

Finansal Suçların Tespitinde Veri Madenciliği Yaklaşımı ve Literatüre Bakış

M. Fevzi ESEN

Yrd. Doç. Dr., İstanbul Medeniyet Üniversitesi
Turizm Fakültesi, Turizm İşletmeciliği A.B.D.
fevzi.esen@medeniyet.edu.tr

Finansal Suçların Tespitinde Veri Madenciliği Yaklaşımı ve Literatüre Bakış

Özet

Sadece Amerika Birleşik Devletleri hisse senedi piyasalarında günlük ortalama işlem miktarının 7 milyar adet olarak gerçekleştiği bile baz alındığında, stratejik, taktik ve operasyonel karar süreçlerindeki problemlerin daha düşük maliyetle ve yüksek güvenilirlikle çözülebilmesi için veri içerisinde saklı bulunan bilgilerin keşfedilmesi gerektiği bir gerçektir. Veri madenciliği olarak adlandırılan bu bilgi keşfi süreci; risk ve portföy yönetimi gibi bankacılık uygulamalarının yanısıra; şirketlerdeki finansal raporlamaların denetlenmesi ve piyasa oyuncuları arasında doğru bilgi akışının sağlanmasında etkin bir şekilde kullanılmaktadır.

Bu çalışmada, 1994 - 2015 yılları arasında yayınlanan 79 adet bilimsel makale, finansal suç kategorisine göre sınıflandırılmış ve veri madenciliği tekniklerine göre değerlendirilmiştir. Çalışmada, veri madenciliği tekniklerinin çoğunlukla bankacılık ve sigorta suçlarının tespitinde kullanıldığı tespit edilmiş olup; finansal suçların veri madenciliğiyle tespiti ve tahminlenmesine yönelik Türkiye'deki çalışmaların yetersiz olduğu sonucuna varılmıştır.

Anahtar Kelimeler: Finansal Suçlar, Veri Madenciliği, Hile Tespiti, Literatür Çalışması

Data Mining Approach In Financial Fraud Detection and a Literature Review

Abstract

Only in USA Stock Exchanges, daily average trading volume is about 7 billion units. Just depending on this statistics, the necessity of information discovery hidden in data is a reality to tackle the problems in strategic, tactical and operational decision processes with lower costs and higher reliability. Information discovery from databases, namely, data mining is an effective method for auditing financial statements in companies and providing flow of information between market players as well as risk and portfolio management as banking applications.

In this study, 79 journal articles related to the subject published 1994-2015 have been classified and evaluated in terms of data mining techniques. It has been found that data mining techniques have been applied most extensively to detection of banking and insurance fraud. However, the findings of literature review show that the number of studies in detection and prediction of financial fraud is not enough in Turkey.

Keywords: Financial Fraud, Data Mining, Fraud Detection, Literature Review

Kabul Tarihi / Accepted: 30.09.2015

1. Giriş

Piyasalarda bilgi kaynaklı problemlerin sebep olduğu hileli işlemlerin piyasa etkinliğini ve yatırımcıların piyasaya olan güvenini sarstığı bir gerçektir. Bilgisizlik ve dikkatsizlik sonucu gerçekleştirilen "hatalar" hariç; kanunlara ve yönetmeliklere aykırı bir biçimde ve yetkilendirilmeden, haksız kazanç sağlamak amacıyla gerçekleştirilen aldattıcı finansal işlemlerden dolayı dünya ekonomisi her sene yaklaşık 2.1 trilyon dolar kayba uğramaktadır (Financial Conduct Authority, 2013).

Miktarı gün geçtikçe artan hileli finansal işlemlere en iyi örnek, mal veya hizmet almaksızın katılımcıların sisteme dahil ettiği nakit üzerinden diğer katılımcıların sürekli artan bir kazanç sağlaması prensibine dayanan ve ilk olarak 1920 yılında Charles Ponzi tarafından gerçekleştirilen titan saadet zinciridir. 2009 yılında piyasaları yaklaşık 50 milyar dolar dolandırarak 150 yıl hapis ve 17.1 milyar dolar para cezasına çarptırılan ve tarihin en büyük ekonomik skandallarından birine imza atan NASDAQ eski başkanı Bernard Madoff vakası; ponzi oyunu, menkul kıymet suçları (securities fraud), yatırım danışmanlığı suçları (investment advisor fraud), hileli finansal raporlama, kara para aklama (money laundering) gibi finansal suçlara örnektir.

Sertifikalı Hile Denetçileri Birliği'ne (ACFE) göre her sene işletmelerin gelirlerinin yaklaşık %5'i finansal hileler dolayısıyla erimekte ve bu kayıp işletme başına senelik ortalama 140 bin dolara mal olmaktadır (ACFE, 2012). 2014 yılında yayınlanan başka bir rapora göre, her üç işletmeden birinin finansal suça karıştığını belirtilmekte ve bu işletmelerin yaklaşık %54'ünün yüzbin dolar ve üzerinde ekonomik kayba uğradığı öngörülmektedir (PricewaterCoopers, 2014). Diğer taraftan, Avrupa'da her sene, toplam sigorta harcamalarının yaklaşık %10'u hileli işlemler dolayısıyla erimekte ve tespit edilemeyen sigorta suçlarından dolayı sadece İngiltere ekonomisi her sene yaklaşık 2,2 milyar euro zarara uğramaktadır (Insurance Europe Report, 2013: 9-10).

Tahmin edilen direkt ekonomik kayıpların yanısıra; verimlilik kaybı, genel işletme giderleri, çalışanların ve piyasa oyuncularının motivasyon kaybı, oynak hisse senedi fiyatları gibi indirekt maliyetler de söz konusu kayıba dahil edildiği takdirde, finansal suçlardan kaynaklanan zararın boyutları daha net anlaşılmaktadır.

Niteliksel olarak değerlendirildiğinde, piyasalarda finansal sahtecilik suçlarının hile, hırsızlık ve kötüye kullanma kapsamında işlendiği; yatırımcıların veya işletmelerin yalan, bilgi saklama ve güveni ihlal etme yoluyla veya şiddet, istismar, teknoloji suçlarıyla organize bir biçimde menfaat temin ettiği belirtilmektedir. Hileli çek, senet ve kredi kartı işlemleri, rüşvet, vergi kaçırma, finansal evrak ve mali tablolarda sahtecilik, kara para aklama, sigortacılık suçları, zimmete geçirme gibi büyük çoğunluğunu beyaz yaka suçlarının oluşturduğu finansal eylemler, ABD ve Avrupa Birliği literatüründe Tablo 1'deki gibi sınıflandırılmaktadır.

Tablo 1. Finansal Suçların Sınıflandırılması (FBI, 2011)

Finansal Suç	Aktiviteler
Bankacılık Suçları	Mortgage Dolandırıcılığı, Kara Para Aklama, Kredi Kart Dolandırıcılığı
Sigorta Suçları	Araç Sigortası Dolandırıcılığı, Sağlık Sigortası Dolandırıcılığı, Konut Sigortası Dolandırıcılığı, Hayat Sigortası Dolandırıcılığı, Tarım Sigortası Dolandırıcılığı
Kıymetli Evrak Suçları	İçeriden Öğrenenler Ticareti, Manipülasyon, Front Running, Ponzi Şemaları, Hedge Fonu Suçları
Diğer Suçlar	Şirket Yolsuzlukları, Vergi Kaçakçılığı, Hileli Finansal Raporlama

Kriminolojik olarak ise finansal suçlar faillerin sosyal statüsüne göre de sınıflandırılmakta olup, "beyaz yaka suçları" olarak adlandırılan ve geleneksel suçlardan ayrılan söz konusu eylemlerin, işletme veya kamu kurumlarında teknik bilgi, beceri ve mesleki konum kullanılarak gerçekleştirildiği gözlenmektedir. Finansal suçların büyük çoğunluğu beyaz yaka suçları arasında incelenmekte olup bazı vakalar hırsızlık, soygunculuk, gasp gibi adi suçlarla organize gerçekleştirilmektedir.

Bu çalışmada, finansal suçların tespitinde kullanılan veri madenciliği tekniklerinden bahsedilmiş olup, literatürde bulunan çalışmalar söz konusu suç kategorilerine göre gruplandırılarak incelenmiştir.

2.Finansal Suç Tespitinde Veri Madenciliği

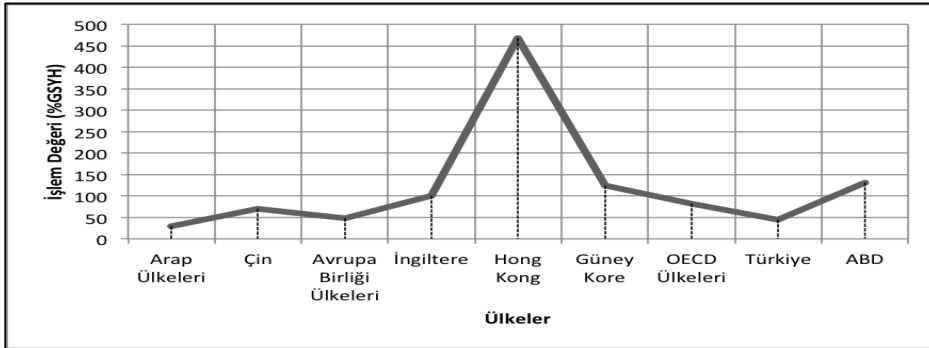
Finansal işlemler içerisindeki hileli davranışları, aktiviteleri tespit edebilmek ve karar vericilerin finansal suçların etkisini azaltabilecek uygun stratejiler geliştirmelerini sağlamak amacıyla finansal suç tespitinin etkin şekilde gerçekleştirilmesi gerekmektedir. Piyasalarda gün geçtikçe karmaşılaşan işlem karakteristikleri ve artan işlem hacmiyle birlikte iş ilişkilerinin de genişlemesi, finansal suçların tespitinde denetçi raporları, mülakatlar, sübjektif yorumlar ve çeşitli teknik göstergeler gibi klasik tekniklerin bütünüyle ve derinlemesine kullanılmasını gerekli kılmaktadır.

Günümüzde finansal faaliyetler sonucu ortaya çıkan kantitatif ve kalitatif verinin karar mekanizmalarında kullanılmak üzere çeşitli tekniklerle bir araya getirilip saklı trendlerin ve anlamlı ilişkilerin ortaya çıkarılmasında veri madenciliği etkin bir şekilde kullanılmaktadır. Finansal suçların tespitinde önemli bir analitik yaklaşım olan veri madenciliği; istatistik, matematik, yapay zeka, örüntü tanıma ve makine öğrenmesi gibi hibrit tekniklerle geniş veritabanlarından faydalı bilgiyi çıkarıp tanımlama ve profilendirme işlemi olarak belirtilmektedir (Turban vd., 2010).

Çok büyük hacimdeki verilerden yeni trendler, biçimler, modeller bulma süreci olarak adlandırılan veri madenciliği, ham verinin tek başına sunamadığı bilgiyi, bilgisayar teknolojileri kullanılarak ortaya çıkaran keşif süreci olarak tanımlanmaktadır (Groth, 2000: 3). Metodolojisi 1996 yılı sonlarında “The Cross-Industry Standart Process for Data Mining (CRISP – DM) konsorsiyumu tarafından oluşturulan veri madenciliği, verilerin çeşitli tekniklerle temizlenmesinin ardından tespit edilen problemin amacına uygun olarak, çeşitli analitik modellemelerle çıkarımlar yapılması süreçlerini içermektedir (Larouse, 2004: 215). İlk örnekleri 1990 yılında görülen veri madenciliği uygulamalarına; finansal suç tespitinin yanı sıra; astronomi, biyoloji, pazarlama, ulaştırma ve lojistik gibi birçok alanda rastlanmaktadır (Meyer ve Cannon, 1998: 186).

Ölçümler, gözlemler ya da araştırma yoluyla elde edilen nitel ve nicel veriler toplanıp gruplandırıldıktan sonra işlenip, dönüştürülmekte ve neticesinde anlam kazanmaktadır. 2012 yılı itibariyle finansal işlemlerin de dahil edildiği elektronik ortamda depolanan veri hacmine bakıldığında, veri içerisinde gizli kalmış kuralların, ilişkilerin, trendlerin keşfedilebilmesi için istatistik, makine bilgisi, veritabanı gibi disiplinlerin etkin şekilde kullanılması gerektiği anlaşılmaktadır.¹ Yapılan araştırmalarda, söz konusu verinin %3'lük bir kısmının ön-işleme süreçlerinden geçirilip analiz için hazır hale getirildiği; sadece %0,5'lik kısmının analiz edildiği belirtilmektedir (EMC², 2014).

Grafik 1. 2012 yılı Global Piyasalarda Menkul Kıymet İşlemlerinin GSYH içerisindeki Payı (World Bank, 2014)



Global piyasalarda gerçekleştirilen menkul kıymet alım satım işlemi ve hacmi, finansal piyasalarda suç tespiti ve takibinin etkin şekilde gerçekleştirilmesinin ekonomik açıdan önemine dikkat çeken en iyi göstergelerinden birisidir. 2012 yılı

¹Araştırmalara göre, 2012 yılı itibariyle global veri arzı 2.8 trilyon gigabyte olarak ölçülmüş olup bu rakamın her sene %23 oranında artacağı tahmin edilmektedir (Bkz.: "EMC² Data Report", (Çevrimiçi) <http://www.emc.com>, 21.04.2014).

Dünya Bankası raporlarına göre, global piyasalarda işlem gören toplam menkul kıymet hacminin ülkelerin toplam GSYH içindeki payı ortalama %70.1 olarak belirtilmektedir. Yani ülkelerin ürettiği tüm nihai mal ve hizmetlerin büyük çoğunluğunu menkul kıymet ticareti oluşturmaktadır (Grafik 1).

Spesifik olarak bakıldığında, sadece New York Borsasında günlük ortalama 170 milyar dolar değerinde 2,6 milyon adet işlem gerçekleştirilmektedir (NYSE, 2014). Kredi kartı yoluyla yapılan hileli işlemlere bakıldığında ise, 2012 yılında ABD'de 26.2 milyar adet işlem gerçekleştirildiği ve 3,2 trilyon dolar işlem hacmi gerçekleştirildiği belirtilmektedir (Federal Reserve Bank, 2013). İşlemlerin yaklaşık %0,1'inin hileli olduğu düşünülürse, günümüzde dahi ekonomik kaybın senelik 2,7 milyar dolar civarında olduğu tespit edilmektedir (Lisboa vd., 2000: 141-158). Söz konusu değerlere bakıldığında, hile tespiti için işlemlerin elektronik ortamda düzenlenmesi ve uygun sayısal yaklaşımlarla analiz edilmesi sürecine sistematik olarak yaklaşılması gerektiği anlaşılmaktadır.

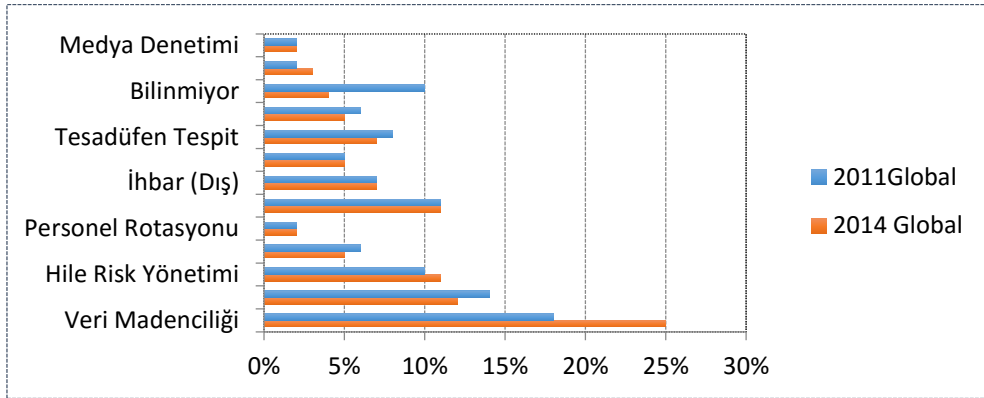
Çok sayıda değişken davranışının incelenmesini gerektiren durumlarda etkin şekilde kullanılan veri madenciliği; verinin kontrolü ve analizinin klasik istatistik metotlarla zorlaştığı durumlarda kullanılabilir (Bozdoğan, 2004: 6-7). Finansal suç tespiti problemlerini statik zaman boyutundan çıkarıp dinamik ve durum odaklı açıdan değerlendirebilen veri madenciliği, geçmiş örüntüleri etkin bir şekilde modellemekte; veri içerisindeki karmaşık ve yeni örüntüleri algılayıp öğrenerek, akılcı kararlar verme becerisi kazanmakta ve verinin çok yönlü analizine imkan tanımaktadır (Cahill vd., 2002: 911-930). Karar vericilere ayrıntılı bilgi ve hile tespitinde farkındalık sağlayan veri madenciliği, yüksek hassasiyette suç tespiti sağladığı için birçok vakanın gerçekleştirilmesini engelleyebilmektedir (Singleton, 2010: 160-163).

Veri madenciliği, veri setindeki normal davranış üzerine profil geliştirilip anomalilerin tespit edilmesini mümkün kılan analitik yaklaşımlar bütünüdür. Problemin farklı açıdan ele alınıp tanımsal ve tahminsel olarak modellenebilmesini sağlayan veri madenciliği, bilgisayar destekli ve hızlı, etkin algoritma temelli çalıştığı için klasik yöntemlere kıyasla hile tespitinde zaman ve maliyet etkindir (Clifton vd., 2010: 1). Olasılık yöntemlerle finansal işlemlerde oluşabilecek hileli durumları önceden tahminleyebilen veri madenciliği, hilenin izlenmesi ve kontrolü süreçlerindeki manelliği azalttığı için, suç tespitinin etkisi ve etkinliği arasındaki ikilemi çözümleme potansiyeline sahiptir (Wang, 2010: 50-53).

Finansal suçlarla mücadele edebilmek amacıyla piyasalarda hile veya usulsüzlükten kaynaklanan işlemlerin engellenebilmesi için çeşitli düzenlemeler getirilmektedir. Piyasa güvenini tesis etmek adına, uyarı ve gözetim faaliyetleri çeşitli kamu ve özel

sektör kuruluşları tarafından yürütülmekte ve piyasalarda kamuyu aydınlatma ilkesi gereği her türlü denetim raporu ve tablolar yatırımcılarla paylaşılmaktadır. Klasik denetim prosedürleriyle tespiti zor olan hile ve usulsüzlükleri, büyük miktarda ve kompleks yapıdaki veriye etkin yaklaşımıyla tespit eden veri madenciliği, finansal hilelerin tespit ve takibinde kullanılan en etkin yaklaşım olarak gösterilmektedir (PricewaterhouseCoopers, 2014).

Grafik 2. Finansal Hile Tespitinde Kullanılan Metodlar (PricewaterhouseCoopers, 2014: 41)



Veri madenciliği temelde iki yaklaşımda sınıflandırılmaktadır. Bunlar, araştırma probleminin özelliklerini, etki eden değişkenleri ve önem derecelerini ortaya koyan klasik istatistik tekniklerin kullanıldığı tanımlayıcı yaklaşımlar ve veride önceden bilinmeyen eğilimleri, örüntüleri keşfetmeye yardımcı olan tahminleyici yaklaşımlardır (Giudici, 2003: 6). Başka bir yaklaşıma göre veri madenciliğinde kullanılan modeller sınıflandırma, kümeleme, birliktelik kuralları ve zaman serileri olarak dört başlıkta incelenmektedir (Han ve Kamber, 2001: 327-443).

Veri madenciliği sınıflandırma tekniklerinde amaç, verinin özelliklerini açıklayarak önceden tanımlanmış sınıflara atamasının yapılmasıdır. Karar ağaçları, naive bayes, kernel kestirimi ve yapay sinir ağları sınıflandırma tekniklerine örnek olarak verilebilir. Tahmin yöntemlerinde en sık kullanılan teknikler ise doğrusal ve doğrusal olmayan regresyondur.

Heterojen bir veri setini homojen gruplara ayırmada kullanılan kümeleme teknikleri, önceden tanımlanmış sınıfları baz almadan benzerliklerine göre bir araya getirmek için kullanılmaktadır. Amaç, birbirine çok benzeyen veriyi bir araya getirip farklı kümelerle bölmektir. Ward, bölünmüş ortalamalar, k-ortalamlar kümeleme içerisinde incelenen hiyerarşik ve hiyerarşik olmayan tekniklere örnek olarak verilebilir.

Bir veya daha fazla deęişken arasındaki ilişkinin varlığı ve gücünü ölçek için kullanılan regresyon analizi, ilişkinin fonksiyonel şeklinin yanı sıra; bir dięeri hakkında tahmin yapılmasına olanak sağlamaktadır. Veri içerisindeki potansiyel ilişkileri tanımlayan birliktelik kuralları, birlikte gerçekleşebilecek olayların olasılık tahminlemesinde bulunmaktadır.

Deęişimin zaman baęlı olduęu serilerde, deęişkenlerin özelliklerini açıklamak ve zaman içerisindeki deęişimini inceleyerek seriyi etkileyen bileşenleri tespit etmek amacıyla kullanılan zaman serileri, gelecek dönemlere ilişkin tahmin çalışmalarını kapsamaktadır. Ekonometri ve matematiksel finasta yaygın kullanıma sahip olan zaman serileri, stokastik yaklaşımların en yaygın kullanım alanına sahip tekniklerinden biridir.

3.Literatür Çalışması

Finansal suçların tespiti, büyük miktardaki veri arasından küçük ve önemsiz görünen kayıtların detaylıca incelenmesiyle mümkün olmaktadır. Finansal piyasalarda hile bazen tekil kayıtlarda tespit edildięi gibi, bazen de dięer işlemlerle komplike bir şekilde büyük örüntülerde tespit edilmektedir. Hileli işlemlerde bulunanların yeni hile tespit tekniklerini aşabilecek nitelikte işlem stratejileri geliştirmeleri veya rutin olarak gerçekleştirilen hileli işlemlerin tespiti için klasik ve modern tekniklerin bir arada kullanılması gerekmektedir. Bu sebeple finansal suçların tespitinde kullanılan veri madencilięi teknikleri, problemin tanımına göre farklılık göstermektedir. Örneğin finansal tablolarda hileli işlemleri belirlemede önceden tespit edilmiş ve sonuçları bilinen vakaların verileri (training data) ilişki - sonuç bağlamında matematiksel bir fonksiyona dahil edilerek başka veriler (test data) için çıkarım yapılabilecek teknikler (gözetimli) olduęu gibi; şirket içi yolsuzluklarda yöneticilerin iş ilişkilerini ortaya çıkaracak veya menkul kıymet alım satım işlemlerindeki anormal aktiviteleri tespit edebilecek baştan bilgi verilmesini gerektirmeyen teknikler (gözetimsiz öğrenme) de bulunmaktadır.

Önceden tespit edilmiş hileli aktiviteleri (normal ve hileli işlemlerin sınıflandırıldığı) yeni gözlemleri sınıflandırmada kullanan gözetimli öğrenme metotlarının etkinlięi ilk olarak Hand (1981) ve McLachlan (1992) tarafından doğrusal ve lojistik diskriminant analizi gibi klasik istatistik tekniklerle kullanılarak ispatlanmıştır. Sonraları, geleneksel yöntemlerin yetersiz kaldığı problemlerin çözümü için yapay zeka tekniklerinin ilk formu sayılan ve mantıksal bilgiyi sisteme kodlayan "kuralla dayalı uzman sistemler" (rule based expert systems) kullanılmaya başlanmıştır (Hand, 1997). Bayes sınıflandırmaları, karar ağacı algoritmaları, yapay sinir ağları gibi doğru sınıflandırma skorunu arttıran teknikler ve birçok hibrit kombinasyonları

suç tespiti problemlerinde kullanılan ilk sayısal yaklaşımlara örnek olarak verilebilir (Chan vd., 1999: 67-74).

Hile tespitinde kullanılan bir diğer yaklaşım ise, Benford kanuna dayanan dijital analizdir. Özellikle 1997 yılından sonra muhasebe denetiminde uygulanan bu teknik, veri kümesi içerisindeki sayılarda kullanılan rakamların olasılık dağılımlarından hareketle anomalilerin tespit edilmesini amaçlamaktadır (Nigrini, 1999: 79-83 ; Wallace, 2002: 16-22 ; Durtschi vd., 2004: 17-34). Buna göre, bir finansal kayıt veya muhasebe verisi içerisindeki rakamlar logaritmik olarak sıralanmakta ve veri içerisinden rasgele seçilen bir sayının ilk hanesinde herhangi bir rakamın gözlenme olasılığı eşit dağılmamaktadır. Olasılıklar, her bir rakam için ayrı olarak hesaplanacağı gibi, rakam kümeleri için ayrıca hesaplanabilmektedir. Gözlemlenen ve teorik olarak tahminlenen veriler arasındaki fark, olası bir hileye işaret etmektedir (Benford, 1938: 551-572).

Bir başka yaklaşımda, finansal hile tespitinde veri madenciliği tekniklerini kantitatif olmayan (deney, anket, mülakat), basit istatistik (histogram, sebep-sonuç ve dağılım diyagramları) ve yapay zekayla sofistike gelişmiş kantitatif teknikler olarak üç kategoride sınıflandırılmaktadır (Fraser vd., 1997: 35-47).

Mevcut tekniklerin başarısının veri setinin hiper parametre değerlerine göre farklılık göstermesi yani bazı veri setleri için başarılı sonuçlar alınamaması ve uygun tekniğin seçiminde yöntem ve parametrelerin teker teker denenerek zaman, maliyet kaybına yol açması veri karakteristiklerinin incelenmesi zorunluluğunu doğurmuştur. Algoritma ve parametre seçiminde kullanılmak üzere geliştirilen "meta öğrenme teknikleri" birçok algoritmayı kombine ederek tahmin başarısını arttırmak amacıyla finansal hile tespitinde kullanılmaya başlanmıştır (Chan vd., 1999: 67-74). Bunun yanı sıra, usulüne uygun olarak gerçekleştirilmiş bir işlemin hileli olarak raporlanması veya hileli işlemin yanlış etiketlenmesi, ortaya çıkarılamaması gibi problemler ciddi maliyet doğurduğu için, hile tespitinde kullanılan sayısal yaklaşımlar genişçe tartışılmıştır (Provost ve Fawcett, 2001: 203-210).

Şekil 1’de görüldüğü gibi, literatürde finansal suçların tespiti için kullanılan teknikler suçun konusu, çeşitliliği, boyutu ve operasyonel karakteristiğine göre farklılık göstermektedir. Örneğin, kredi kartı yoluyla yapılan hileli işlemlerin ve kara para aklama gibi bankacılık suçlarının tespitinde sınıflandırma ve kümeleme teknikleri kullanılırken, sigorta suçlarının tespitinde daha çok ağ modelleri ve görselleştirme teknikleri tercih edilmektedir.

Şekil 1. Finansal Hile Tespitinde Veri Madenciliği Tekniklerinin Kategorizasyonu
(Ngai vd., 2011: 563)



4. Çalışmaların Sınıflandırılması

Finansal kurumlar veya mudiler adına, bankalar veya diğer finansal kuruluşlar tarafından elde tutulan mal, para ve diğer varlıkların, illegal bir şekilde kullanılması veya edinilmesiyle haksız kazanç sağlanması / kaybın engellenmesi bankacılık suçu olarak nitelendirilmektedir. Aktif veya aktif olmayan kredi kartıyla yetkisiz olarak gerçekleştirilen işlemlerin konusu olan kredi kartı dolandırıcılığı ve gayri menkul hakkında gerçeğe aykırı ve yanlış beyanda bulunan veya beyanda bulunmayı ihmal eden borcu veren, satın alan veya kiralayan kişi / kişilerin eylemlerinin konusunu içeren mortgage dolandırıcılığı bankacılık suçları arasında incelenmektedir. Yasadışı bir faaliyet neticesinde elde edilen ekonomik menfaatlerin bütünü olarak adlandırılan kara paranın karmaşık işlemlere tabi tutularak aklanması suçu ise diğer bankacılık suçları arasında yer almaktadır (Cornell School of Law, 2009). Bankacılık suçlarının tespitinde veri madenciliği tekniklerini kullanan bazı çalışmalar sınıflandırma, kümeleme, aykırı değer tespiti, birliktelik kuralları, regresyon ve görselleştirme olarak kategorize edilerek Tablo 2'deki gibi özetlenebilir:

Tablo 2. Bankacılık Suçları

Referans Çalışma	İşlem Türü	Veri Madenciliği Modeli	Kullanılan Teknik
Ghosh ve Reilly (1994), Aleskerov Chan vd. (1999)	Kredi Kartı Dolandırıcılığı	Sınıflandırma	Yapay Sinir Ağları
	Kredi Kartı Dolandırıcılığı	Sınıflandırma	Karar Ağaçları, CART, Bayes Ağları
Wheeler ve Aitken (2000)	Kredi Kartı Dolandırıcılığı	Sınıflandırma	Vaka Bazlı Çözümleme
Quah ve Sriganesh (2008)	Kredi Kartı Dolandırıcılığı	Kümeleme	Kendini Düzenleyen Haritalar
Whitrow vd. (2009)	Kredi Kartı Dolandırıcılığı	Sınıflandırma	Destek Vektör Makinaları, Rastgele Orman Algoritması
Brabazon vd. (2010)	Kredi Kartı Dolandırıcılığı	Kümeleme	Yapay Bağışıklık Sistemi Algoritması
Duman ve Ozelik (2011)	Kredi Kartı Dolandırıcılığı	Sınıflandırma	Genetik Algoritmalar
Bhattacharyya vd. (2012)	Kredi Kartı Dolandırıcılığı	Regresyon	Lojistik Regresyon
Jha vd. (2012)	Kredi Kartı Dolandırıcılığı	Regresyon	Multinomial Lojit Modeller
Xiong vd. (2013)	Kredi Kartı Dolandırıcılığı	Kümeleme	k-means Kümeleme, Dizi Madenciliği
Duman ve Elikucuk (2013)	Kredi Kartı Dolandırıcılığı	Sınıflandırma	Genetik Algoritmalar, Göçmen Kuşlar Optimizasyonu Algoritması
Olsewski (2014)	Kredi Kartı Dolandırıcılığı	Sınıflandırma, Görselleştirme	Özdüzenleyici Haritalar Yöntemi, Veri Görselleştirme
Zhang vd. (2003)	Kara Para Aklama	Regresyon	Korelasyon Analizi Tabanlı Link Keşfi
Gao vd. (2007)	Kara Para Aklama	Sınıflandırma	Ağ Analizi
Liu vd. (2008)	Kara Para Aklama	Kümeleme	Dizi Madenciliği, Öklid Benzerlik Uzaklığı
Gao (2009)	Kara Para Aklama	Kümeleme, Aykırı Değer Tespiti	Küme Tabanlı Lokal Aykırı Faktör Algoritması
Raza ve Haider (2011)	Kara Para Aklama	Sınıflandırma, Aykırı Değer Tespiti	Bayes Ağ Analizi, Entropi Tabanlı Aykırı İndeks Algoritması
Drezewski vd. (2012)	Kara Para Aklama	Kümeleme	Örüntü Madenciliği, BIDE Algoritması
Cao ve Do (2012)	Kara Para Aklama	Kümeleme	CLOPE Algoritması
Didimo vd. (2014)	Kara Para Aklama	Veri Görselleştirme, Kümeleme	Sosyal Ağ Analizi, VisFAN algoritması
Shunnar ve Barry (2011)	Mortgage Dolandırıcılığı	Sınıflandırma, Aykırı Değer Tespiti	Kuadratik Diskriminant Analizi

Belirli bir bedel karşılığında, tahminlenen riskin gerçekleşmesine bağlı olarak ortaya çıkan kaybın telafisini konu alan sözleşme olarak nitelendirilen sigorta kavramı, karşılanması mümkün olmayan zararların tazmininde kullanılan en önemli ekonomik araç olarak görülmektedir. Doğru olmayan veya aldatıcı, eksik bilgi

verilerek tarafların hak ve borçlarını, sigortanın konusu, bedeli veya süresini etkileyebilecek nitelikteki kanıtı değiştirmek, gizlemek veya yok etmek yoluyla sigortalı ve sigortacı arasındaki sözleşmenin çerçevesi dışında haksız kazanç sağlayan her türlü eylemler sigorta suistimalleri kapsamına girmektedir. Başvuru, uygunluk, derecelendirme ve kanıtlandırma süreçlerinde gerçekleştirilen hileli işlemlerin kapsamı tarım, otomotiv, sağlık ve konut sektörü gibi geniş bir yelpazede incelenmektedir. Sigorta suçlarının tespitinde literatürde kullanılan veri madenciliği teknikleri Tablo 3'deki gibi sınıflandırılabilir:

Tablo 3. Sigorta Suçları

Referans Çalışma	İşlem Türü	Veri Madenciliği Modeli	Kullanılan Teknik
Rejesus vd. (2004)	Tarım Sigortası Dolandırıcılığı	Regresyon	Log-Doğrusal Model
Jin vd. (2005)	Tarım Sigortası Dolandırıcılığı	Regresyon	Lojistik Model, Probit Model
Atwood vd. (2006)	Tarım Sigortası Dolandırıcılığı	Regresyon	Değişim Modelleri
Stefano ve Gisella	Araç Sigortası Dolandırıcılığı	Sınıflandırma	Bulanık Mantık Modelleri
Viane vd. (2002), Viane vd. (2004)	Araç Sigortası Dolandırıcılığı	Sınıflandırma	Bayes Yapay Sinir Ağları, Naive Bayes, K- En yakın Komşu, Karar Ağaçları
Brockett vd. (1998; 2002)	Araç Sigortası Dolandırıcılığı	Sınıflandırma	Temel Bileşenler Analizi, RIDIT, Bayes İnanç Ağları
Artis vd. (1999)	Araç Sigortası Dolandırıcılığı	Sınıflandırma	Kesikli Seçim Modelleri
Pathak vd. (2005)	Araç Sigortası Dolandırıcılığı	Sınıflandırma	Bulanık Mantık
Bermudez vd. (2008)	Araç Sigortası Dolandırıcılığı	Sınıflandırma	Bayes Modelleri
Subelj vd. (2011)	Araç Sigortası Dolandırıcılığı	Sınıflandırma, Veri Görselleştirme	Sosyal Ağ Analizi, Link Analizi, Yapay Sinir Ağları
Cox (1995)	Sağlık Sigortası Dolandırıcılığı	Sınıflandırma	Bulanık Mantık
Sokol vd. (2001)	Sağlık Sigortası Dolandırıcılığı	Veri Görselleştirme	Veri Görselleştirme
Major ve Riedinger (2002)	Sağlık Sigortası Dolandırıcılığı	Sınıflandırma	Polimorf Mantık
Yamanashi vd. (2004)	Sağlık Sigortası Dolandırıcılığı	Kümeleme	Sonlu Karma Modeller
Yang ve Hwang (2006)	Sağlık Sigortası Dolandırıcılığı	Sınıflandırma	Birliktelik Kuralları
Musal (2010)	Sağlık Sigortası Dolandırıcılığı	Kümeleme	Uzaklık Analizi
Kirlidog ve Asuk (2012)	Sağlık Sigortası Dolandırıcılığı	Sınıflandırma	Destek Vektör Makineleri
Shin vd. (2012)	Sağlık Sigortası Dolandırıcılığı	Aykırı Değer Analizi	Kompozit Derece Analizi
Thornton vd. (2013)	Sağlık Sigortası Dolandırıcılığı	Sınıflandırma	Çok Değişkenli Şema Analizi

Sermaye veya mal piyasalarında hileli ve yanıltıcı faaliyetler kapsamına giren manipülasyon, içeriden öğrenenler ticareti suçları, fiyat oluşumuna etki eden

ayrıcılık bilginin kullanımı ve çeşitli yöntemlerle yapay fiyat oluşturulması temeline dayanmaktadır. Manipülasyonda, bilinçli olarak yanlış yönlendirici faaliyetlerle menkul kıymet fiyatlarının yükseltilmesi, düşürülmesi veya stabilize edilmesi amaçlanırken, içeriden öğrenenler ticaretinde ise fiyatlara etki edebilecek ayrıcalıklı bilginin kullanılarak anormal kazanç elde edilmesi veya olası zararlardan sakınılması amaçlanmaktadır. Önceden hareket etme (front running), ponzi şemaları, açığa satış istismları ve hedge fonu suçları diğer kıymetli evrak suçları kategorisinde incelenmektedir. Kurumsal denetimi gerektiren şirket içi finansal bilgilerin suistimali, içeriden öğrenenlerin kendi aralarındaki bilgi paylaşımları ve herhangi finansal bir suçun örtbas edilmesi gibi suç oluşturan vergi kaçakçılığı ve yolsuzluklar da diğer suçlar kategorisinde incelenmektedir (Ngai vd., 2011: 562).

Tablo 4. Menkul Kıymet Suçları

Referans Çalışma	İşlem Türü	Veri Madenciliği Modeli	Kullanılan Teknik
Böhme ve Holz (2006)	Manipülasyon	Regresyon	Çok Değişkenli Regresyon
Öğüt vd. (2009)	Manipülasyon	Sınıflandırma, Regresyon	Yapay Sinir Ağları, Destek Vektör Makinaları, Diskriminant Analizi, Regresyon Analizi
Mongkolnavin ve Tirapat (2009)	Manipülasyon	Sınıflandırma	Birliklilik Kuralları
Zaki vd. (2010)	Manipülasyon	Sınıflandırma	C5.0 Algoritması
Diaz vd. (2011)	Manipülasyon	Sınıflandırma, Regresyon	Doğrusal ve Lojistik Regresyon Modelleri, C5.0 Algoritması, k-En yakın Komşu Analizi, Akran Grupları Analizi
Kim ve Shon (2012)	Manipülasyon	Sınıflandırma	Birliklilik Kuralları
Kirkland vd. (1998)	İçeriden Öğrenenler Ticareti	Sınıflandırma	Yapay Sinir Ağları
Safer (2002)	İçeriden Öğrenenler Ticareti	Sınıflandırma	Çok Değişkenli Regresyon Modelleri, Bulanık Mantık, Kural Tabanlı Algoritmalar, Metin Madenciliği
Goldberg vd. (2003)	İçeriden Öğrenenler Ticareti	Regresyon, Kümeleme	Lojistik Regresyon, Yapay Sinir Ağları, C4.5 Algoritması
Donoho (2004)	İçeriden Öğrenenler Ticareti	Sınıflandırma	Sosyal Ağ Analizi, Mantıksal Modelleme
Gupta ve Hossain (2011)	İçeriden Öğrenenler Ticareti	Kümeleme	Ağ Analizi, Zaman Serileri
Tamersoy vd. (2013)	İçeriden Öğrenenler Ticareti	Kümeleme	

Şirketlerin finansal tablolarında değişikliğe yol açacak şekilde kayıtların tahrif edilmesi, makyajlanması, yanlış yansıtılması ve raporlama prosedürünün kasten yanlış uygulanması şeklinde ortaya çıkan hileli finansal raporlamalar ve vergi mükelleflerinin ödemekle yükümlü oldukları vergilerin ilgili kurumlara yanlış beyan edilerek veya gizlenerek ödenmemesi durumu olarak adlandırılan “vergi kaçakçılığı”, diğer ilişkili finansal suçlar arasında incelenmektedir. Ayrıca, tüzel veya

gerçek kişilerin yetkilerini, bulundukları statünün avantajlarını kullanarak özel çıkar sağlama amacıyla kullanması olarak nitelendirilen şirket yolsuzlukları da diğer ilişkili finansal suçlar arasında incelenmektedir. Söz konusu finansal suçların tespitinde literatürde kullanılan veri madenciliği teknikleri Tablo 5’deki gibi özetlenebilir:

Tablo 5. Diğer İlişkili Finansal Suçlar

Referans Çalışma	İşlem Türü	Veri Madenciliği Modeli	Kullanılan Teknik
Beasley (1996)	Hileli Finansal Raporlama	Regresyon	Lojit Regresyon
Green ve Choi (1997)	Hileli Finansal Raporlama	Sınıflandırma	Yapay Sinir Ağları
Bell vd. (2000)	Hileli Finansal Raporlama	Regresyon	Lojistik Regresyon
Bai vd. (2008)	Hileli Finansal Raporlama	Sınıflandırma ve Regresyon	CART ve Regresyon Ağaçları
Spathis vd. (2002)	Hileli Finansal Raporlama	Sınıflandırma, Regresyon	Diskriminant Analizi, Lojit Modeller
Kotsiantis vd. (2006)	Hileli Finansal Raporlama	Sınıflandırma	Stacking Variant Metodolojisi, C4.5 Algoritması, Bayes Ağları
Kirkos vd. (2007)	Hileli Finansal Raporlama	Sınıflandırma	Karar Ağaçları, Bayes İnanç Ağları, Sinir Ağları
Glancy ve Yadav (2011)	Hileli Finansal Raporlama	Kümeleme	Hiyerarşik Kümeleme
Ravisankar vd. (2011)	Hileli Finansal Raporlama	Sınıflandırma, Regresyon	Genetik Algoritmalar, Lojistik Regresyon, Destek Vektör Makinaları, Çok Katmanlı İleri Beslemeli Sinir Ağları
Chen (2011)	Hileli Finansal Raporlama	Sınıflandırma, Regresyon	C5.0, CART, CHAID, Karar Ağaçları, Lojistik Regresyon
Huang vd. (2014)	Hileli Finansal Raporlama	Kümeleme	Gelişen Hiyerarşik Özdüzenleyici Haritalar
Wu vd. (2012)	Vergi Kaçakçılığı	Sınıflandırma	Birliktelik Kuralları
Gonzalez ve Velasquez (2013)	Vergi Kaçakçılığı	Kümeleme	Özdüzenleyici Haritalar
Deshmukh vd. (1997)	Şirket Yolsuzluğu	Sınıflandırma	Bulanık Mantık
Eining vd. (1997)	Şirket Yolsuzluğu	Regresyon	Lojistik Regresyon
Welch vd. (1998)	Şirket Yolsuzluğu	Sınıflandırma	Evrimsel Algoritmalar
Koskivaara (2004)	Şirket Yolsuzluğu	Sınıflandırma	Yapay Sinir Ağları
Holton (2009)	Şirket Yolsuzluğu	Kümeleme	Naive Bayes

Türkiye’de finansal suçların veri madenciliği ile tespiti konusunda yayınlanmış çok az sayıda bilimsel çalışma tespit edilmiştir. Bu çalışmaların birkaçı hilenin gerçekleşmeden önce bir erken uyarı sistemi ile tespit edilmesi ve gerçekleşikten sonra tespit edilmesi konularına odaklanırken, diğer kısmı muhasebe hilelerinin

tespiti için prosedür geliştirilmesi ve konu hakkında Avrupa Birliği – ABD düzenlemelerini ele almaktadır. Finansal suçların veri madenciliği teknikleriyle değerlendirildiği tespit edilen çalışmalar Tablo 6’daki gibi özetlenebilir:

Tablo 6. Finansal Suçların Tespitinde Veri Madenciliği : Türkiye Uygulamaları

Referans Çalışma	İşlem Türü	Veri Madenciliği Modeli	Kullanılan Teknik
Küçükocaoğlu v.d. (1997)	Manipülasyon	Regresyon	Lojistik Regresyon Analizi
Koyuncuğil ve Özgülbaş (2007)	Menkul Kıymet Suçları	Sınıflandırma	CHAID
Ata ve Seyrek (2009)	Hileli Finansal	Sınıflandırma	Karar Ağaçları, YSA
Kırloğlu ve Ceyhan	Hileli Finansal	Sınıflandırma	K-En Yakın Komşu Algoritması
Karaatlı v.d. (2014)	Hileli Finansal	Sınıflandırma	Yapay Sinir Ağları
Terzi ve Şen (2015)	Hileli Finansal	Sınıflandırma	Yapay Sinir Ağları

5. Sonuç

Bu çalışmada, finansal hilelerin tespiti ve önlenmesinde kullanılan veri madenciliği yaklaşımları esas alınarak, geniş bir literatür taraması yapılmış ve her bir çalışmada kullanılan veri madenciliği modeli ve tekniği sınıflandırılarak finansal suç kategorisine göre tablolar halinde verilmiştir. Bu açıdan çalışmanın finansal suçların tespitine ilişkin araştırmalar açısından alternatif bir kaynak olacağı düşünülmektedir. Çalışmada gerçekleştirilen literatür araştırmasının sonuçları şu şekilde özetlemek mümkündür:

1. Finansal suçların tespitinde veri madenciliği tekniklerinin 79 çalışma içerisinde 22 çalışma ile (%28) en çok bankacılık suçlarının tespitinde kullanıldığı ve bu suçlar içerisinde en çok kredi kartı suçları tespitinin olduğu (%62) tespit edilmiştir. Bankacılık suçları arasında en az çalışmanın mortgage dolandırıcılığı konusunda yapıldığı göze çarpmaktadır. Ayrıca, bankacılık suçlarının tespitinin, en çok (%62) sınıflandırma teknikleri kullanılarak gerçekleştirildiği görülmektedir. Kredi kartı dolandırıcılığının tespitinde yoğun olarak sınıflandırma teknikleri kullanılırken, kara para aklama suçunun tespitinde ise kümeleme tekniklerinin hakim olarak kullanıldığı gözlenmektedir. Sınıflandırma tekniklerinin “verinin önceden tanımlanmış sınıflara atanması” amacı göz önünde bulundurulduğunda, gerçekleştirilen çalışmaların birçoğunda önceden tespit edilmiş hileli işlemler verisinin kullanıldığı anlaşılmaktadır. Hileli işlemler verisinin bulunmadığı durumlarda ise, işlemleri benzer özelliklerine göre ayırmak için kümeleme yaklaşımları kullanılmaktadır.

2. Sigorta suçlarının veri madenciliği ile tespitinde, sigorta kayıtlarında hileye işaret edebilecek değişkenlerin seçimi ve söz konusu işlemlerin skorlanarak normal işlemlerden farklılaşıp farklılaşmadığı konularına odaklanılmaktadır. Buna göre, özellikle sigorta sağlayıcısı konumunda bulunan kurum ve kuruluşların keşifsel davranışlarını ölçmede kullanılabilecek ölçeğin oluşturulması ve sosyo-demografik ölçütler göz önünde bulundurularak sigorta harcamaları içerisindeki anormal trendlerin keşfedilmesi gerçekleştirilen çalışmalarca amaçlanmaktadır.
3. Menkul kıymet suçları açısından bakıldığında çalışmalarda, sermaye piyasalarında menkul kıymet fiyatları üzerinde etkili olabilecek içerideki bilgi, spam mesajlar, piyasa açıklamaları ve bildirimlerin etkilerinin tespit edilmesi ve tahminlenmesi amaçlanmaktadır. Özellikle içeriden öğrenenler ticareti çalışmalarında, gün içi işlem trafiği yakın takibe alınarak, şirketlerin kamuyu aydınlatma beyanları mercek altına alınmış ve ileriye/geriye yönelik menkul kıymet fiyatları, hacmi ve bazı işlem karakteristikleri bazında değerlendirilmeler yapılmıştır. Bu çalışmaların bazılarında organizasyon içerisinde yönetici ve çalışan statü gruplarının işlemlerinin hisse senedi fiyatlarına etkisiyle birlikte, sosyal ağ içerisindeki ilişkileri modellenerek değerlendirilmiştir. Fakat bununla birlikte, aykırı değer analiziyle menkul kıymet suçlarının tespiti üzerine herhangi bir çalışmaya rastlanmamıştır.
4. Hileli finansal tablo raporlamalarına etki eden faktörlerin (iç kontrol yapısı, maaş oranları, agresif yönetim politikası, denetim komitesi etkinliği vb.) ve hile indikatörü olan değişkenlerin (karlılık, likidite, etkinlik oranları vb.) belirlenmesi amacıyla kullanılan veri madenciliği tekniklerinin, vergi kaçakçılığının tespiti konusunda sınırlı olan literatürü göze çarpmaktadır. Bu doğrultuda, ilgili alana ilişkin problemin veri madenciliği süreci açısından değerlendirilip, analiz edilmesi tavsiye edilebilir. Ayrıca, probleme özgü çalışılan veri seti içerisindeki anormal örüntülerin tespiti için klasik sayısal yöntemlerin tespit edemediği, “anormal gözlemin tespitinin başka bir gözlemce engellenmesi” veya “normal bir gözlemin aykırı bir gözlem olarak tespit edilmesi” gibi problemlerin önüne geçebilmek için dayanıklı ve hibrit sayısal yaklaşımların uygulanması tavsiye edilmektedir.

Finansal suçların tespitinde veri madenciliği tekniklerinin kullanımı konusunda Türkiye uygulamaları incelendiğinde, yapılan çalışma sayısının çok az olduğu tespit edilmiştir. Söz konusu çalışmaların mali tablo denetimi konusuna odaklandığı; vergi

kaçakçılığı, şirket yolsuzluğu gibi konularda çalışmanın olmadığı belirlenmiştir. Ayrıca, sigortacılık ve bankacılık suçları konusunda da hiçbir çalışmaya rastlanmamıştır. Konunun önemi göz önünde bulundurulduğunda, çalışmaların yürütülebilmesi için gerekli veri paylaşımının ilgili kurumlar tarafından araştırmacılara sunulması ve çalışmaların makine öğrenmesi, istatistik gibi kantitatif teknikler kullanılarak multidisipliner yürütülmesi önerilmektedir.

Kaynaklar

ESKİŞEHİR OSMANGAZİ ÜNİVERSİTESİ İİBF DERGİSİ

ACFE (2012), "ACFE Report to the nations on occupational fraud and abuse, Technical report- Global fraud survey 2012", <http://www.acfe.com> (Erişim: 01.06.2014).

Aleskerov, E., Freisleben, B., Rao, B. (1997), "CARDWATCH: A Neural Network-Based Database Mining System for Credit Card Fraud Detection", Proc. of the IEEE/IAFE on Computational Intelligence for Financial Engineering, 220-226.

Artis, M., Ayuso, M., Guillen, M. (1999), "Modelling different types of automobile insurance fraud behaviour in the Spanish market", Insurance: Mathematics and Economics, 24, 67-81.

Ata, H. Ali, Seyrek, İbrahim H. (2009), "The Use Of Data Mining Techniques In Detecting Fraudulent Financial Statements: An Application on Manufacturing Firms", Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 14(2), 157-170.

Atwood, J.A., Robison, J.F., Shaik, S. (2006), "Estimating the prevalence and cost of yield-switching fraud in the federal crop insurance program", American Journal of Agricultural Economics, 88(2), 365-381.

B. Bell, Timothy, V. Carcello, J. (2000), "A Decision Aid for Assessing the Likelihood of Fraudulent Financial Reporting", Auditing: A Journal of Practice & Theory, 19(1), 169-184.

Bai, B. J., Yang, Yen, X. (2008), "False financial statements: Characteristics of China's listed companies and cart detecting approach", International Journal of Information Technology & Decision Making, 7, 339-359.

Beasley, Mark S. (1996), "An Empirical Analysis of the Relation between the Board of Director Composition and Financial Statement Fraud", The Accounting Review, 71(4), 443-465.

Benford, F. (1938), "The law of anomalous numbers", Proceedings of the American Philosophical Society, 78 (4), 551-572.

Bermudez, Ll., Perez, J.M., Ayuso, M. Gomez, E., Vazquez, F.J. (2008), "A Bayesian dichotomous model with asymmetric link for fraud in insurance", Insurance: Mathematics and Economics, 42(2), 779-786.

Bhattacharyya, S., Jha, S., Tharakunnel, K., Christopher Westland, J. (2011), "Data mining for credit card fraud: A comparative study", [Decision Support Systems](#), 50(3), 602-613.

Bozdoğan, H. (2004), Statistical Data Mining and Knowledge Discovery, Washington: A CRC Press Company.

Böhme, Rainer, Holz, Thorsten (2006), "The Effect of Stock Spam on Financial Markets", <http://ssrn.com/abstract=897431>, 04.05.2014 (Erişim: 12.05.2015).

Brabazon, A., Cahill, J., Keenan, P., Walsh, D. (2010), "Identifying online credit card fraud using Artificial Immune Systems", IEEE Congress on Evolutionary Computation, IEEE, 1-7.

Brockett, P. L., R. A. Derrig, Xia, X. (1998), "Using Kohonen's Self-Organizing Feature Map to Uncover Automobile Bodily Injury Claims Fraud", Journal of Risk and Insurance, 65, 245-274.

Brockett, P. L., Derrig, R.A., Golden, L.L., Levine, Alpert, A. M. (2002), "Fraud Classification Using Principal Component Analysis of RIDITs", Journal of Risk and Insurance, Vol. 69, Issue 3, s. 341-372.

Cahill, M., Chen, F., Lambert, D., Pinheiro, J., Sun, D. (2002), "Detecting Fraud in the Real World," Handbook of Massive Datasets, 911-930.

Chan, P. K., Fan, W., Prodromidis, A. L., Stolfo, S. J. (1999), "Distributed data mining in credit card fraud detection", IEEE Intelligent Systems, 14(6), 67–74.

Chan, P. K., Fan, W., Prodromidis, A. L., Stolfo, S. J. (1999), "Distributed data mining in credit" .

Chen, M.Y. (2011), "Predicting corporate financial distress based on integration of decision tree classification and logistic regression", Expert Systems with Applications, 38(9), 11261-11272.

Cox, E. (1995), "A Fuzzy System for Detecting Anomalous Behaviors in Healthcare Provider Claims", Intelligent Systems for Finance and Business, 111-134.

Deshmukh, A., Romine, J., Siegel, P.H. (1997), "Measurement and combination of red flags to assess the risk of management fraud: a fuzzy set approach", Managerial Finance, 23(6), 35-48.

Díaz, D., Theodoulidis, B., Sampaio, P. (2011), "Analysis of stock market manipulations using knowledge discovery techniques applied to intraday trade prices", Expert Systems with Applications, 38, 12757–12771.

Didimo ,Walter, Liotta, Giuseppe, Montecchiani, Fabrizio (2014), "Network visualization for financial crime detection", Journal of Visual Languages & Computing, 25(4), 433–451.

D.T. Larouse (2004), *Discovering Knowledge in Data: An Introduction to Data Mining*, John Wiley & Sons .

Don, M., Casey, C. (1998), *Building a Better Data Warehouse*, USA: Prentice Hall

Donoho, S. (2004), "Early detection of insider trading in option markets", In *Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining*, 420-429.

Drezewski, Rafal, Sepielak, Jan, Filipkowski, Wojciech (2012), "System supporting money laundering detection", *Digital Investigation*, 9(1), 8–21

Duman, Ekrem, Eliküçük, İlker (2013), "Applying Migrating Birds Optimization to Credit Card Fraud Detection", *Trends and Applications in Knowledge Discovery and Data Mining Lecture Notes in Computer Science*, 7867, 416-427.

Duman, E., Ozelik, M.H. (2011), "Detecting credit card fraud by genetic algorithm and scatter search", *Expert Systems with Applications*, 38, 3057-13063.

Durtschi, C., Hillison, W., Pacini, C. (2004), "The Effective Use of Benford's Law to Assist in detecting fraud in accounting data", *Journal of Forensic Accounting*, 5, 17-34.

E. Turban, J.E. Aronson, T.P. Liang, R. Sharda (2010), *Decision Support and Business Intelligence Systems*, Prentice Hall.

Eining, M., Jones, D.R., Loebbecke, J.K. (1997), "Reliance on decision aids: an examination of auditors' assessment of management fraud," *Auditing: A Journal of Practice & Theory*, 16(2), 1–19.

EMC² (2014), "EMC² Data Report", <http://www.emc.com>, (Erişim: 21.04.2014).

EMC² (2014), "The digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East", <http://www.emc.com>, (Erişim: 16.03.2015).

European Commission (2015), "Financial Crime", (Çevrimiçi) http://ec.europa.eu/internal_market/company/financial-crime, (Erişim:04.04.2015).

Federal Bureau of Investigation (2011), "Financial Crimes Report to the Public of 2011", <http://www.fbi.gov/stats-services/publications>, (Erişim: 04.04.2014).

Federal Reserve Bank (2013), "The 2013 Federal Reserve Payments Study", <http://www.frbervices.org>, (Erişim: 11.05.2015).

Federal Bureau of Investigation New York Divisio (2014), "White-Collar Crime", <http://newyork.fbi.gov/dojpressrel/pressrel08/nyfo121108.htm> (Eriřim: 22.04.2014).

Financial Conduct Authority (2013), "The changing face of financial crime: Report of 2013", <http://www.fca.org.uk> (Eriřim: 03.05.2015).

Fraser, I. A. M., Hatherly, D. J., Lin, K. Z. (1997) "An empirical investigation of the use of analytical review by external auditors", *The British Accounting Review*, 29(1), 35–47.

Gao Z. (2009), "Application of cluster-based local outlier factor algorithm in anti-money laundering", *Conference on Management and Service Science*, 1-4.

Gao, Zengan, Ye, Mo (2007), "A framework for data mining based anti-money laundering research", *Journal of Money Laundering Control*, 10(2), 170 – 179.

Ghosh, S., Reilly, D.L. (1994), "Credit Card Fraud Detection with a Neural-Network," *Proceedings on 27th Hawaii International Conference on System Sciences: Information Systems: Decision Support and Knowledge-Based Systems*, 3, 621-630.

Giudici, P. (2003), *Applied Data Mining Statistical Methods for Business and Industry*, Wiley.

Glancy, F. H., Yadav, S. B. (2011), "A computational model for financial reporting fraud detection: On quantitative methods for detection of financial fraud", *Decision Support Systems*, 50(3), 595–601.

Goldberg, H. (2003), "The NASD Securities Observation, News Analysis & Regulation System (SONAR)", *Conference Proceedings of Innovative Applications of Artificial Intelligence*, 11-18.

González, P.C., Velásquez, J.D. (2013), "Characterization and detection of taxpayers with false invoices using data mining techniques", *Expert Systems with Applications*, 40(5), 1427–1436.

Green, B. P., Choi, J. H. (1997), "Assessing the risk of management fraud through neural network technology", *Auditing: A Journal of Practice & Theory*, 16, 14–28.

Gupta, S., Hossain, L. (2011), "Towards near-real-time detection of insider trading behaviour through social networks", *Computer Fraud & Security*, 2011(1), 7-16.

Han, J., Kamber, M. (2001), *Data Mining: Concepts and Techniques*, USA: Morgan Kaufman Publishers, Academic Press.

Hand, D.J. (1992), *Discrimination and Classification*, Chichester: Wiley.

Hand, D.J. (1997), *Construction and Assessment of Classification Rules*, Chichester: Wiley.

Holton, C. (2009), "Identifying disgruntled employee systems fraud risk through text mining: a simple solution for a multi-billion dollar problem", *Decision Support Systems*, 46(4), 853–864.

Huang, S.Y., Tsaih, R.H., Yu, F. (2014), "Topological pattern discovery and feature extraction for fraudulent financial reporting", *Expert Systems with Applications*, 41(9), 4360–4372.

Insurance Europe (2013), "The Impact of Insurance Fraud", <http://www.insuranceeurope.eu/> (Erişim: 19.03.2015).

Jha, Sanjeev, Guillen, Montserrat, Westland, J.C. (2012), "Employing transaction aggregation strategy to detect credit card fraud", *Expert Systems with Applications*, 39(16), 12650–12657.

Jin, Yufei, R.M., Roderick, Little, B.B. (2005), "Binary choice models for rare events data: a crop insurance fraud application", *Applied Economics*, 37(7), 841-848.

Karaatlı, M., Senal, S., Öztürk, M.S. (2014), "Denetim Planlaması Aşamasında Analitik İnceleme Tekniği olarak Yapay Sinir Ağları Kullanımı: Bir Firma Uygulaması", *Ege Academic Review*, 14(4), 637-648.

Khoa Cao, Dang, Do, Phuc. (2012), "Applying Data Mining in Money Laundering Detection for the Vietnamese Banking Industry", *Intelligent Information and Database Systems Lecture Notes in Computer Science*, 7197, 207-216.

Kırlioğlu, H., Ceyhan İ. (2014), "Mali Tablo Denetiminde Ön Analitik İnceleme Tekniği olarak Veri Madenciliğinin Kullanımı: Borsa İstanbul Uygulaması", *Akademik Yaklaşımlar Dergisi*, 5(1), 13-36.

Kirkland, J.D., Senator, T.E., Hayden, J.J., Dybala, T., Goldberg, H.G., Shyr, P. (1998), "The NASD Regulation Advanced Detection System (ADS)", *Association for the Advancement of Artificial Intelligence*, 20(1), 55-67.

Kirkos, E., Spathis, C., Manolopoulos, Y. (2007), "Data Mining Techniques for the detection of fraudulent financial statements", *Expert Systems with Applications*, 32(4), 995-1003.

Kirlidog, Melih, Asuk, Cuneyt (2012), "A Fraud Detection Approach with Data Mining in Health Insurance", *Procedia - Social and Behavioral Sciences*, 62, 989–994.

- Koskivaara, E. (2004), "Artificial neural networks in analytical review procedures", *Managerial Auditing Journal*, 19(2), 191 – 223.
- Kotsiantis, S., Koumanakos, E. D., Tampakas, Tzelepis, V. (2006), "Forecasting Fraudulent Financial Statements Using Data Mining", *International Journal of Computational Intelligence*, 2, 104-110.
- Koyuncugil, S., Özgülbaş, N. (2007), "Detecting Financial Early Warning Signs In Istanbul Stock Exchange by Data Mining", *International Journal of Business Research*, 7(3).
- Küçükkoçaoğlu, G., Benli, Y.K., Küçüksözen, C. (1997), "Finansal Bilgi Manipülasyonunun Tespitinde Yapay Sinir Ağı Modelinin Kullanımı", *İMKB Dergisi*, 9 (36), 1-30.
- Lisboa, P., Edisbury, B., Vellido, A. (2000), *Business Applications of Neural Networks: The State-Of-The-Art of Real-World Applications*, World Scientific Pub Co.Inc.
- Liu, X., Zhang, P., Zeng, D. (2008), "Sequence Matching for Suspicious Activity Detection in Anti-Money Laundering", *Intelligence and Security Informatics Lecture Notes in Computer Science*, 5075, 50-61.
- Major, J. A., Riedinger, D. R. (2002), "EFD: A Hybrid Knowledge / Statistical Based System for the Detection of Fraud", *Journal of Risk and Insurance*, 69(3), 309-324.
- McLachlan, G.J. (1992), *Discriminant Analysis and Statistical Pattern Recognition*, New York: Wiley.
- Mongkolnavin, J., Tirapat, S. (2009), "Marking the Close analysis in Thai Bond Market Surveillance using association rules", *Expert Systems with Applications*, 36(4), 8523- 8527.
- Musal, R. M. (2010), "Two models to investigate Medicare fraud within unsupervised databases", *Expert Systems with Applications*, 37(12), 8628-8633.
- Nigrini, M.J. (1999), "I've got your number", *Journal of Accountancy*, 187(5), 79–83.
- NYSE (2014), "NYSE Market Data", <http://www.nyxdata.com/Data-Products/Product-Summaries>, (Erişim: 15.05.2014).
- Olszewski, Dominik (2014), "Fraud detection using self-organizing map visualizing the user profiles", *Knowledge-Based Systems*, 70, 324–334.
- Ögüt, H., Aktaş, R., Alp, A., Doğanay, M.M. (2009), "Prediction of financial information manipulation by using support vector machine and probabilistic neural network", *Expert Systems with Applications*, 36(3), 5419-5423.

Pathak, J., Vidyarthi, N., Summers, S.L. (2005), "A fuzzy-based algorithm for auditors to detect elements of fraud in settled insurance claims", *Managerial Auditing Journal*, 20(6), 632-644.

Phua, Clifton, Lee, Vincent, Smith, Kate, Gayler, R. (2014), "A Comprehensive Survey of Data Mining-based Fraud Detection Research", <http://www.arxiv.org>, (Eriřim: 18.02.2015).

PricewaterhouseCoopers (2014), "Global Economic Crime Survey of 2014", <http://www.pwc.com>, (Eriřim 23.06.2014).

PricewaterhouseCoopers (2014), "Global Economic Crime Survey of 2014", <http://www.pwc.com>, (Eriřim: 23.06.2014).

Provost, F., Fawcett, T. (2001), "Robust classification for imprecise environments", *Machine Learning*, 42, 2001, 203–210.

Raza, Saleha, Haider, Sajjad (2011), "Suspicious activity reporting using dynamic bayesian networks", *Procedia Computer Science*, 3, 987–991.

Quah T. S, Sriganesh, M. (2008), "Real-time credit card fraud using computational intelligence", *Expert Systems with Applications*, 35(4), 1721-1732.

Rao, C.A. (2005), *Handbook of Statistics*, Elsevier.

Ravisankar, P., Ravi, V., Raghava, G., Bose, I. (2011) "Detection of financial statement fraud and feature selection using data mining techniques", *Decision Support Systems*, 50(2), 491-500.

Rejesus, R.M., Little, B., Lovell, A. (2004), "Using Data Mining to Detect Crop Insurance Fraud: Is there a Role for Social Scientists?", *Journal of Financial Crime*, 12(1), 24-32.

Robert, Groth (2000), *Data Mining: Building Competitive Advantage*, USA: Prentice Hall PTR.

Safer, Alan M. (2002), The application of neural networks to predict abnormal stock returns using insider trading data, *Applied Stochastic Models in Business and Industry*, 18(4), 381-389.

Shin, H., Park, H., Lee, J., Jhee, W. C. (2012), "A scoring model to detect abusive billing patterns in health insurance claims", *Expert Systems with Applications*, 39(8), 7441-7450.

Shunnar, T., Barry, P.M., Shunnar, T. (2011), "Tracking Fraudulent Activities in Real Estate Transactions", *International Federations of Surveyors*, 1-15.

Singleton, T.W. (2010), *Fraud Auditing and Forensic Accounting* 4th edition, USA: John Wiley and Sons.

Sokol, L., Garcia, B., Rodriguez, J., West, M., Johnson, K. (2001), "Using data mining to find fraud in HCFA health care claims", *Topics in Health Information Management*, 22(1), 1-13.

Spathis, C., Doumpos, M., Zopounidis, C. (2002), "Detecting falsified financial statements: a comparative study using multicriteria analysis and multivariate statistical techniques", *The European Accounting Review*, 11(3), 509–535.

Stefano, B., Gisella, F. (2001), "Insurance Fraud Evaluation: A Fuzzy Expert System", *Proc. of IEEE International Fuzzy Systems Conference*, 1491-1494.

Subelj, L., Furlan, S., Bajec, M. (2011), "An expert system for detecting automobile insurance fraud using social network analysis", *Expert Systems with Applications*, 38(1), 1039-1052.

Tamersoy, A., Xie, B., Lenkey, S.L., Routledge, B.R., Chau, D.H., Navathe, S.B.(2013), "Insider Trading: Patterns & Discoveries from a Large Scale Exploratory Analysis", *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 797-804.

Terzi, S., Şen, İ.K. (2015), "Adli Muhasebede Hilelerin Tespitinde Yapay Sinir Ağı Modelinin Kullanımı", *Uluslararası İktisadi ve İdari İncelemeler Dergisi*, 14, 477-490.

The World Bank (2014), "Financial Indicators", <http://data.worldbank.org/indicator>, (Erişim: 13.05.2014).

Thornton, D., Mueller, R. M., Schoutsen, P., Van Hillegersberg, J. (2013), "Predicting Healthcare Fraud in Medicaid: A Multidimensional Data Model and Analysis Techniques for Fraud Detection", *Procedia Technology*, 9, 1252-1264.

Viaene, S., Derrig, R., Baesens, B., Dedene, G. (2002), "A Comparison of State-of-the-Art Classification Techniques for Expert Automobile Insurance Claim Fraud Detection", *Journal of Risk and Insurance*, 69(3), 373-421.

Viaene, S., Derrig, R., Dedene, G. (2004), "A Case Study of Applying Boosting Naive Bayes to Claim Fraud Diagnosis", *IEEE Transactions on Knowledge and Data Engineering*, 16(5), 612- 620.

Wallace, W.A. (2002), "Assessing the quality of data used for benchmarking and decision-making", *The Journal of Government Financial Management*, 51(3), 16-22.

Wang, S. (2010), "A Comprehensive Survey of Data Mining-Based Accounting-Fraud Detection Research", International Conference on Intelligent Computation Technology and Automation, 1, 50-53.

Wang Richard Y., B. Kon, Henry, Madnick, Stuart E. (1993), "Data Quality Requirements Analysis and Modeling", Ninth International Conference of Data Engineering Proceedings Book, 1.

Welch, J., Reeves, T.E., Welch, S.T. (1998), "Using a genetic algorithm-based classifier system for modeling auditor decision behavior in a fraud setting", International Journal of Intelligent Systems in Accounting, Finance & Management, 7(3), 173-186.

Wheeler, R., Aitken, S. (2000), "Multiple Algorithms for Fraud Detection", Knowledge-Based Systems, 13(3), 93-99.

Whitrow, C., Hand, D.J., Juszczak, P., Weston, D., Adams, N.M. (2009), "Transaction aggregation as a strategy for credit card fraud detection", Data Mining and Knowledge, 18(1), 30-55.

Wu, R.S., Ou, C.S., Lin, H.Y., Chang, S.I., Yen, D.C. (2012), "Using data mining technique to enhance tax evasion detection performance", Expert Systems with Applications, 39(10), 8769-8777.

Xiong, Tengke, Wang, Shengrui, Mayers, E. Andre, Monga (2013), "Personal bankruptcy prediction by mining credit card data", Expert Systems with Applications, 40(2), 665–676.

Yamanishi, K., Takeuchi, J., Williams, G., Milne, P. (2004), "On-Line Unsupervised Outlier Detection Using Finite Mixtures with Discounting Learning Algorithms", Data Mining and Knowledge Discovery, 8, 275-300.

Yang, W. S., Hwang, S. Y. (2006), "A process-mining framework for the detection of healthcare fraud and abuse", Expert Systems with Applications, 31(1), 56-68.

Yoonseong, Kim, So Young, Sohn (2012), "Stock fraud detection using peer group analysis", Expert Systems with Applications, 39(10), 8986–8992.

Zaki, Mohamed, Theodoulidis, Babis, Diaz, David (2010), "Stock-Touting Through Spam E-Mails: A Data Mining Case Study", Journal of Manufacturing Technology Management, 22(6), 770-787.

Zhang M., Hua, Jingzhou, Zhang, Ruofei, Salerno, John J., Philip S. Yu (2003), "Applying data mining in investigating money laundering crimes", Association for Computing Machinery, 747-752.