetler Borsası Web Sitesi, İMKB Akademik Çalışma Verileri, <http://www.imkb.gov.tr/>, İstanbul Menkul Kıymetler Borsası

[3]. Türkiye Cumhuriyet Merkez Bankası Web Sitesi, T.C.M.B Merkez Bankası Verileri, <http://www.tcmb.gov.tr/>, Türkiye Cumhuriyet Merkez Bankası