BAŞKENT ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI
MUHASEBE VE FİNANSMAN DOKTORA PROGRAMI

FİNANSAL İŞLEMLERDE GENETİK ALGORİTMA YÖNTEMİ İLE HİLE TAHMİNİ

HAZIRLAYAN

ELİF SENEM GÜDÜ

DOKTORA TEZİ

TEZ DANIŞMANI

DOÇ. DR. SONER GÖKTEN

ANKARA - 2022

BAŞKENT ÜNİVERSİTESİ

SOSYAL BİLİMLER ENSTİTÜSÜ

DOKTORA TEZ ÇALIŞMASI ORİJİNALLİK RAPORU

Tarih: 15/02/2022

Öğrencinin Adı, Soyadı: Elif Senem Güdü

Öğrencinin Numarası: 21610381

Anabilim Dalı: İşletme Anabilim Dalı

Programı: Muhasebe ve Finansman Doktora Programı

Danışmanın Unvanı/Adı, Soyadı: Doç. Dr. Soner Gökten

Tez Başlığı: Finansal İşlemlerde Genetik Algoritma ile Hile Tahmini

Yukarıda başlığı belirtilen Yüksek Lisans/Doktora tez çalışmamın; Giriş, Ana Bölümler ve Sonuç Bölümünden oluşan, toplam 89 sayfalık kısmına ilişkin, 15/ 02 / 2022 tarihinde tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı %10'dur. Uygulanan filtrelemeler:

1. Kaynakça hariç

2. Alıntılar hariç

3. Beş (5) kelimeden daha az örtüşme içeren metin kısımları hariç

"Başkent Üniversitesi Enstitüleri Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Usul ve Esaslarını" inceledim ve bu uygulama esaslarında belirtilen azami benzerlik oranlarına tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Öğrenci İmzası:…….

ONAY

Tarih: 15/02/2022

Öğrenci Danışmanı Unvan, Ad, Soyad, İmza:

Doç. Dr. Soner Gökten

……

TEŞEKKÜR

Tezimin yazılması sürecinde çeşitli kişilerin yardım ve desteklerini aldım. İlk olarak tez danışmanım Doç. Dr Soner Gökten'e tezimin yazılması ve tamamlanması sürecinde verdiği rehberlik ve destekten dolayı sonsuz teşekkürlerimi sunarım.

Ayrıca, bu tez ve doktora eğitimine başlamamda teşviki bulunan Prof Dr. Güray Küçükağaoğluna,mesleğine olan saygısı ve öğrencilerine olan sevgisi ile örnek aldığımız değerli hocamız Prof Dr. Nalan Akdoğan'a teşekkürü bir borç bilirim.

Uzun bir yolculuk olan doktora eğitimimde her zaman bana destek olan sevgili eşim Dr. Tamer Güdü'ye ve tezime ayırdığım zamandan dolayı bana gösterdiği anlayıştan dolayı kızım Deniz Güdü'ye teşekkür ederim. Doktora eğitimime başlamamda ve devam etmemde beni teşvik eden annem Nigar Erkenci ve babam Hüseyin Erkenci'ye teşekkür ederim.

ÖZET

Günümüzde, hile ve suistimal denetimi finansal denetim çerçevesi altında ele alınmaktadır. Hile ve suistimal sonucu kurumların karşılaştığı kayıpların büyüklüğü hile denetimine finansal denetim çerçevesinde yaklaşımın bazı noktalarda yetersiz kaldığına işaret etmektedir. Bu durum, hile denetiminin finansal denetimin sınırları dışına çıkma zorunluluğunu ortaya koymaktadır. İç kontrol, yeni büyük veri çalışmaları, hukuk ve insan psikolojisi günümüzde hile denetiminde ihtiyaç duyulan alanlar olarak karşımıza çıkmaktadır. Uğranılan kayıpların büyüklüğü disiplinlerarası yeni bir yaklaşım ile hile denetiminde hile ve suistimal gerçekleşmeden önce yapılan tespitinin önemini göstermektedir.

Bu çalışmada bir hile tahmin modellemesi yapılmıştır. K. RamaKalyani ve D. Uma Devi (2012) çalışması referans makale olarak alınmıştır. Belirtilen makalede, yazarlar beş adet hile kuralı belirleyerek veri setinde genetik algoritma yöntemi ile hileli işlemleri tespit etmektedir. Bu tezde yazarların Java dilinde yazdıkları kod Matlab'a çevrilmiştir. Yazılan Matlab kodu veri setine uyarlanmış ve yazarların Java kodunda ulaştığı sonuçlara ulaşılmıştır. Ardından, yeni bir metot önerilmiş, veri setine uyarlanmış ve yeni sonuçlara ulaşılmıştır.

Üçüncü bölümde Matlab kodu yeni, benzer ve daha büyük bir veri setine yeni uyarlanmış ve bazı kurallar değiştirilerek hileli işlemler tespit edilmiştir. Bu çalışmada, bazı hile kurallarının tespit edilerek küçük veya büyük veri setlerine uyarlanabileceğini görülmektedir. Bu yöntem, denetçinin veri setlerinde kırmızı bayrak olarak tespit edilen bazı işlemlere yoğunlaşmasına yardımcı olabilir.

Anahtar Kelimeler: Hile Denetimi, Hile Tahmin Modellemesi, Kırmızı Bayrak

ABSTRACT

Fraud Audit is considered within the context of financial audit, today. The huge amounts of loss faced by institutions due to fraud and abuse indicate the inadequacy of the approaching the "fraud audit" within the context of financial audit at some points, today. This situation reveals the necessity of fraud audit to exceed the limits of financial audit. Internal control, big data analytics, law and human psychology are the new areas that are needed in fraud audit, today. The huge magnitude of losses faced indicates the importance of forecasting of fraud and abuse before the offence is committed with a new interdisciplinary approach.

A fraud forecast modelling is done in this research. The article of K. RamaKalyani and D. Uma Devi (2012) is taken as reference article. In the mentioned article, the authors determine five rules of fraud and detect the fraud transactions by genetic algorithm. In this thesis, the code of the rules that is in Java language in the reference article is changed to Matlab. The code in Matlab is applied to the data set and the results are found in line with the result of the code of the authors in Java.

Later, an alternative method is proposed and applied to the data set and new results are found. In the third chapter, the Matlab code is applied to a new, similar and bigger data set and some rules are modified and the fraud transactions are detected. The output of the research shows that some fraud rules may be determined and applied to small or big data sets. This method may assist the auditor to focus on some transactions on the data sets that are detected as red flags.

Keywords: Fraud Audit, Fraud Forecast Modelling, Red Flag

İÇİNDEKİLER

TABLOLAR LİSTESİ

ŞEKİLLER LİSTESİ

SİMGELER VE KISALTMALAR LİSTESİ

SAS	:	Denetim Standartlarına İlişkin Beyan
ACFE	:	Uluslararası Suistimal ve Hile Tespit Uzmanları Birliği
TÜRMOB	:	Türkiye Serbest Muhasebeci Mali Müşavirler ve Yeminli Mali Müşavirler Odaları Birliği
TESMER	:	Türkiye Serbest Muhasebeci Mali Müşavirler ve Yeminli Mali Müşavirler Temel Eğitim ve Staj Merkezi
CFO	:	Mali İşler Yöneticisi
CEO	:	İcra Kurulu Başkanı
KPMG	:	Klynveld Peat Marwick Goerdeler Denetim Firması
PwC	:	PricewaterhouseCoopers Denetim Firması

GİRİŞ

Hile ve suistimal denetimi her geçen gün önem kazanan alanlardan biridir. Hile ve yolsuzluk denetimine kurumların ilgisi her geçen gün kurumların karşılaştığı hile ve suistimal vakalarının artmasından kaynaklanmaktadır. Kurumların uğradığı hile ve suistimaller giderek daha profesyonel yöntemler ile gerçekleştirilmekte ve buna paralel olarak uğranılan kayıplar artmaktadır.

Karşılaşılan tehdit alışılan denetim yöntemleri ile tespit edilememekte ve hile denetimi ve tespitinde daha farklı bir yaklaşıma ihtiyaç duyulmaktadır. Tehditin büyüklüğü günümüz teknolojisinden faydalanmayı zorunlu hale getirmektedir. Birçok kurum önceden uyarı sistemi anlamına gelen "kırmızı bayrak"ları denetim sistemine dahil etmekte ve oluşabilecek hile ve suistimal tehditlerinden kırmızı bayraklar sayesinde suistimal gerçekleşmeden önce haberdar olmaya ve önlem almaya çalışmaktadır.

Bu çalışmada bu konunun seçilmesinin nedeni suistimale uğrayan kişi ve firmaların çoğu kez olay gerçekleştikten sonra bu durumla yüzleşmeleridir. Çoğu zaman, en iyi denetim firmaları tarafından denetlenen kurumlar bile kurumlarında yıllarca ve organize olarak yürütülen hile faaliyetlerinden haberdar olmamakta ve sonrasında beklenmedik bir zaman kesitinde tahmin etmedikleri kişiler tarafından suistimale uğrayarak yüksek maddi kayıplar ile karşılaşmaktadır.

Denetimde klasik denetim araçları hileyle mücadelede yetersiz kalmaktadır ve yeni açılımlara ihtiyaç duyulmaktadır. Disiplinlerarası iş birliği hilenin önceden tahminde önem kazanmaktadır. Tek bir disiplin günümüzde hile ile mücadelede yetersiz kalmakta ve hile ve suistimal sonrası gerçekleştirilen denetimler sadece zararın tespitine ve firmaların iç kontrol sistemlerini yeniden gözden geçirmelerini sağlamakta, hile ve suistimal sonucu uğranılan kaybı geriye getirememektedir.

Bu çalışmanın amacı hilenin önceden tahmininde veri çalışmalarını kullanarak hilenin önceden tahmini çalışmalarına bir denetim aracı olarak katkı sağlamaktır. Tez üç bölüme ayrılmıştır;

İlk bölümde hile kavramı, usulsüzlük denetimi, hilenin tespit edilmesi ve hile tespit teknolojileri incelenmektedir. Öncelikle hile kavramı ve gelişimi incelenmiştir. Ardından hile denetimi ve hile denetiminde kırmızı bayraklar kavramı açıklanmıştır. Hile denetiminde gelişime açık yönler üzerinde durularak ve hile denetiminde önceden tahmin mekanizmasının önemine işaret edilmiştir.

Ardından literatürde kullanılan yaygın hile tahmin yöntemleri açıklanmıştır. Çalışmada Naive Bayes, Destek Vektör Makinesi ve Tek sınıflı Destek Vektör Makinesi, Yapay Sinir Ağları, Karar Ağaçları ve Rastgele Orman, K-En Yakın Komşu Algoritması ve K-Ortalamalar, Açı-temelli Uç Değer tespiti, Kümeleme Bazlı Yerel Uç Değer Faktörü tespiti, Özellik Sınıflandırma, Histogram temelli uç değer, Lokal Aykırı Değer, Minimum Kovaryans Determinantı, Vaka Bazlı Tüme Varım, Evrimsel Algoritmalar yöntemleri incelenmiştir.

Daha sonra hile tespit teknolojilerindeki yaklaşımlar incelenmiştir. Hile tespit teknolojilerinde; literatürdeki kredi kartı hile analiz yaklaşımları üzerinde durulmuştur. Diğer alanlardaki çalışmalar; telekomünikasyon hile analiz yaklaşımları, finansal tablolar hile analiz yaklaşımları, hisse senedi hile tahmin yaklaşımları, sigorta hile analiz yaklaşımları, karapara aklamayı önleme sistemleri, bilgisayar ihlal tahmin sistemleri, veri setine dayalı hile tahmin yaklaşımları, gerçek zamanlı hile tahmin sistemleri, hile tahmininde makine öğrenmesi konuları genel çerçevede açıklanmıştır.

İkinci bölümde K. RamaKalyani ve D. Uma Devi (2012) makalesi referans alınarak bir çalışma yapılmıştır. Yapılan çalışmada referans makalenin içeriği anlatılmış ve oluşturulmuş hile kuralları açıklanmıştır. Makalede oluşturulan Java kodu Matlab koduna çevrilmiş ve Matlab kodu ile java kodu çıktıları ile uyumlu sonuçlara ulaşılmıştır. Ek olarak alternatif bir yaklaşım önerilmiş ve sonuçları ilgili bölümde açıklanmıştır

Üçüncü bölümde yazılan kod yeni ve daha büyük bir veri setine uyarlanmış, kurallar yeni veri setine göre güncellenmiştir. Güncellenen kod yeni veri setine uyarlanmış yeni kurallar üzerinden çalıştırılarak veri setinde şüpheli işlemler tespit edilmiştir.

Sonuç bölümünde varılan sonuçlar özetlenmiştir.

BÖLÜM I.

HİLE DENETİMİ VE HİLE TAHMİNİ

1.1. Hile Kavramı

Hile, bir kişinin satın alınmasını veya bir kişinin diğeri üzerinde gerçeği saptırarak insan zekasının yaratabileceği çeşitli yöntemleri kapsayan kapsamlı bir terimdir. Hile ile ilgili kesin ve net bir tanımlama koymak zordur çünkü hile sürprizi, düzenbazlığı, kurnazlığı ve diğer kişinin aldatıldığı adil olmayan tüm yöntemleri içerir. Hile tanımının sınırlarını çizen sınırları bir anlamda insan zekasının ve dolandırıcılığının sınırlarıdır.

Hile aşağıdaki özellikleri içeren bir aldatmacadır;

1. Açıklama

2. Maddi bir konuya ilişkin

3. Yanlış olan

4. Ve kasti olarak veya gözü kara biçimde

5. İnanılan

6. Ve kurbana göre davranılan

7. Ve kurbanın zararına yol açan (Albrecht ve diğ., 2004).

Hile, gücü veya varlıkları kötüye kullanma, zimmete para geçirme gibi hukuk dışı fiiller olarak da tanımlanmaktadır. Doğan ve diğ. (2017) hukuki boyutu ile hile haksız bir kazanç veya hukuki olmayan fayda elde etmek için bilinçli bir aldatmaca veya mağdurun hukuki bir hakkından mahrum bırakmak anlamını taşır. (Cambridge Dictionary. 2021: https://dictionary.cambridge.org/tr/.)

1.1.1. Hilenin Önemi

Hile ve suistimal konusunun önemine hile ve suistimal sonucu uğranılan kayıpların büyüklüğü işaret etmektedir. 2020 yılında Uluslarası Suistimal İnceleme Uzmanları Derneği "Report to Nations" raporu bu alanda yayınlanan bilinen çalışmalardan biridir. Bu çalışma mesleki hile ve susitimali inceleyen 2504 vaka ve 125 ülkeyi ve $3.6 milyar $'dan fazla kaybı kapsamaktadır.

Rapora göre hile uzmanlarının temel bulguları organizasyonların her yıl gelirlerinin %5'ini hile ve suistimal nedeni ile kaybettiğini göstermektedir. Tahmin, 2019 yılında $90.52 trilyon gelir üzerinden %5 kayıp varsayımı ile yapılmıştır ve toplamda her yıl için $4.5 trilyondan fazla kayba denk gelmektedir.

Çalışma ortalama kaybın vaka başına 1.509.000$ olduğunu göstermektedir. Tipik bir hile vakasının çözülmesi ortalamada 14 ay gibi uzun bir süreyi almakta ve ayda 8.300 $ maliyete yol açmaktadır.

Çalışmada ulaşılan diğer dikkat çekici sonuçlar şunlardır. Hile ve suistimal vakalarına karşı çalışan farkındalığının arttırılmasının önemi ulaşılan çıktılar aracılığı ile de doğrulanmaktadır. Hileye karşı olan suistimal farkındalığı hile ile ipuçları ve bilginin kurumsal raporlama mekanizmasına iletilmesini hızlandırmaktadır. Araştırma sonuçlarına göre, hile ve suistimale ilişkin elde edilen ipuçlarının %56'sının hileye karşı farkındalığı olan personelden alınırken, ipuçlarının %37'sinin konu ile ilgili yeterince farkındalığı olmayan personelden geldiği görülmektedir.

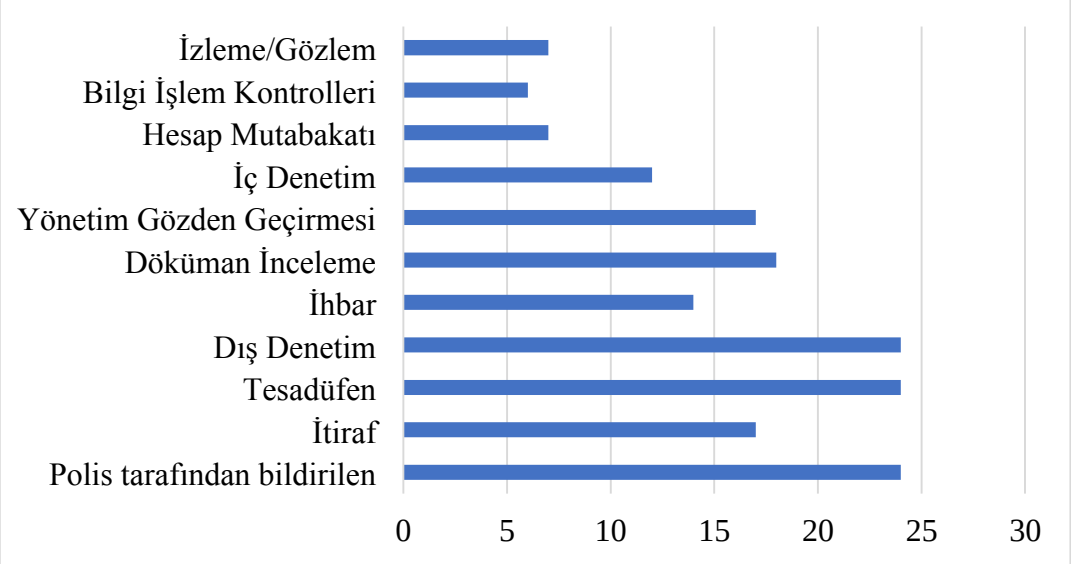
Vaka tiplerine göre de uğranılan kayıpların farklılık gösterdiği gözükmektedir. Araştırma sonucuna göre, varlık suistimali tüm vaka tipleri içerisinde en çok rastlanan ve en az maliyetli vaka tipidir. Varlık suistimal vakaları toplam vakaların %86'sı ve ortalamada $100.000 kayba yol açtığı görülmektedir. Mali tablo yolsuzluğu en az rastlanan vaka tipi olmakla beraber (%10) en yüksek maddi kayba ortalamada $954.000 yol açtığı görülmektedir. Bu vakaların %43'ü ihbarlar aracılığı ile bildirilmekte ve ihbarların yarısı çalışanlardan gelmektedir.

Son 10 yıl içerinde kırmızı hat, hile karşıtı politika, çalışanlar ve yöneticiler için hile eğitimleri gibi hileyle ilgili kontrollerin arttığını gözlemlenmektedir. İç kontrollerin yetersizliğinin hile vakalarının üçte birine katkısı olduğunu çalışma sonuçlarındandır. Hileye karşı iç kontrollerin varlığı hem maddi kaybın tutarını azaltmakta hem de daha hızlı tespite olanak sağlamaktadır. Firma sahipleri veya yöneticilerin tüm yolsuzluk vakaları içerinde %20 oranında suistimal olayına karıştıkları görülürken diğer seviyedeki çalışanlara göre daha büyük (ortalamada $600.000) kayıpları çalıştıkları veya sahip oldukları kuruma verdikleri görülmektedir.

Dolandırıcı profili incelendiğinde, hileyi düzenleyen kişilerin %72'sinin erkek olduğu ve tutar olarak da kadınlara göre daha yüksek tutarlarda ($150.000 erkek/ $85.000 kadın) çalıştıkları kurumları zarara uğrattıkları görülmektedir. Çalışma sonuçlarına göre, vakaların ağırlıklı olarak operasyon, muhasebe, üst yönetim ve satış departmanlarından geldiği gözlenmiştir.

Firma büyüklüğünün de hile risklerini belirleyen etmenler arasında olduğu görülmüştür. Şöyle ki küçük firmalarda büyük firmaların dört katı daha fazla oranında çek ve ödeme ile ilgili usulsüzlüklere, iki katı daha fazla fatura ve maaş bordrosu yolsuzluğuna rastlanmıştır. Çalışmada, hilenin tespit edilme yöntemi ve süresi arasında ekteki grafikteki ilişki tespit edilmiştir.

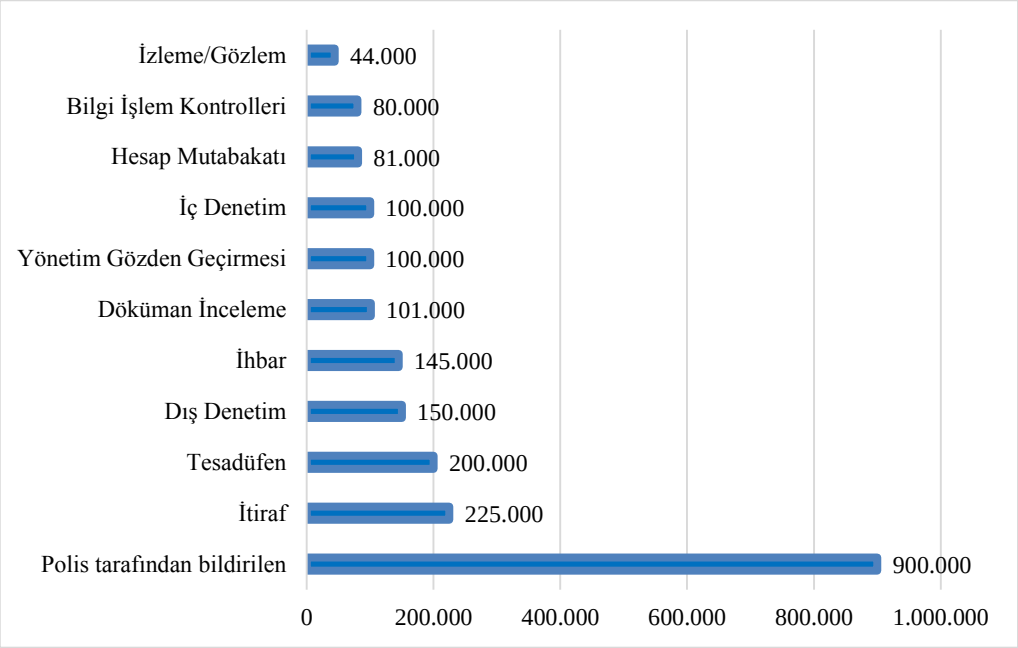
Şekil 1 Hilenin Tespit Yöntemi/ Süresi (Ay)



Kaynak: ACFE Report to the Nations, 2020.

Hile tespit yöntemlerinin süreleri incelendiğinde polis tarafından bildirilen, dış denetim veya tesadüfen tespit edilen vakaların tespit edilme sürelerinin ortalamada iki yıla yakın olduğu görülürken, göreceli en hızlı tespit yönteminin bilgi işlem kontrolleri olduğu görülmektedir.

Şekil 2 Hilenin Tespit Yöntemi/Ortalama Kayıp ($)

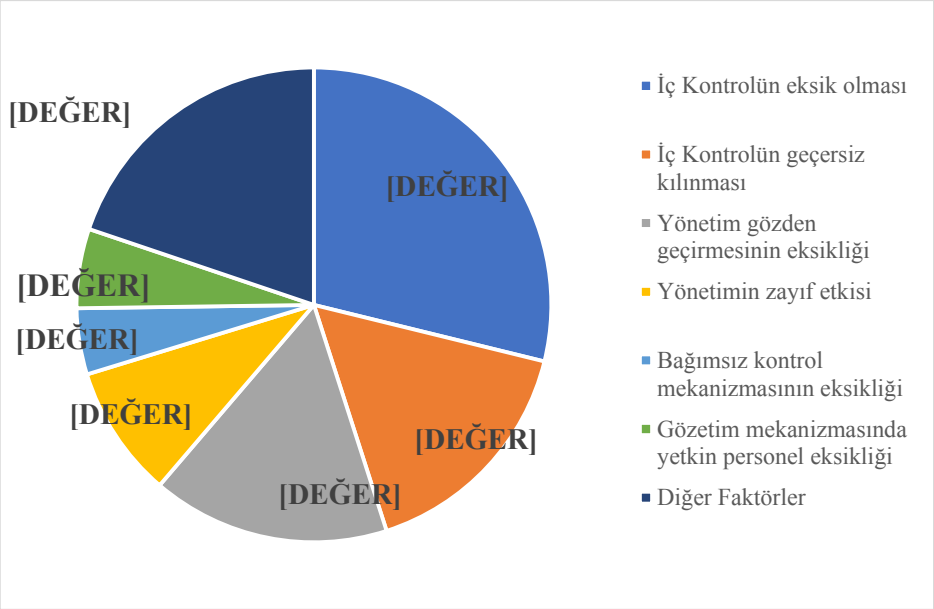


Kaynak: ACFE Report to the Nations, 2020.

Diğer yöntemlere göre, polis tarafından bildirilen dolandırıcılık vakalarının kuruma kaybı daha yüksek tutarlı olduğu görülmektedir.

Hile ile iç kontrolün ilişkisi diğer sık incelenen bir konudur. Bu araştırmada da diğer birçok çalışmaya paralel olarak iç kontrolün eksik olmasının yüksek oranda hile ile olan ilişkisine işaret edilmektedir (%32). İç kontrolün geçersiz kılınması ve yönetimin gözden geçirmesinin eksikliği faktörlerinin ikisi birlikte değerlendirildiğinde (%18 ve %18 toplamda %36 oranı) bu durum iç kontrolün eksikliğinin yansıra iç kontrol ve kurum üst yönetimin iş birliğinin önemine işaret etmektedir.

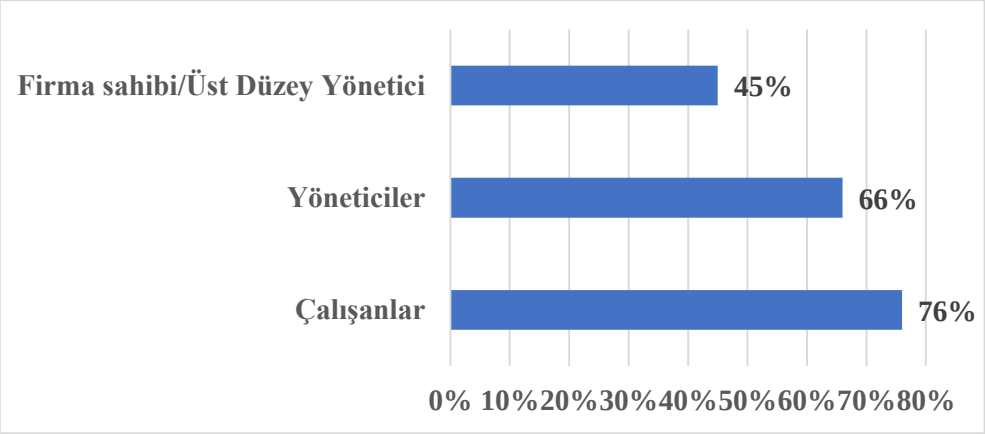
Şekil 3 Mesleki Hileye katkısı olan İç Kontrol Zayıflıkları



Kaynak: ACFE Report to the Nations, 2020.

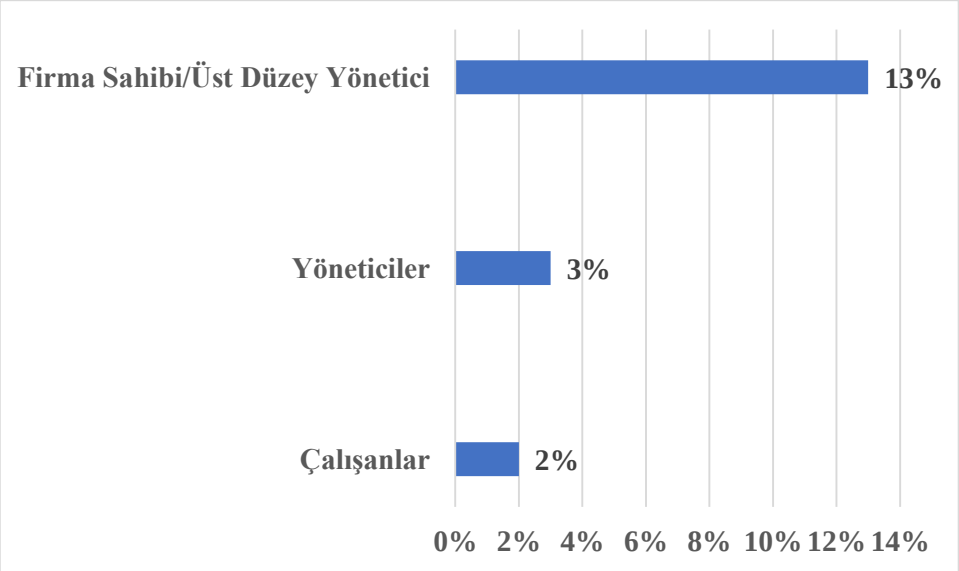
Çalışmanın diğer bir bulgusu hile tespitinden sonra kurumların büyük bir oranda dolandırıcıyı cezalandırma eğiliminde olduğunu göstermektedir. İş akdinin sonlandırılması bu cezalandırma mekanizmalarından biri iken ekteki grafikte görüleceği üzere bu durumun çalışanlara firma sahibi/üst düzey yöneticiye göre daha yüksek oranda uygulandığı görülmektedir. Aynı şekilde ekteki diğer grafikte de görüleceği üzere ceza alınmayan durumların üst düzey yöneticilerde çalışanlara göre daha yüksek olduğu sonucuna varılmaktadır.

Şekil 4 Dolandırıcılık Yüzünden İş Akdinin Sonlandırılması



Kaynak: ACFE Report to the Nations, 2020.

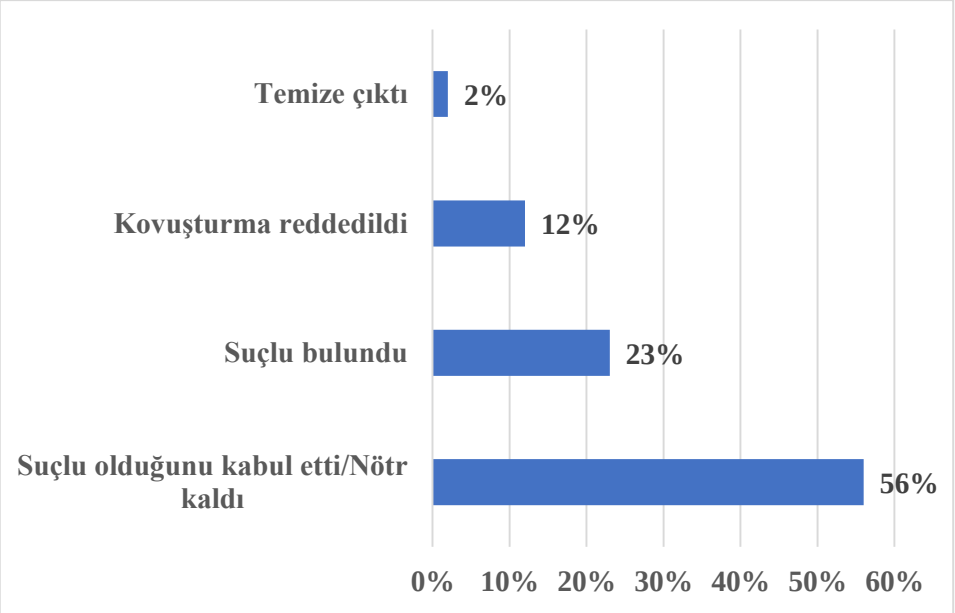
Şekil 5 Ceza Alınmayan Durumlar



Kaynak: ACFE Report to the Nations, 2020.

Hukuki yollara başvurmakta hile ile karşılaşıldığında alınan aksiyonlardan biridir. Hukuki yollara devredilen vakaların yarısından fazlasında dolandırıcılıkla suçlanan kişi kendisi suçlu olduğunu kabul etmekte veya nötr kalmaktadır.

Şekil 6 Hukuki Yollara Devredilen Vakaların Sonuçları



Kaynak: ACFE Report to the Nations, 2020.

1.1.2. Hile Kavramının Gelişimi

Hilenin önemi karşılaşılan kayıpların büyüklüğü ile ölçülebilirken, hile kavramının ortaya çıkışı incelendiğinde hile ile insanoğlunun tanışmasının paranın varlığı ile eş zamanlı olduğu görülmektedir (Experian, 2021).

İlk hile vakasına yazıtlarda Milattan önce 300. yılda karşılaşılmaktadır. Hegestratos adında bir Yunanlı gemi tüccarı gemisi ve taşıdığı yük için bir gemi ipotek sözleşmesi yapar. Anlaşmaya göre taşınan yük mısır satıldığında tüccar borç aldığı para ve faizini borç para aldığı kişiye geri verecek borç veren kişide gemi üzerindeki ipoteği kaldıracaktır.

Tarihin kayda alınmış ilk hile girişimi olarak, Hegestratos boş gemiyi batırma girişiminde bulunur, böylece mısırı ayrıca satacaktır. Gemiyi batırmaya çalışan Hegestratos mürettebi tarafından yakalanmış ve kaçarken boğularak ölmüştür.

Tarihte eski diğer bir kayıtlı dolandırıcılık vakası Roma İmparatorluğu döneminde MS 193. yılında yaşanmıştır. Praetorian Korumaları adı verilen bir grup asker Pertinax adındaki

imparatoru öldürür ve Roma İmparatorluğu tahtını açık arttırmaya çıkarırlar. En yüksek teklifi Juliaunus adında bir adam verir. Teklif ordudaki her bir asker başına 250 parça altındır. Verilen teklif toplamda günümüz parası ile 1 milyar £'ne denktir.

Gerçekte askerlere kendilerine ait olmayan bir varlığı, imparatorluğu satmaktadır ve bu durum günümüz finansal dolandırıcılığına denk gelmektedir. Juliaunus hiçbir zaman imparator olarak kabul edilmemiştir ve sonrasında tarihte 5 imparator dönemi adı verilen iç savaş başlamıştır.

Diğer bilinen tarihi bir dolandırıcılık vakası 1821 yılına aittir. Gregor McGregor ordudaki bir İskoçyalı generaldir. Etrafta savaşta yaptığı kahramanlıklarından bahsetmektedir. Aynı zamanda bir adayı savaşta ele geçirdiğini ve kendisinin bu adanın prensi olduğu dedikodusunu çevresine yaymaktadır. Yatırımcılarına bu adada lüks evler ve cennet gibi bir hayat vaat etmektedir. Birçok insan sürü halinde hareket ederek gerçekte var olmayan bu evleri almış ve sterlini generalin sahte parası ile değiş tokuş etmiştir.

Kayıtlı ilginç bir dolandırıcılık vakası 1911 yılında Arjantinli Eduardo de Valfiernoya aittir. Bu kişi isimsiz bir Lourve çalışanına Leonardo Da Vincinin Mona Lisa'sını çalması için para vermiştir. Eduardo'nun aslında tabloya ihtiyacı yoktur ve tabloyu istememektedir. İnsanlara ünlü resimlerin sahtelerinin yapılarak yeraltı koleksiyoncularına satılabileceğini göstermek istemiştir.

Saadet zinciri günümüzde de yaygın bir dolandırıcılık tipidir. İlk olarak 1920 yılında Charles Ponzi adındaki bir Amerikalı tarafından uygulanmıştır. Bu kişi %5 kar marjı ile yatırımcılara posta çeki satmaktadır. Ardından bunu genişletmiş ve %50 kar sözü ile yatırımcılardan para toplamaya başlamıştır. Sonraki yatırımcılardan topladığı para ile ilk yatırımcılara söz verdiği karı geri ödemektedir. Yakalandığında 10 milyon $'ı vardır ve ülkeden kaçmıştır.

Son 15 yıldır, dolandırıcılık sektörler arasında artmıştır ve şu an trilyonlar bazında dolandırıcılık faaliyeti yapılmaktadır. Teknolojideki gelişmeler insanların hizmetlere online olarak güvenli ve hızlı erişimini artmıştır. Bu durum eş zamanlı olarak dolandırıcılık faaliyetlerinde artmasına neden olmuştur.

1.1.3. Hile Teorileri

Artan dolandırıcılık faaliyetlerine paralel olarak akademik dünyada da bu konuya ilişkin yapılan araştırmalar artmıştır. Hile teorileri hile çalışmalarına hilenin tanımlanması, çeşitleri, belirtileri ve önlenmesine bir çerçeve oluşturmuştur.

Hile teorisinin gelişimi incelendiğinde insan davranışları, çevresel etkenler, aktörler ve sistemlerinin teoriler içerisinde önemli değişkenler olduğunu görülmektedir. Literatüre "Beyaz Yakalı Suçu" –terimini 1940 yılında ilk kez kazandıran Edwin H. Sutherland kazandırmıştır. Sutherland (1940)'a göre; beyaz yakalıların işlediği suçlar ile diğer suçlar arasında bazı farklar bulunmaktadır. Bunlardan biri beyaz yakalıların çoğunlukla profesyonellerden oluşması, bir diğeri faillerin profesyonel olmasından dolayı kapsam dahilinde değerlendirildikleri yasal davaların diğer davalara göreceli olarak daha hafif olmasıdır. Son olarak, beyaz yaka suçlarının daha fark edilmez olmasından dolayı, failleri ve mağdurları tespit etmekte yaşanan zorluklardır (Vardar, 2019).

Sutherland'ın beyaz yakalıların işlediği suçlarla ilgili ortaya koyduğu teori, dolandırıcılıkla ilgili teoriler içinde zamanla önemli bir referans noktası haline gelmiştir. Beyaz yakalı suçu kavramı hem zaman içerisinde kavram olarak kendisini korumuş hem de daha fazla değişkeni içeren hile üçgeni teorisine yol açmıştır.

Hile teorisi ve diğer takip eden teorilerin önemli bir kurumsal çerçeve oluşturmakla beraber kurumsal hilelerin çoğunluğunu açıklayamadığı literatürde tartışılmaktadır. Hile denetiminin uzun zaman boyunca pasif yaklaşımı benimsediği görülmektedir. Bu yaklaşımda iç kontrolü baz alan bir yaklaşım izlenmiş ve iç kontrolün hile belirtilerini veya kırmızı bayrakları iç kontrol sistemi içerisinde tespit edebilme yetisi temel alınmıştır. Hile denetiminde pasif veya reaktif yaklaşımlar ele alınmıştır. Pasif yaklaşımda rastgele ve önyargısız olarak belirli denetim prosedürleri izleme yöntemi izlenirken, reaktif yaklaşımda gerçekleşmiş hile vakalarına karşın alınan aksiyonlar örneğin soruşturma açılması ve iddiaların incelenmesi gibi yöntemler izlenmektedir.

Diğer taraftan proaktif denetim yaklaşımında, denetim süreçleri herhangi bir hile iddiası veya şüphesi olmasa da kurulan bir denetim mekanizması ile işletilmektedir. Bu çerçevede, pasif yaklaşımlar yerine proaktif yaklaşımların günümüz hile vakalarının tespitinde daha fazla

tercih edildiği görülmektedir. Hile teorileri; hile üçgeni ve diğer belli başlı hile teorileri bu yaklaşım çerçevesinde açıklanmaktadır.

1.1.3.1. Hile Üçgeni

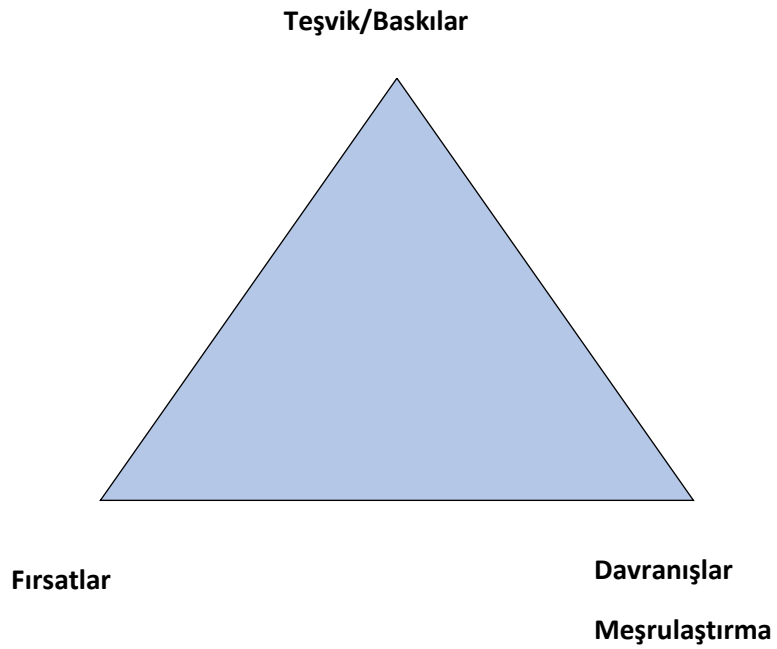
Geleneksel hile denetiminde, kırmızı bayraklar hileli işlemlerin tespitinde işaret olarak kullanılmaktadır. SAS No.53'ün (SAS: Statement on Auditing Standards) yerine getirilen SAS No.82 hile risk faktörlerini yani kırmızı bayrakları ilk tanımlayan düzenlemelerden biridir. Bu standardın yerini daha sonra SAS No.89 almıştır. SAS No.89'da denetimde hileli finansal raporlamanın tespitini için kullanılması istenen 42 kırmızı bayrak açıklanmaktadır (Yücel, 2013).

SAS 99 standardı Amerikan denetçilerinin denetim sırasında finansal hile konusuna önem ve dikkat göstermeleri gerektiğini vurgular. Denetim yevmiye kayıtlarının, düzeltmelerin, muhasebe tahminlerinin ve olağandışı işlemlerin detaylı incelenmesi ve hile ile ilişkili önemli yanlışlıkların araştırılmasını içerir (Moyes ve diğ., 2009).

99 nolu denetim standardı işletmede gerçekleştirilen muhasebe hilelerine üç faktörün etkisinden bahseder; özenti/baskılar, fırsatlar ve davranış/bahanelerdir. Bu üç faktöre hile üçgeni denilmektedir. "Hile üçgeni" hile teorilerinin temel teorisidir. Bu üç unsur teşvik/baskılar, fırsatlar ve davranışlar/meşrulaştırmadır. Hilenin olduğu bir ortamda bir teşvik veya baskı vardır. İkinci olarak hilenin oluşması için bir fırsat vardır. Üçüncü olarak hileyi yapan kişinin kendi davranışını meşrulaştırma tutumu görülmektedir.

IFRS standartlarında ISA 240, Türkiye Denetim Standartlarında 240 nolu Bağımsız Denetim Standardı, bağımsız denetçinin finansal tablolarda hile denetimine ilişkin sorumluluklarını tanımlar. SAS 99, ISA 240 ve ülkemizdeki 240 nolu denetim standartlarının hepsi kırmızı bayrakları üç kategoride incelemektedir. Bu üç kategori ilerleyen bölümde detayları anlatılacak hile üçgeninde temelini oluşturan teşvik, baskı ve meşrulaştırmadır.

Şekil 7 Hile Üçgeni



Yazında hile üçgeni çerçevesinde yapılan çalışmalara rastlandığı gibi, hile üçgenin yetersiz kaldığı ve geliştirilebilecek yönleri de olduğunu belirten çalışmalar mevcuttur. Yapılan çalışmalar hile üçgenin genelde bir veya iki bacağını baz almıştır. Hile üçgenin üç bacağını da baz alarak yapılan çalışmalar göreceli olarak daha sınırlıdır.

Hile üçgeni teorisinin üç bacağını da dikkate alarak yapılan çalışmalardan biri olan, Bell ve Carcello (2000) çalışmalarında bir grup hile yapılan şirket içerisinden seçtikleri örneklemde hile üçgeni şartlarının tamamının bu işletmelerde geçerli olduğunu tespit ettiler. Hernandez ve Groot (2007) teşvik ve baskıların ve fırsatların denetim ortakları tarafından daha yüksek hile riski ile ilişkilendirildiğini yaptıkları çalışmada tespit ettiler. Rezaee (2004)'te incelediği beş hile vakasında hile üçgeni koşullarının incelediği firmalarda geçerli olduğunu gösterdi. Diğer çalışmalar genelde üçgenin bir bacağı üzerinde yoğunlaşmış ve aşağıda açıklanmıştır.

1.1.3.1.1. Teşvik/Baskılar

Hile üçgeni teorisine göre hilenin gerçekleşmesi için ilk koşul hilenin oluştuğu ortamda teşvik/baskı olmasıdır. Literatürde teşvik/baskı ve hile ilişkisi üzerine olan çalışmaların

bazılarında gelirlerin olduğundan farklı gösterilme nedenlerini incelemiştir. Bu nedenler arasında analistlerin piyasa beklentilerini karşılamak, tazminat ve teşvik yapıları, dış borçlanma ihtiyacı veya düşük performans nedenleri gösterilmektedir.

Bu çalışmalardan biri olan Dechow, Sloan ve Sweeney (1996) 92 firma üzerinde yaptığı çalışmada firmalar arasında finansal borçlanmayı daha düşük maliyet ile sağlayabilmek için gelir kalemleri üzerinde hile yapıldığı tespit edilmiştir. Dechow ve diğ. (1996) ve Erickson, Hanlon ve Maydew (2006) yılında yaptıkları araştırmada yönetim hisse teşvikleri ile muhasebe hileleri arasındaki ilişkiyi araştırmışlardır. Çalışmada, 1996-2003 döneminde hileli bir firma grubunu araştırılmış ve ancak hisse teşvikleri ve firmanın hileli finansal bilgi raporlaması arasında bir ilişki bulunamamıştır (Erickson ve diğ., 2006).

Efendi, Srivastava ve Swanson (2007) yaptıkları araştırmada tekrar eden finansal bildirimde bulunan firmalar üzerinden seçtikleri bir grup firma örneğinde gerçeğe aykırı finansal beyan olasılığının CEO'nun şirket hisseleri üzerinde opsiyona sahip olduğu zaman arttığını gözlemlemişlerdir (Efendi ve diğ., 2007). Benzer şekilde, Burns ve Kedia (2006) yaptıkları çalışmada hisse opsiyonları ile hileli raporlama arasında ilişki tespit etmişlerdir. Beneish (1999a) SEC uygulamalarına tabi bir grup firma üzerinde yaptığı çalışmasında yöneticilerin gelirlerin yüksek olduğu dönemlerde şirket hisselerini satma eğiliminde olduklarını göstermiştir. Beneish'e göre bu durum içeriden öğrenmenin gelirlerin normalden fazla gösterilmesinin bir nedeni olabileceğini göstermektedir (Beneish, 1999b).

Summers ve Sweney (1998) yılında yaptıkları çalışmada Beneish ile benzer sonuçlara ulaşmış ve içeriden bilgi sızdırma ile hile arasında ilişki tespit etmişlerdir. Lie (2005) yaptığı çalışma hisse opsiyonun hileli davranışa teşvik sağladığına dair kanıt içermektedir.

Rosner (2003) düşük performanslı firmalar örneklemi üzerinden yaptığı çalışmasında firmaların gelir arttırıcı muhasebe hilelerine daha fazla başvurup başvurmadığını ve başarısız firmalarda denetçilerin hileli finansal raporlama ile daha fazla karşılaşıp karşılaşmadığını sorgulamıştır. Bulgular bu tip firmalarda düşük nakit akışları ve karın olduğu yıllarda gelirlerin olduğundan fazla gösterildiğini ve yüksek nakit akışı ve kar edilen yıllarda bu hileli gösterime yönelik düzeltmelerinin yapıldığını göstermektedir.

1.1.3.1.2. Fırsatlar

Denetim standartları No.99 Bölüm 316'da hileli finansal tablo oluşturma fırsatını arttıran faktörleri açıklamaktadır. Bu faktörler içinde bulunulan endüstrinin yapısı, işletmenin dikkat çekici, karışık veya ilişkili taraf işlemleri, yönetimin etkin olmayan şekilde denetimi, birçok farklı yasal yapıdan oluşan karışık bir organizasyon yapısı ve kontrollerin eksik denetlenmesinden veya engellenmesinden kaynaklı etkin olmayan kontroller yer almaktadır.

Albrecht ve Albrecht (2003) yaptıkları çalışmada etkin bir kontrol yapısının hileli faaliyetlerin yapılma olasılığını en aza indirmede veya ortadan kaldırmada en önemli basamak olduğunu belirtmektedir. Birçok çalışma zayıf kurumsal yönetim belirtisi olarak görülen etkin olmayan yönetim denetiminin hile olasılığı ile olan ilişkisini göstermiştir. Örneğin, Dechow ve diğ. (1996) yaptıkları çalışmada gelirler üzerinde manipülasyon yapan firmaların birçoğunun yönetim kurullarının yeterince bağımsız olmadığını, daha yüksek olasılıkla bir yönetim kurulu başkanlığı ve CEO için tek kişilik bir yapıya sahip olduklarını ve CEO'nun aynı zamanda firmanın da kurucusu olduğunu, daha düşük bir olasılıkla bir denetim komitesi ve dışarıdan bir ortağa sahip olduğunu göstermiştir

Beasley (1996) yaptığı çalışmanın bulguları hileli finansal raporlama yapan firmalarda bağımsız denetim kurulu üyesi oranının daha düşük olduğu yönündedir. Farber (2005) bulgusu da bu çalışmayı destekler niteliktedir. Hilenin tespit edildiği firmalar zayıf kurumsal yönetime sahip ve daha az denetim kurulu üyesi olan, daha az denetim toplantısı yapılan, denetim kurulunda daha az finans uzmanı bulunan, daha az sayıda 4 büyük denetim firmasından hizmet alan ve CEO'ları daha yüksek oranda yönetim kurulu başkanı olan firmalardır.

Abbott, Parker ve Peters (2004) denetim komitesi yapısının bağımsızlığının, faaliyet seviyesinin, finansal uzmanlığın finansal tabloların yeniden düzenlenmesi veya hileli olması üzerindeki etkisine işaret etmektedirler

Araştırmacılar, 1991-1999 arasında finansal tablolarını yeniden düzenleyen 88 firma ile birlikte 44 hileli raporlama yapan firmayı incelediler. Denetim komitesinin bağımsızlık ve faaliyet seviyesinin finansal tabloların yeniden düzenlemesi ile negatif yönlü ilişkili olduğunu

tespit ettiler. Aynı şekilde denetim komitesinde finansal uzmanlığı bulunan bir üyenin olması ile finansal tabloların yeniden düzenlenmesi arasında da negatif yönlü ilişki tespit edildi.

McMullen ve Raghunandan (1996) aynı şekilde finansal raporlama problemi yaşayan firmaların daha az olasılıkla bağımsız üyelerden oluşan denetim komitesine sahip olduğunu gösterdiler. Diğer araştırmalar denetim kurulu üyelerinin finansal ve denetim okuryazarlığının finansal raporların kalitesini arttırdığını göstermektedir (McDaniel ve diğ., 2002; Bedard ve Johnstone, 2004).

Loebbecke, Eining ve Willingham (1989) yaptıkları çalışmada finansal hile deneyimi yaşamış olan denetim ortakları ile yaptıkları ankette hileli finansal raporlama olasılığının kararların baskın olarak yönetim tarafından verildiği ve zayıf iç denetim yapısı olan firmalarda daha yüksek olduğunu gösterdiler.

Akademik yazın, daha zayıf kurumsal yönetime sahip olan firmaların daha yüksek hileli finansal bilgi raporlama olasılığına sahip olduğunu göstermektedir. Kurumsal yönetimin zayıf olduğu işletmelerde hile fırsatlarının daha fazla olduğu görülmektedir.

1.1.3.1.3. Davranışlar/Meşrulaştırma

Hile teorisinin davranışlar ve meşrulaştırma bacağı ile ilgili yapılan çalışmalara incelendiğinde muhasebe standartlarının hile yapan yönetici üzerindeki etkisi üzerine çalışmalar dikkat çekmektedir. Muhasebe standartlarının kesin ya da esnek olarak tanımlanması yöneticinin hile karşısındaki tutumunu etkileyebilmektedir. Örneğin, Nelson, Elliott ve Tarpley (2002) yaptıkları çalışmada muhasebe standartlarının kesinliği veya net olmayışının yapılan muhasebe hilelerinin tipini etkilediğini tespit etmişlerdir.

Aynı alanda, Hernandez ve Groot (2007) yaptıkları çalışmada dört büyük denetim ortağının risk değerlendirmeleri inceleyerek denetim firmasının müşteri seçiminde veya kabulünde yöneticinin dürüstlüğü ve etik değerlerinin hile riski değerlendirmesinde en önem verilen faktör olduğu yönünde sonuca ulaştılar. Diğer faktörler gelir tahakkuklarının ve muhasebe tahminlerinin olduğundan fazla gösterilmesi olarak belirtilmiştir

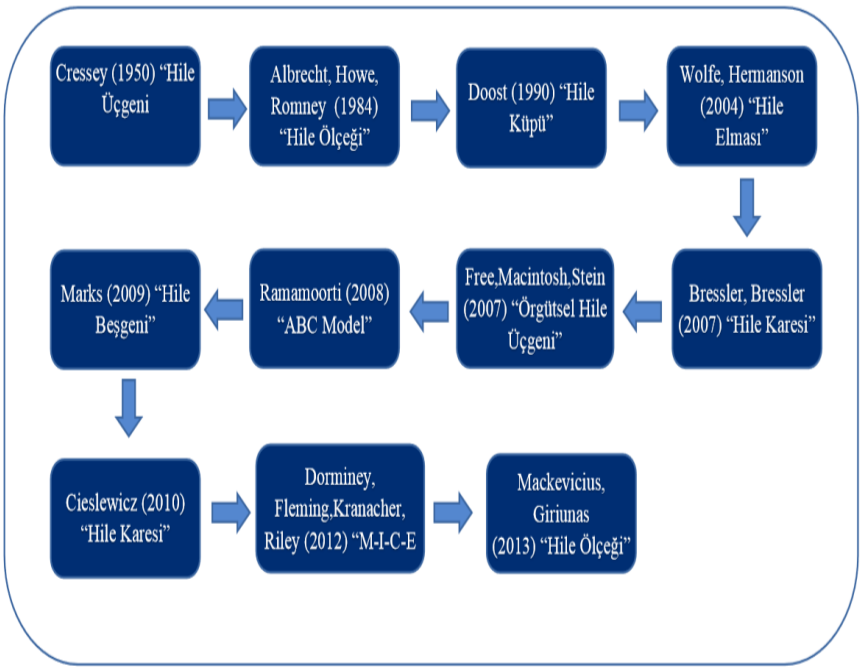
Gillett ve Uddin (2005) CFO'nun hileli finansal raporlamaya dair tutumunun hileli raporlama niyeti üzerinde büyük etkisi olduğunu gösterdiler. Aynı çalışma, teşvik yapısının hileli finansal raporlama niyeti üzerinde aynı derecede etkili olmadığını gösterdi.

Özet olarak, akademik araştırmalar hile üçgeninde ve denetim standartlarında tanımlanan teşvik, fırsat ve davranışlar ile hilenin varlığı arasında ilişki olduğunu göstermektedir. Bu durum bu teşvik ve fırsatları tanımlayan kontrol listelerinin denetim planının oluşturulmasında faydalı olup olmayacağını sorusunu getirdi. Literatürde, kontrol listelerinin hileyi tespit etmedeki faydası konusunda tek bir sonuca ulaşılmadığı görülmektedir (Hogan ve diğ., 2008).

1.1.3.2. Paradigma Değişimleri

Hile üçgeni halen birçok akademik çalışmanın ve denetim standartlarının temelini oluştursa da zaman içerisinde diğer teorilerde literatürde yerini almıştır. Hile üçgeninden sonra hile elması, hile karesi ve hile ölçeği gibi farklı modeller tartışılmıştır. Yazında, çeşitli hile teorileri olup önemli olanlarının kronolojik sıralaması aşağıdaki şekilde özetlenmiştir.

Şekil 8 Hile Modellerinin Kronolojik Sıralaması



Kaynak: Vassiljev ve Alver, 2016; Free, 2015.

1.1.3.3. Hile Küpü

Bilgisayar dolandırıcılıklarının, ilgili kurumla doğrudan bağlantılı kişiler tarafından gerçekleştirilebildiği gibi günlük operasyonlarla doğrudan teması olmayanlar yabancılar tarafından da gerçekleştirildiği görülmektedir.

Doost (1990) makalesinde bu konuyu işlemiş ve hile küpü teorisini açıklamıştır. Teoride, bilgisayar suçlarının işlenmesinde önemi olan üç farklı boyut olduğunu belirtilmektedir; bireyin çevre ile olan ilişkisi, alanında uzmanlık ve harekete geçebilmesi için gerekli içgüdü. Modelde, bu üç etmenin kişilerin hile yapmasına etkileri üzerinde durulmaktadır.

1.1.3.4. Hile Elması

Hile üçgeninden sonra bilinen teorilerden biri de hile elması teorisidir. Hile elması modeli Wolfe ve Hermanson (2004) tarafından geliştirilmiştir. Hile üçgeninde yer alan baskı, fırsat ve haklı göstermeye ek olarak kişinin kabiliyeti dördüncü unsur olarak hile elması teorisinde yer almaktadır. Kabiliyetin kaynağı, hilekârın bulunduğu makamdan / pozisyondan veya organizasyon yapısı olabilir. Hile elmasında yer alan unsurları aşağıdadır;

- Baskı: İhtiyacım var, yapmak zorundayım,
- Fırsat: Sistemdeki zayıflık sayesinde hileyi yapabilirim,
- Haklı Gösterme: Hilenin alacağım riske değeceği konusunda kendini ikna etme,
- Kabiliyet: Bunu başarmak için gerekli özelliklere ve yeteneklere sahibim.

Teoriye göre kabiliyete sahip olan kişinin özellikleri şunlardır;

- Kişinin şirket kural ve uygulamaları istismar edecek pozisyon ve yetkiye sahip olması,

- Hilekârın iç kontrol zayıflıklarını, fark edecek ve kullanacak derecede zeki olması,
- Hilekârın yakalanmayacağına inanan güçlü bir egoya ve güvene sahip olması ve eğer yakalanırsa da farklı bahaneler ile bu durumun üstesinden gelebileceğine inanması,
- Gerektiğinde diğer kişileri de hile yapmaya zorlaması veya ikna edebilmesi
- Stresle kolay başa çıkabilmesi,
- Kesin ve tutarlı yalan söyleme özelliğine sahip olmasıdır.

1.1.3.5. Hile Ölçeği

Mackevicius ve Giriunas (2013) makalelerinde belirtilen "Hile Ölçeği" göre, hile modelinin unsurları şunlardır;

- Motivasyon,
- Koşullar,
- Kapasite,
- Farkındalık/ Kabullenme

Motivasyon, kişinin yaptığı hilenin altındaki temel nedenidir. Bu nedenler ekonomik, psikolojik, kişisel ekonomik veya ekonomik olmayan nedenler olabilir.

Koşullar aşağıdaki şekilde özetlenebilir;

a) Kurum yöneticilerin dürüstlük ve etik değerlere yaklaşımı;

b) Kurumun çalışanları;

c) Kurumun organizasyonel yapısı;

d) Kurumun finansal durumu;

e) Kurumun faaliyetlerinin organizasyon yapısı

f) Muhasebe, denetim ve iç kontrol sistemi;

g) Dış koşullar

Olasılıklar; hilenin yapılabileceği olasılıklar olarak değerlendirilir. Olasılıklar çalışanın kurum içerisindeki rolü ve gücü, kurumun varlık ve muhasebe kayıtlarına ulaşabilme gücü olabilmekte ve bu durum kurumun iç kontrol yapısının etkinliği ile de ilişkilendirilmektedir.

Dördüncü unsur farkındalık ve kabullenmedir. Bu özellik hilenin gerçekleştirilme performansı ile de ilişkilendirilebilir çünkü hilenin gerçekleştirilme performansı kişinin dürüstlüğü ve ahlakının hile yapmasına izin verecek kişisel özellikleri ile ilişkili bir durumdur.

1.1.3.6. ABC Modeli

Model Ramamoorti, Morrison, Koletar ve Pope (2013) hile bileşenlerinin analizi olarak sunulmuştur. Bu modelde hileye iten unsurlar A, B, C harfleri ile belirtilir. ABC modelinde "A" harfi hile yapanı hile yapmaya teşvik eden bireysel özelliklerini ifade eder. B harfi grup ve ortak yapılan hile davranışlarını, C harfi hileye yol açan toplumsal faktörleri belirtir.

1.1.3.7. Hile Beşgeni

Kuram Marks (2009) tarafından geliştirilmiştir. Hile üçgeni teorisinin hile davranışını açıklamada bazı açılardan eksik kaldığını öngörerek, teoriye iki unsur daha eklenmesini önermektedir. Hile beşgeni teorisinde hile üçgenine ek olarak kibir ve beceri unsurları da eklenmiştir. Kibir dolandırıcı kişinin kurum kural ve prosedürlerine uyması gerekmediğine olan kişisel inancı olarak tanımlanabilir. Beceri ise dolandırıcının iç kontrol mekanizmalarını geçersiz kılma gücü, içinde bulunulan iç kontrol zayıflığını kendi avantajına çevirebilme sosyal becerisi, çevrelerindeki diğer kişileri güç kullanarak veya baskı altına alarak kontrol etmeleri ve bu şekilde yapılan usulsüzlüğü örtbas etme becerisi olarak nitelendirilebilir.

1.2. Usulsüzlük Denetimi

1.2.1. Usulsüzlük Denetiminde Önemlilik Seviyesinin Tespiti

Bağımsız Denetim Standartlarında önemlilik finansal tablolarda yer alan tek veya topluca yanlışlıkların finansal tablo kullanıcılarının kararlarını etkileme olasılığıdır. Bir yanlışlık bir finansal tablo kullanıcısının vereceği kararı olumsuz yönde etkiliyorsa o yanlışlık önemli kabul edilmekte, etkilemiyorsa da önemsiz olarak varsayılır

BDS 240 göre denetçi mali tablolarda hile kaynaklı "önemli yanlışlık" risklerini belirlemek ve değerlendirmekle sorumludur. Hile kaynaklı önemli yanlışlık risklerinin tespiti daha zor olabilir ve denetçilerin bu eylemleri ortaya çıkarması için ayrıntılı değerlendirmeleri gerekir (Türedi, 2020). Kontrol listelerinin kullanılması, soruşturma tekniklerinin kullanılması, analitik inceleme teknikleri, veri madenciliği bu araçlar arasında değerlendirilebilir.

1.2.2. Usulsüzlük Denetiminde Kırmızı Bayraklar

Bir şirketin operasyonlarında veya şirket yapısı içerisinde, hile kaynaklı önemli yanlışlık riskinin tespit edilmesini sağlayan uyarılar yer alabilmekte, hatta gerek tahrif edilmiş veya doğruluğu sorgulanabilir belge ve dokümanların gerekse de çalışanlar arasındaki alışılmışın dışındaki ilişkilerin varlığı gibi hilenin gerçekleştiğine ilişkin birtakım işaretler görülebilmektedir (Coenen, 2008). Bu işaretlere "Kırmızı Bayraklar" denilmektedir. Şirketlerin bu tür işaretleri tanımaları muhtemel bir hile vakasını tespit edebilmeleri açısından önem taşıdığından, en ufak bir işaret bile göz ardı edilmemesi önerilmektedir.

Kırmızı bayraklar olarak tanımlana işaretlerin herhangi bir suistimale konu olup olmadığının araştırılması önem arz etmektedir. Hile belirtileri muhasebe yapısındaki bazı işaretlerden, iç kontrol sistemindeki zafiyetlere, sistemsel analitik işaretlere ve çalışanların olağandışı yaşam ve davranış şekillerine uzanan geniş bir yelpazede yer alabilir (Güneş, 2014).

Kırmızı bayraklar hile üçgeni teorisi ile uyumlu olarak 3 ayrı grupta incelenmektedir. Fırsat kırmızı bayrakları, kişilerin etkisiz iç kontroller, yetersiz gözetim ve iç kontrollere

uymayan yöneticiler gibi ortamlar sayesinde hileli işlemleri rahatlıkla yapabilecek kişiler için uygun durumlardır. Baskı kırmızı bayrak giderleri bütçenin altında gösterme, satış ve karları yüksek gösterme gibi hile yapmak için finansal teşvikleri olan çevrelerdir. Rasyonelleştirme kırmızı bayrakları kişilerin hile yapmak için belirli davranış ve yeteneklere sahip olmalarını ve doğru olduğuna inandıkları nedenlerle hileyi rasyonalize etmeleri durumudur.

Kırmızı bayraklar firmanın veya kişinin kendi çıkarı için hileye eğilim doğuran olaylar, durumsal stres, kişilik özellikleri veya olanaklar olabilir (Romney ve diğ., 1980). Kırmızı bayrakların bilinmesi denetçinin hilenin oluşmasına yol açan koşulları daha iyi anlamasını sağlamaktadır.

Yapılan araştırmalar denetçilerin hile vakalarının sadece %20'sini tespit edebildiklerini göstermektedir (Wells, 1990). Daha yüksek oranda hile tespiti için kırmızı bayrakların denetçiler tarafından daha etkin kullanılması önerilmektedir (Cottrell ve Albrecht, 1994).

Bu kapsamda, hile denetimindeki kontrol listelerine kırmızı bayrak adı verilen hile belirtileri dahil edilmektedir. SAS No.99 kırmızı bayrakları risk faktörleri olarak tanımlamaktadır ve daha sonra bu risk faktörlerini hile üçgeninde yer alan teşvik/fırsat/davranışlar alanlarında sınıflandırmaktadır. Ekteki bölümde üçgenin bileşenlerine göre hile denetiminde SAS 99'a göre kullanılan kırmızı bayrak listesi açıklanmaktadır.

1.2.2.1. Teşvik ve Baskı Kırmızı Bayrakları

- Teşvik ve Baskı Kırmızı bayrakları şunlardır;

- -Borsa veya borç geri ödemeleri yeterliliklerini sağlayabilmek için sınırlı kapasiteye sahip olmak

- -Özellikle aynı sektördeki firmalara göre beklenmeyen hızlı büyüme veya normal dışı karlılığa sahip olmak

- -İflas, kapanma veya el değiştirme risklerini doğuran operasyonel kayıplar

- -Yönetim kurulu veya yöneticilerin kurum üzerinde belirgin finansal çıkarlarının bulunması

- -Yüksek şirket hedeflerini gerçekleştirmeye bağlı olan yönetici ikramiyeleri

- -Yatırım analistlerinin, kurumsal yatırımcıların, belli başlı yatırımcıların veya üçüncü tarafların kurumla ilgili gerçeküstü karlılık beklentileri, açıklamaları veya yıllık rapor mesajları

- -İş birleşmeleri, sözleşme ihalelerine ilişkin algılanan veya gerçek olumsuz finansal raporlama sonuçlarının etkileri

- -Yönetimin basın açıklamaları veya yıllık raporlarında gerçekçi olmayan karlılık veya eğilim beklentisi içerisinde olması

- -Firmanın operasyonlarından negatif nakit akışına sahip olması veya gelir veya gelir büyümesi raporlarken nakit akışı yaratamamak

- -Yöneticilerin veya yönetim kurulunun firmaya ait kişisel garantileri aracılığı ile verdikleri borçlarının olması

- -Firmanın faaliyet gösterdiği sektörde yüksek rekabet olması veya piyasanın doyuma ulaşarak giderek düşen kar marjları ile karşı karşıya kalınması

- -Müşteri talebi veya iş kapasitesinde belirgin düşüşler

- -Firmanın faaliyet gösterdiği sektörde veya ekonominin genelinde yaşanan daralmalar

- -Rekabetçi kalmak için gerekli olan temel araştırma ve geliştirme faaliyetlerinde ve sermaye harcamaları için ek borç veya özsermaye finansmanına ihtiyaç duymak

- -Teknolojideki hızlı değişimlere, üretilen ürünün modasının geçmesine veya faiz oranlarına karşı yüksek düzeyde hassasiyete sahip olmak

- -Yeni muhasebesel, yasal veya düzenleyici gerekliklikler (Yücel, 2013).

1.2.2.2. Fırsat Kırmızı Bayrakları

Fırsat kırmızı bayrakları şunlardır;

- -Finansal raporlama ve iç kontrol sistemi üzerinde etkisiz yönetim kurulu veya denetim komitesi gözetiminin olması
- -Temel iç kontrollerin yetersiz takibi
- -İspatı zor olan sübjektif yargı veya belirsizlikleri içeren varlık, kaynak, gelir ve gider tahminleri
- -İşin normal süreci dışında belirgin ilişkili taraf işlemlerinin varlığı veya denetim görmeyen ilişkili taraflarla veya başka bir denetim firması tarafından denetlenen ilişkili taraflarla olan işlemler
- -Sahibi olmayan kişiler tarafından yönetilen firmalarda bir kişi veya kişilerin yönetimde baskın olması
- -Yöneticinin veya yönetim kurulunun sıkça değişmesi
- -Normal dışı yönetim kanalları veya hukuki varlıklar içeren çok karışık organizasyonel yapı
- -Özün önceliği sorularını doğuran yılsonu kapanışına doğru dikkat çekici, normal dışı ve çok karışık işlemlerin varlığı
- -İş süreci ile ilişkili olmayan vergisiz ülkelerde açılmış dikkat çekici banka hesapları veya şube operasyonları
- -Raporlanabilir durumları içeren verimsiz muhasebe ve bilgi sistemleri

- -Yüksek işe giriş-çıkış oranları, etkin olmayan muhasebe, iç kontrol ve bilgi teknolojileri personeli

- -İşletmede kontrol gücü olan organizasyonu veya kişileri belirlemede belirsizlik olması ve güçlük çekilmesi

- -Emsallerine uygun olmayan veya uygunsuz işlemlerle sonuçlanabilecek şart veya koşulların tedarikçi veya müşterilere dikte edilmesini sağlayan belirli bir sektöre hakim olma kapasitesi veya güçlü bir finansal varlık

- -Farklı kültürlerin ve iş ortamlarının olduğu yetki alanlarında uluslararası sınırlar çerçevesinde yürütülen veya yerleştirilmiş belirgin faaliyetler (Moyes ve diğ., 2009).

1.2.2.3. Meşrulaştırma Kırmızı Bayrakları

Meşrulaştırma kırmızı bayrakları şunlardır;

- Kıymetli evrak hukuku ihlali veya haklarında dolandırıcılık iddiası gibi üst yönetim, yönetim kurulu üyelerine ait geçmişte vakaların bulunması

- Özellikle denetçilerin çalışmasının çerçevesini etkilemek amacıyla yapılan girişimler dahil denetçi ile olan iletişimde baskın yönetim davranışı

- Vergi nedeniyle gelirlerin en düşük seviyede raporlanmasına yönelik uygunsuz araçların yönetim tarafından uygulanma isteği

- Raporlama, denetim ve muhasebe konularında şu anki veya önceki denetçi ile yapılan tartışmaların sık olması

- Yönetimin iç kontrollerdeki raporlanabilir şartları zamanında düzeltememesi

- İşletmenin hisse fiyatları veya gelirini arttırmaya yönelik yönetimin aşırı ilgisi

- Önemli tahminlerin belirlenmesi veya muhasebe prensiplerinin seçimine finans dışı yönetimin aşırı ilgisi

- Denetim raporunun yayınlanması veya denetimin tamamlanmasına ilişkin gerçekçi olmayan zaman sınırlaması talebi denetçi üzerinde gerçekçi olmayan taleplerin olması

- Gerçekçi olmayan veya iddialı hedefleri gerçekleştirilmesi için üçüncü tarafların, kredi verenler, analistlerin yönetim tarafından kabul edilmesi

- Denetim komitesi, yönetim kurulu ile iletişime geçmesini engellemek veya kişilere veya bilgiye erişimini uygun olmayan şekilde engellemek amacıyla denetçinin üzerinde resmi veya gayrı resmi engeller olması

- Önemlilik ilkesini göstererek yönetimin marjinal veya uygun olmayan muhasebeyi meşrulaştırmak için yönetimin tekrar eden girişimlerinin olması

- Yönetim tarafından kuruma ait etik standartların, kurum değerlerinin etkin olmayan şekilde uygulanması, desteklenmesi ve iletişime geçilmesi

1.2.2.4. Kırmızı Bayrak Gelişim Alanları

Araştırmalar, hile belirtilerine sık rastlanmasına rağmen tüm belirtilerin her zaman hile ile sonuçlanmadığını göstermektedir (Albrecht ve diğ., 1986). Aynı zamanda hile belirtilerini birleştirip ağırlık vererek toplam hile riskini değerlendirmek ve bir denetim planı oluşturmanın kolay olmadığı belirtilmektedir (Patterson ve Noel, 2003).

Kırmızı bayrak adı verilen hile belirtilerini içeren kontrol listelerinin hile tespitinde kullanılması konusunda araştırmalar mevcuttur. Kontrol listelerinin hilenin tespitindeki faydalarına dair çıktılar farklı sonuçlar vermektedir. Örneğin, bir grup araştırmacı Pincus (1989) hile belirtilerini içeren kontrol listesinin kullanımının hilenin tespitinde etkisi olmadığını sonucuna varırken diğer bir grup araştırmacı etkisi olduğu sonucuna ulaşmıştır (Asare ve Wright, 2004).

Kontrol listelerinin hile denetimi üzerindeki etkisi araştırıldığı gibi hileye dair denetim standartlarının hile denetimi üzerindeki etkisi de araştırılmıştır. Bulgular kontrol listelerinin etkisi gibi bu alanda da farklıdır. Sadece SAS No.89'da tanımlanan hile risk faktörlerine dayalı olarak yapılan denetim değerlendirmeleri bazı durumlarda yetersiz kalmaktadır göstermektedir. Benzeri çalışmalar SAS No.89 öncesi ve sonrası denetçi davranışlarını incelemiştir. Bulgular SAS No.89 sonrası dikkat ve planlamanın hile risk faktörlerinden sonrası arttığını ancak denetim sırasında uygulanan prosedürlerinde fazla bir değişiklik gözlemlenmediğini göstermiştir (Glover ve diğ., 2003; Zimbelman, 1997).

Wilks ve Zimbelman (2004) yaptıkları ilk çalışmada denetim çalışmasının SAS No.89'da hile üçgeni çerçevesinde teşvik/fırsat/davranışlar olarak bölümlenmesinin denetçinin hile riskini tespit etmeye faydası olup olmadığını ölçtüler. Bulguları bu bölümlemenin denetçi risk hassasiyetini arttırdığı yönündedir.

1.2.3. Hilenin Tespit Edilmesi

Hile riskinin tespit edilmesine ilişkin yöntemler literatürde klasik yöntemler ve proaktif yöntemler olarak ikiye ayrılmaktadır.

Klasik yöntemler; iç denetim, bağımsız denetim, çalışanların gözlenmesi, ihbar hatları, etik kurallara uyum yöntemleridir. Vardar (2019), klasik yöntemlerde pasif yaklaşım öngörülmektedir. Pasif yaklaşımda hile denetimi istek ve ihbar olduğunda yani sadece iç kontrol sistemine dışarıdan bir etki geldiğinde, içeriden gelen etkiye verilen tepki olarak gerçekleşmektedir.

Proaktif yöntemler ise denetimde önceden önlem alma mekanizmalarıdır. Proaktif yöntemler; analitik inceleme teknikleri, veri madenciliği, sürpriz denetimler, sürekli denetimler, gözetim, dijital analiz, fısıltı yöntemleridir.

1.2.3.1. Veri Madenciliği

Proaktif yöntemlerin hile tespitinde daha etkin sonuçlar verdiği görülmektedir. Veri madenciliği de proaktif hile tespiti yöntemleri arasında önemli bir araç olarak yer almaktadır. Hile alanında kullanılan veri madenciliği, belirli bir hile durumuna işaret eden anomalileri

veya kalıpları tanımlamak için işlemsel verileri elde etme ve analiz etme işlemidir. Bu noktada, veri madenciliği, yığınlanmış bilgisayar verileri arasından çeşitli istatistik ve matematiksel testler aracılığıyla bu verilerdeki gizlenmiş olguları bulmaya yarayan, tespit edilebilmesi zor ilişkileri ortaya çıkaran ve çeşitli tahminler yapılabilmesini sağlayan veri tabanı teknolojisini ifade etmektedir. Şüphesiz ki çeşitli ticari veri madenciliği yazılımlarının tespit ettiği anomaliler, hilenin bir ispatı değil, kırmızı bayraklarıdır.

Hile alanında veri madenciliği bazı anormallik kalıplarını tespit emek için verilerin elde edilme ve analiz etme işlemidir. Veri madenciliği büyük veri setlerinde istatistiksel ve matematiksel yöntemler ile veri setleri içerisinde ilişki kalıplarını tespit etmekte kullanılmakta ve bu noktadan yola çıkılarak hile tahmini yapılabilmektedir. Bu noktada belirtilmesi gereken önemli bir husus anormallik tespitlerinin usulsüz bir işleme işaret edebilme olasılığı kadar edememe olasılığına da sahip olduğu yani kesinlik taşımadığıdır. Diğer bir deyişle tespit edilen anormallik kalıbı hilenin bir ispatı değil, kırmızı bayrak olarak nitelendirilmelidir.

Veri madenciliğinde hile riskine en uygun veri seti veya alt grubu seçilir. Veriler temizlenir, birbiri ile olan ilişkileri tespit edilir ve dönüştürülür. Veri temizliğinde tespit edilmesi hedeflenen hile riskinden çok uzakta veya alakasız olan verilerin veri setinden çıkarılması ve düzenlemesi yapılır. Veri madenciliğinde iki tip çalışma sonucu alınabilmektedir. Bunlardan ilki veri temizliği sonrasında veri setindeki değişkenler arasında ilişkilerin tanımlanmasıdır. İkincisi ise veri setindeki ilişkilere dayanarak tahmin modellemesi yapılmasıdır.

Veri madenciliğinde izlenen başlıca yöntemler şunlardır,

1.2.3.1.1. Navie Bayes Yöntemi

Naive Bayes sınıflandırıcı algoritması bir veri setindeki sınıfları tespit etmek için kullanılan algoritmalardan biridir. Literatürde algoritmanın büyük veri setleri içerisinde hızlı ve iyi sonuçlar verdiği belirtilmekte ve yaygın olarak kullanılmaktadır.

Yöntemin temeli, Bayes istatistik teorisine dayanmaktadır. Bayes teoremi bir değişken için içinde bulunulan olasılık dağılımı içerisinde koşullu olasılık ile marjinal olasılık arasındaki ilişkiyi tanımlar. Bayes teoremi, P(c|x) in olma ihtimalini, P(c), P(x) ve P(x|c) den

hesaplar. Naive Bayes sınıflandırıcısında c sınıfındaki x lerin olma ihtimali, diğer ihtimallerden bağımsızdır. Bu duruma "sınıf durumundan bağımsızlık" denir. Bayes teoreminin matematiksel ifadesi aşağıdadır;

$$P(c|x) = \frac{P(x|c).P(c)}{P(x)} \tag{1.1}$$

$$P(c|x) = P(x_1|c).P(x_2|c) \cdots P(x_n|c).P(c) \tag{1.2}$$

- P(c|x) x sınıfında c nin olma ihtimalidir
- P(c) c nin olma ihtimalidir
- P(x|c) c sınıfında x in olma ihtimalidir
- P(x) x in olma ihtimalidir.

1.2.3.1.2. Destek Vektör Makinesi ve Tek Sınıflı Destek Vektör Makinesi

Destek vektör makinesi yönteminde iki sınıfı birbirinden ayırabilecek en uygun hiperdüzlem aranır. Bu yöntem yanlış sınıflandırma olasılığını en aza indirecek çözümler sunabilmektedir.

Tek sınıflı vektör makinesi bir sınıf veya bir hedef sınıfı diğer veriden ayırmakta kullanılır. Hedef sınıftaki veri noktaları gibi belirgin bir sınıf veriyi diğer veriden ayırmak için eğitilirler. Bu konuda literatürde iki Kernel fonksiyonu ile sonuçlanan iki farklı yaklaşım bulunmaktadır; Schölkopf ve Tax and Duin yaklaşımları.

Tek-sınıf destek vektör makinesi yöntemi diğer destek vektör makinesi yöntemleri arasında sık kullanılan bir yöntemdir ve bu yöntemde veri bir fonksiyon ile öznitelik uzayına taşınır, ardından bir hiperdüzlem vasıtasıyla başlangıç noktasından en uzağa taşınır. Bu durumda optimizasyon normal destek vektör makinesindekine benzer ikinci dereceden problemin çözümü ile sağlanır;

$$\min\left(1/2\min\left(\frac{1}{2\|w\|^2}+\frac{1}{vl\sum\xi_i}-\rho\right)\right. \quad (1.3)$$

$$\left(w * \Phi(x_i)\right) \geq \rho - \xi_i \; i = 1,2,\ldots,l \; ; \; \xi_i \geq 0 \quad (1.4)$$

Denklemde;

w ve ρ: hiperdüzlem parametreleri,

Φ: çekirdek fonksiyon,

v: hatalı sınıflandırılan değerlerin (aykırı değer) izin verilen oranı,

l: eğitim kümesindeki nesnelerin sayısı

ξ: hata parametresini ifade eder.

Transduktif Vektör Makineleri yöntemi önceden tanımlanmış bir gruba ait olan bir noktayı sınıflandırma kararında güvenilirlilik ölçülerini önermek için bu kategori altında veri madenciliğinde tercih edilen yöntemlerden biridir. Algoritmik rastgelelik teorisininin hesaplanması bu yöntemde kullanılmaktadır. Transduksiyon Güvenilirlik Makinelerinde güvenilirlik ölçüsü kullanılmaktadır. Kolmogorov tarafından geliştirilen rastgelelik teori algoritması transduktif güvenilirlik tahmin sürecinin temelidir.

Veri madenciliğinde kullanılan geleneksel yöntemlerin aksine transduksiyon bireysel noktaların güvenilirliğini sağlayan yeni ve henüz sınıflandırılmamış noktaların eğitimi gibi yöntemler sağlayabilir. Bu yöntemde noktanın belirli bir sınıfa dahil olup olmadığının testi p-değeri üzerinden yapılmaktadır. Boş hipotezin ilgili noktanın var olan sınıflandırmaya dahil olduğu durumda, p-değerinin küçük olması boş hipotezin reddedilme olasılığını yani ilgili noktanın aykırı değer olma olasılığını arttırmaktadır. Transduksiyon yöntemini güvenilirlik testi olarak kullananlar, tuhaflık ölçütü denilen bir p-değeri fonksiyonu kullanmaktadırlar (Mzila ve Dube, 2013).

1.2.3.1.3. Yapay Sinir Ağları Yöntemi

Bu yöntemde biyolojik sinir ağlarının yapısını ve fonksiyonunu örnek alınmaktadır. Beyindeki nöronlara benzer olarak YSA yönteminde çeşitli katmanlardan oluşan nötronlar bulunmaktadır. İleri beslemeli sinir ağları yöntemi popular bir metottur. YSA yönteminde üç katman bulunmaktadır;

a) Girdi Katmanı

b) Gizli Katman

c) Çıktı Katmanı

Girdi katmanı girilen verileri, çıktı katmanı verilerin dışarı çıkışını ifade eder. Modelde girdi katmanında bağımsız değişkenler, çıktı katmanında ise bağımlı değişkenler bulunmaktadır. İki katman arasında yer alan gizli katman girdi katmanından gelen verileri çıktı katmanına iletir.

Yapay Sinir Ağları yönteminde hedef ve gerçekleşen çıktı arasındaki hata oranına bağlı olarak nöron ağırlıklarını ayarlayan veri setlerini öğrenmek için eğitim algoritması kullanmaktadır. Genel olarak veri setlerinde öğrenmek için yapay sinir ağları yöntemi eğitim algoritmasını geriye dönük besleme algoritması olarak kullanmaktadır.

Çalışmalar geleneksel yöntemler ile karşılaştırıldığında sinir ağları modelinin hile tahmininde daha başarılı olduğunu göstermektedir (Fanning ve diğ., 1995; Green ve Choi, 1997).

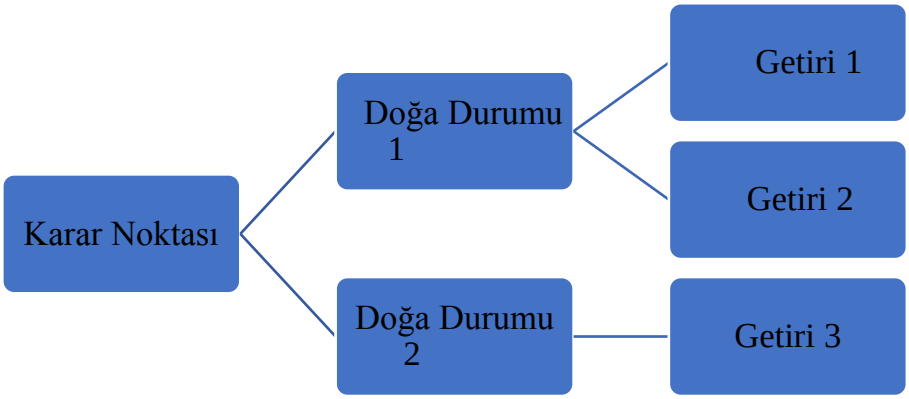
1.2.3.1.4. Karar Ağaçları ve Rastgele Orman Yöntemi

Veri madenciliğinde sık kullanılan yöntemlerden biri “Karar Ağacı” yöntemidir. Bu yöntemde hesaplanan çeşitli olasılıklar hesaplanmakta ve ağacın dallarını oluşturmaktadır. Bu yöntem ile, dallardan köke kadar istenilen yere ulaşılarak çeşitli olasılıklar hesaplanabilmektedir. Yöntem, büyük veri setlerini daha küçük setlere bölerek veri setlerinde tahmin edici ve tanımlayıcı özellikleri sağlamaktadır. Bu yöntemin yaygın olarak

kullanımının en önemli nedeni ağaç yapılarının oluşturulmasında kullanılan kuralların anlaşılabilir ve sade olmasıdır (Vardar, 2019).

Yukarıda da belirtildiği üzere karar ağaçları literatürde tercih edilen bir sınıflandırma ve örüntü tanımlama algoritmasıdır. Karar Ağaçları sınıflandırması işlemi çok aşamalı veya ardışık bir yapıdır. Bu mekanizma ekteki yapıya sahiptir.

Şekil 9 Karar Ağacının Yapısı



Rastgele Orman Yöntemi Karar Ağaçları Yönteminin bir çeşididir. Hile tahmin çalışmalarında sık olarak kullanılmaktadır. Bu yöntemde birçok sınıflandırma ağacı bulunmaktadır. Girdi verileri sınıflandırma ağaçlarının tamamından geçirilmektedir. Sınıflandırma ağaçları girdi verilerinin sınıflandırılması ile yapılmaktadır. Girdi veriler tüm sınıflandırma ağaçlarına girip oylanır. Ağaç yapılarından çıkan sonuçlara göre en fazla oyu alan sınıfa atama yapılır.

Sınıflandırmada ağaçlarının aşağıdaki basamakları izler:

- N sayıda eğitim verisi için orijinal veriden rastgele N sayıda veri toplanır.

• Eğer M tane girdi veri var ise, her bir düğüm için m<M olacak şekilde bir m sayısı belirlenir. m sayısı kadar veri M girdi verisinden rastgele seçilir ve bu düğümlerin en iyi şekilde ayrılması için kullanılır. Ağaç yapılarının oluşturulmasında bu m değerleri sabit kalır.

Her ağaç mümkün olan en büyük ölçüde gelişir. Budama yapılmaz. Ağaç yapıları gelişirken oluşabilecek orman hata oranı 2 şeye bağlıdır:

-Ormandaki iki ağaç arası korelasyon ve

-Ormandaki her bir ağacın dayanıklılığı

Oluşturulan ormandaki 2 ağaç arası korelasyonun artması orman hata oranın artması anlamına gelir. Her bir ağacın bireysel dayanıklılığının arttırılması orman hata oranını düşürür. m sayısının azaltılması korelasyon ve dayanıklılığın azalmasını, arttırılması ise artmasını sağlamaktadır. En uygun m aralığı genellikle oldukça geniştir.

Bu yöntemde, verilerin 2/3' ü eğitim, 1/3' ü de test amaçlı olarak ağaç yapılarını oluşturmak için kullanılır. Oransal dağılım kullanılan veri seti içerisine rastgele olarak gerçekleştirilir.

1.2.3.1.5. K-En Yakın Komşu Algoritması

Cover ve Hart (1967), K-NN algoritmasını geliştirmişlerdir. Bu algoritmada, örnek veri noktasının bulunduğu sınıfın ve en yakın komşunun, k değerine göre sınıflandırılmaktadır. Bu algoritma, bilinen, eski, basit ve etkili örüntü sınıflandırma bir örüntüleme yöntemidir ve yaygın olarak kullanılmaktadır. Taşcı ve Onan (2016) K-NN algoritmasının tercih edilme nedenleri arasında eğitiminin olmamasına rağmen gerçekleştirilmesinin kolay olmasıdır. Ek olarak analitik olarak izlenebilme, yerel bilgilere uyarlanabilme, paralel gerçekleştirilebilmekte ve uygun, gürültülü eğitim verilerine karşı dirençli olma gibi avantajlara sahiptir

K-NN algoritması, örnek tabanlı öğrenme algoritmalarındandır. Bu tip algoritmalarda öğrenme algoritmalarında, öğrenme işlemi eğitim setindeki verilere bağlı olarak gerçekleşir. Yeni karşılaşılan bir örnek, eğitim setinde yer alan örneklere göre sınıflandırılmaktadır. K-NN

algoritmasında, eğitim setinde yer alan örnekler n boyutlu sayısal niteliklerle sınıflandırılır Örnekler n boyutlu uzayda bir noktayı temsil edecek şekilde n boyutlu bir örnek uzayında bulunur.

Bilinmeyen bir örnek sisteme girdiğinde, eğitim setinden ilgili örneğe en yakın k tane örnek belirlenir. İkinci adımda, yeni örneğin sınıf etiketi, k en yakın komşusunun sınıf etiketlerinin çoğunluk oylamasına göre atanır. Bu yöntem ile K-NN algoritması, büyük eğitim setlerinde etkin sonuçlar verdiği görülmektedir.

K-NN algoritması basit bir yapıya olduğundan az sayıda parametre ile çalışmaktadır. K-NN algoritmasında, sınıf etiketinin çoğunluk oylamasına dayalı olarak belirlenmesi özellikle simetrik olmayan dağılıma sahip veri setlerinde sıklıkla görülen sınıfların, yeni örneklerin sınıf etiketlerinin belirlenmesinde daha baskın bir role sahip olmalarına yol açmaktadır. Bu soruna çözüm olarak, K-NN algoritmasının uzaklık ölçütünün etki değerine farklı şekillerde ağırlık değeri atayan yöntemlerle çalışılmaktadır.

K-NN algoritmasının bazı dezavantajları da bulunmaktadır. Bunlardan biri yüksek bir hesaplama maliyetidir. Sınıf etiketi belirlenmek istenen örneğin, veri setinde yer alan örnekler ile arasındaki uzaklığın belirlenmesi, büyük eğitim veri setlerinde yüksek maliyete yol açmaktadır. Soruna çözüm olarak, K-NN algoritması farklı boyut azaltma yöntemleri ile ya da daha güçlü veri yapıları ile birlikte kullanılmaktadır.

K-NN algoritmasının diğer bir dezavantajı, ilgisiz özniteliklerin varlığında sınıflandırma modeli oluşturmada kullanılmasında ortaya çıkmaktadır. Bu durumlarda eğitim için gereken süre uzundur. Ek olarak algoritma, yüksek miktarda bellek alanına ihtiyaç duymaktadır. Veri setinin büyüklüğü ve öznitelik boyutu arttıkça işlem yükü ve maliyetin yükselmektedir. Algoritmanın performansı k komşu sayısı, uzaklık ölçütü ve öznitelik sayısı gibi parametre ve özelliklere bağlı olarak değişkenlik göstermektedir. KNN algoritması, çok boyutlu veri setlerinde etkin değildir, yüksek bellek gereksinimi olması ve komşu sayısı, uzaklık ölçütü gibi parametrelere duyarlılığı nedeni ile tercih edilmemektedir (Taşcı ve Onan, 2016).

1.2.3.1.6. K-Ortalamalar Yöntemi

K-Ortalamalar kümeleme algoritması (K-Means clustering algorithm) MacQueen (1967) tarafından önerilmiştir ve bu grupta sık kullanılan algoritmalardan biridir. Bu yöntemde, kayıtlar önceden belirlenen küme sayısına göre gruplandırılır. Her biri tek kayıttan oluşan k adet küme ile analiz yapılır. Her yeni kayıt en yakın olan kümeye atanır. Kümeye yeni bir kayıt eklendikçe küme ortalaması tekrar hesaplanır. Tüm kayıtlar ilgili kümelere atanır. Tüm kayıtlar ilgili kümelere atandıktan sonra, atandıkları küme ortalamasından daha yakın küme ortalaması olur ise kayıtların yerleri tekrar değiştirilir.

K-Ortalamalar kümeleme algoritması büyük verileri işlemedeki başarısı ile veri madenciliğinde en yaygın kullanılan algoritmalardan biridir. K-Ortalamalar kümeleme algoritmasının uygulama aşamaları aşağıdaki şekildedir:

Adım 1. Başlangıç küme merkezlerini belirlemek için k tane merkez seçilir. Küme

Merkezleri rastgele ya da çeşitli metotlar yardımıyla seçilebilir.

Adım 2. Kaydın seçilen merkezlere uzaklığı ölçülür. Sonuca göre her bir kayıt k kümeden en yakın olan kümeye atanır.

Adım 3. Kümelerin merkezleri, kümedeki kayıtların ortalaması alınarak tekrar hesaplanır.

Adım 4. İkinci ve üçüncü adımlar küme merkezleri değişmeyene kadar tekrarlanır. (Üstünel, 2018).

Ortalama en yakın komşu yöntemi her ağırlık merkezi ile ona en yakın ağırlık merkezi arasındaki uzaklığı ölçer. Ardından en yakın komşu mesafelerinin ortalaması hesaplanır. Ortalama mesafenin varsayımsal rastgele dağılımın ortalamasından az olduğu durumlarda, özelliklerin dağılımı kümelenmiş olarak analiz edilir. Ortalama mesafenin varsayımsal rastgele dağılımın ortalamasından daha büyük olduğu durumlarda ise özellikler dağılmış olarak değerlendirilir. En yakın ortalama komşu oranı tespit edilmiş ortalama mesafenin beklenen ortalama mesafeye bölünmesi ile bulunur.

Ortalama En Yakın Komşu Oranı aşağıdaki formülle hesaplanmaktadır;

$$ANN = \frac{Do}{De} \tag{1.5}$$

Do= Her bir özellik ve en yakın komşusu arasında gözlemlenen ortalama mesafe

De= Rastgele dağılıma göre özelliklerin birbirleri arasındaki beklenen mesafe

$$Do = \sum_{i=1}^{n} \frac{di}{n} \tag{1.6}$$

$$De = \frac{0{,}5}{\sqrt{\frac{n}{A}}} \tag{1.7}$$

$$z = \frac{Do - De}{SE} \tag{1.8}$$

$$SE = \frac{0.26136}{\frac{\sqrt{n^2}}{A}} \tag{1.9}$$

Eğer endeks birden küçük ise eğilim kümelenme yönünde, birden büyük ise eğilim dağılım yönündedir. (ArcGisPro 2020:https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-statistics/h-how-average-nearest-neighbor-distance-spatial-st.htm).

1.2.3.1.7. Açı Bazlı Uç Değer Tespiti

Mesafelerin karşılaştırılması artan veri çeşitliliği ile giderek daha yetersiz hale gelmektedir. Veri içerisindeki şablonları belirleyebilmek için daha farklı yaklaşımlara ihtiyaç duyulmuştur. Bu yüzden sadece vektörler arasındaki mesafenin ölçülmesine değil, vektörlerin yönünün tayinine de ihtiyaç duyulmuştur. Vektörler arasındaki açıların karşılaştırılması birbirine benzer noktaların ve aykırı noktaların ayırt edilmesine yardımcı olmaktadır. Yöntemin yaklaşımı şu şekilde oluşturulmuştur. Açıların sapması kümelemenin sınırındaki noktalar için daha küçük olacaktır. Aykırı noktaların açı sapması içerideki noktalara göre daha az olmaktadır.

Bu değerlendirmelerin sonucunda açı-bazlı aykırı faktör tespit yöntemi nesnelerin birbirinden uzaklaşma yönlerini tanımlayabilmektedir. Eğer bir nokta için gözlemlenen açıların çeşitliliği fazla ise, noktanın her yöndeki noktalar tarafından çevrilme olasılığı yüksektir ve nokta kümenin içerisindedir. Eğer belirli bir noktaya ait tespit edilen açıların çeşitliliği az ise diğer noktaların belirli yönlerde olma olasılığı daha yüksektir. Bu durum ilgili noktanın birlikte gruplanan bir grup noktaya göre başka bir yerde olduğunu göstermektedir. Birbirlerine göre benzer açıları olduğu halde P noktasına göre göreceli olarak daha küçük açıların varlığı P'nin aykırı değer olduğunun göstergesi olarak kabul edilmektedir (Kriegel ve diğ., 2008).

Formülasyon olarak yöntem aşağıdaki şekilde ifade edilmektedir;

Verilen bir veritabanında $D \subseteq \vec{R}^d$, bir nokta $A \in \vec{D}$, ve bir norm $\|.\| : R^d \rightarrow \overrightarrow{R_0^+}$. İçsel çarpım $<.,.> : R^d \times R^d \rightarrow \vec{R}$ olarak gösterilmektedir. $\vec{B}, \vec{C} \in D$ noktaları için, BC fark

vektörü C – B'yi ifade eder. Açı bazlı uç değer ABOF(A) $\vec{A}$'nın tüm fark vektörleri ile D'deki tüm noktalara olan ağırlıklandırılmış mesafeye göre hesaplanan açılarının farkıdır.

$$ABOF(A) = \frac{VAR_{B,C \in D}(AB.AC)}{\|A.B\|^2.\|A.C\|^2} \quad (1.10)$$

1.2.3.1.8. Kümeleme Bazlı Yerel Uç Değer Faktör Tespiti

Kümeleme bazlı yerel aykırı faktör tespitinde veri seti içerisindeki yoğunluk alanını tespit edilmektedir. Bu yöntemde, her kümenin yoğunluğu tahmin edilmektedir. K-ortalama yöntemi veri setini kümelemek için kullanılır. Sonrasında bu yöntemde kümeler büyük ve küçük olarak iki sınıfa ayrılır. Aykırı değer hesaplaması her noktanın kümenin orta noktasına olan mesafesinin ölçülmesi ile hesaplanır. Küçük kümelemelerde büyük kümelemeye olan mesafe kullanılmaktadır. Bu yöntemin avantajı k-en yakın komşu yöntemi yerine kümeleme yöntemini kullanması iken yöntemin dezavantajı k-değerine karşı olan hassasiyetidir.

1.2.3.1.9. Özellik Sınıflandırma Yöntemi

Özellik sınıflandırma yöntemi yeni bir eğitim yöntemi olarak farkı veri Modellerini orijinal alt veri setlerine göre eğitir. Bazı durumlarda bazı özelliklerin az eğitimi söz konusu

olabilmektedir çünkü belirgin özellikler daha az belirgin özelliklere baskın olabilmektedir. Bu soruna çözüm olarak özellik sınıflandırma yönteminde özellik alt kümeleri sürekli olarak seçilmektedir her bir alt küme üzerinde farklı modeller eğitilmekte ve bireysel modellerin ortalaması alınmaktadır (Sutton ve diğ., 2005).

Histogram bazlı aykırı değer tespiti bir anormallik tespit algoritmasıdır. Bu yöntemle veri setindeki her bir özellik için tek değişkenli bir histogram oluşturulmaktadır. Bu yöntem özelliklerin birbirinden bağımsız olduğu varsayımına dayanır. Eğer kullanılan veri seri geniş dağılım sergileyen çok sayıda boyuta sahip bir veri seti ise özellikler arasındaki bağımsızlık varsayımı daha az önemli hale gelebilmektedir.

Histogramdaki her bir bölümün yüksekliği yoğunluk olasılığını temsil etmektedir. Her bir özelliğin hesaplamadaki eşit ağırlığını sağlamak için bölümlerin toplam uzunluğu 1 olacak şekilde standartlaştırılmıştır. Anormal vakalarının daha yüksek, normal vakaların daha düşük değeri olması için bu yöntemde hesaplanan değerlerin tersi alınmaktadır. Değerler ekteki formül ile hesaplanmaktadır. Formülde d özelliklerin sayısını, v özelliklere ait vektörü, $hist_i$ (v) her bir özelliğe ait yoğunluk tahminini vermektedir.

$$HBOS(v) = \sum_{i=0}^{d} \log\left(\frac{1}{hist_i(v)}\right) \tag{1.11}$$

Bu yöntemin üç avantajı bulunmaktadır; yüksek hesaplama hızı, puan bazlı tespit ve öğrenme aşamasının olmaması. Tespit edilecek anormallik sayısını bulabilmek için analiz edilecek veri miktarının büyük olduğu durumlarda bu yöntem ile daha iyi sonuç alınmaktadır.

Anormallik tespitinde, puanlama bazlı tespit yöntemlerinde, her bir noktaya bir aykırı değer puan verilmektedir. Bu durum ikili çıktı yöntemlerine göre bir avantajdır çünkü aykırı değer puanı tahminin güvenilirliğini ve kesinliğini tahmin etmekte faydalı olmaktadır. Ek olarak öğrenme gerektirmeyen bir yöntemdir. Yöntemde tek belirlenmesi gereken değer histogramdaki bölümlerin sayısıdır (Paulauskas ve Baskys, 2019).

1.2.3.1.10. Lokal Aykırı Değer Yöntemi

Lokal aykırı değer yöntemi yoğunluk-bazlı aykırı değer tespit yöntemlerinin bir türüdür. Bu yöntemin önemli bir sınırlaması yoğunluğu değişen aykırı değerlerin tespitidir. Yoğunluğun değişmesi yoğunluk bazlı yöntemlerde rastlanan bir problemdir. Lokal aykırı değer yöntemi ilgili nokta ve komşu noktaların yoğunluğunu birlikte hesaplamaya katar. Bu yöntemde noktalar için aykırı değer puanlaması yapılır böylece veri setindeki aykırı değerler tespit edilir. Yöntemin temel farklılığı veri noktasının göreceli yoğunluğunu hesaplamasıdır.

Yönteme göre göreceli yoğunluk ekteki formül ile hesaplanmaktadır;

$$X\ noktasının\ göreceki\ yoğunluğu = \frac{X\ noktasının\ yoğunluğu}{Tüm\ komşu\ noktaların\ ortalama\ yoğunluğu} \quad (1.12)$$

İlgili noktanın yoğunluğunun komşu noktaların yoğunluğundan düşük olması ilgili noktanın aykırı değer olduğuna işaret etmektedir (Kotu ve Deshpande, 2015).

1.2.3.1.11. Minimum Kovaryans Determinantı

Normal dışı veriyi tespit etmenin bir yolu normal verinin belirli bir dağılımı olduğunu varsayarak aykırı değerlerin düşük yoğunluğa sahip noktalar olduğunu varsaymaktır. Eliptik olarak dağılım gösteren veri setlerinde (Gauss dağılım gibi) hesaplama her bir noktanın ortalamaya olan uzaklığını gösteren Mahalanobis mesafesini hesaplayarak yapılmaktadır. Aykırı değerler belirli bir eşik değerinin üzerinde kalan değerlerdir. Mahalanobis mesafe hesaplamasında dağılımın varyans-kovaryans bilgileri kullanılmaktadır. Bu bilgiler bilinmediğinde veri setinden tahmin edilmeleri gerekmektedir.

Değişken tahminlerinde aykırı değerlerin varlığı hesaplamayı etkileyebilmektedir. Örneğin aykırı değerler veri setinin ortalamasını kendilerine doğru çekerek yapay olarak veri setinin kovaryans matrisini olduğundan büyük gösterilmesine neden olmaktadır. Aykırı değer önceden tespit edildiğinde bu değerler hesaplamadan çıkarılmaktadır, tahmin çalışması öncesinde genelde bu bilgi araştırmacıda bulunmamaktadır.

Minimum Kovaryans Determinant metodu varyans-kovaryans matrisinin tahmininde anormalliklerin etkisini en aza indirgemek için kullanılmaktadır. Örnek olarak belirli bir

büyüklükteki tüm olası veri alt setlerinin hesaplamaya alındığı varsayamında determinantı en küçük olan veri seti tahminleri seçilerek, seçilen kovaryans matrisi tutarlılık faktörü ile çarpılır.

Determinantı minimize etme fikrinin temeli dağılımın genişliğine dayanır. Bu yöntemde en dar dağılımlı alt veri seti tercih edilmektedir. Bu şekilde anormallikler daha kolay dışlanabilmektedir. Bu yöntem alt veri setlerinde sistemsel olarak hesaplama yapmayı kolaylaştırmaktadır. (Intuitive explanation of Minimum Covariance Determinant (MCD)2021: https://stats.stackexchange.com/questions/475636/intuitive-explanation-of-minimum-covariance-determinant-mcd).

1.2.3.1.12. Vaka Bazlı Tüme Varım Yöntemi

Vaka bazlı tümevarım yönteminde yeni problemlere çözüm olarak geçmişte yaşanan sorunların çözümünde kullanılan yöntemler baz alınır. Bu yöntemde dört temel aşama yer almaktadır;

1. Yeniden edinmek: Sorgudaki eski vakalar yeniden gözden geçirilir, eski vakaların çözümleri çıkarılır. Tüme varım yöntemleri ile sorgudaki vakaların ortak özellikleri çıkarılır ve sorguda herhangi bir yeniden öğrenme yapılmaz.

2. Yeniden kullanma: Eski vakaların en iyi çözümleri çıkarılır ve yeni vakaya uygulanır. Yeni çözüme yönelik olarak yeni genellemeler veya özelleştirmeler oluşabilir.

3. Gözden geçirme: Çözüm gözden geçirilir ve çözümün beklenen çıktıyı sağlayıp sağlayamadığına bakılır.

4. Yeni vakanın çözümü olarak oluşturulmuş yeni çözüm; hafızada güncellenmiş çıktılara ve korumanın ne şekilde yapılacağına dair belirlenmiş olan kurallara göre korunur veya korunmaz.

Bu yöntemde süreçte ne yapılacağı ve nasıl yapılacağı tariflenir. Bu yöntem aynı zamanda 4R yaklaşımı olarak da bilinmektedir.

1.2.3.1.13. Evrimsel Algoritmalar

Genetik algoritma Charles Darwinin doğal evrim teorisinden esinlenmiştir. Amaç daha güçlü bir yeni nesil üretmektir. Genetik algoritma evrimsel algoritma çeşitlerinden biridir. Doğal seleksiyon bazı bireylerin hayatta kalması ve soytürleri arasındaki farklar nedeni ile çoğalma sürecidir. Kalıtsal olarak taşınan özelliklerdeki bazı değişiklikler bir nesilden diğerine geçmektedir.

Darwine göre, tüm türler bir seçim sürecinden geçerek yaşarlar. Bireylerin hayatta kalma, savaşma ve üreme yeteneklerini arttıran küçük, kalıtımla kazanılmış değişimlerin doğal seçimleridir. Bu süreçte hayatta kalanlar doğal seleksiyon denen bir süreçten geçerler.

Aynı süreç evrimsel algoritmalar içinde geçerlidir. Doğal seleksiyon bir topluluktaki en güçlü olanların seçilmesi ile başlar. Bu bireyler kendi özelliklerini taşıyan yeni nesilleri yaratırlar. Böylece özellikler nesiller arasında aktarılır.

Tüm canlı topluluklarındaki değişimler vardır. Çünkü bireyin genlerinde değişim olabilmekte ve alt soylar değişimleri kalıtsal olarak devir alabilmektedir. Bireylerin hayatları boyunca, genleri çevreleri ile etkileşim halindedir. Süreç içerisinde belirli özelliklere sahip bireyler hayatta kalmakta ve topluluk evrimleşmektedir. Doğal seleksiyon süreci ile evrimsel algoritma benzerlik göstermektedir. Evrimsel algoritma dört aşamadan oluşmaktadır; başlatma, seçme, genetik uygulayıcılar ve son verme.

1.2.3.1.13.1. Topluluk ve Gen Seçimi

Doğal seleksiyon bazı soytürleri tercih eder. Uzun evrim dönemleri sonrası bu süreç özel biyolojik nişlere sahip toplulukların oluşmasıyla sonuçlanabilir. Topluluğun biyolojik evrim sürecinde doğal seleksiyon gerekli bir süreçtir.

Eğer ebeveynler güçlü ise, yavrular güçlü olacaktır ve hayatta kalma olasılıkları da daha yüksek olacaktır. Süreç tekrar ederek devam eder. En sonunda en güçlü bireylerden oluşan topluluğa erişilir.

Algoritmanın performansını maksimize eden faktörler kombinasyonuna ulaşmak hedeftir. İki durumda en iyi sonuç kabul edilebilir. İlk durumda maksimum tekrarla algoritma çalıştırılır, ikincisinde belirli bir performans seviyesine kadar algoritma çalıştırılır. Yazında başka yöntemlerde uygulanmaktadır.

Bir algoritmanın performansını objektif olarak ölçmek için genetik algoritmada uygunluk fonksiyonu kullanılır. Bu fonksiyon aynı zamanda hedef fonksiyonudur. Belirli bir hedefe ulaşmak için bir sonuca ne kadar yakın olduğumuzu özetler. Genetik algoritmada uygunluk fonksiyonları yavrunun en iyi çözüme olan mesafesini hesaplar.

Genetik algoritma araştırma bazlı bir optimizasyon tekniğidir. Bu yöntem ile normalde çözülmesi uzun süren problemlere en iyi çözümler bulunmaktadır. Rassal olarak yaratılmış her bir durum için olası çözümlerin havuzu veya popülasyonu oluşturulmaktadır. Her bir rassal oluşturulmuş fonksiyon için uygunluk fonksiyonu hesaplanmaktadır. İlk nesilde oluşan çözümlerin bazıları ikinci nesil için ebeveyn olmaktadır. Böylece topluluk içerisinde ebeveyn seçimi gerçekleştirilmektedir.

Genetik algoritmada aynı zamanda çaprazlama denilen bir çoğalma yöntemi bulunmaktadır. Aday ebeveyn havuzunda her bir aday uygunluk fonksiyonuna sokulmakta ve daha yüksek puana sahip olan adaylara eşleşme için daha yüksek olasılık verilmektedir. Bu sürece çaprazlama adı verilmektedir.

Genetik algoritmada başka uygulanan bir yöntem mutasyondur. Mutasyon popülasyondaki bazı bireylere uygulanan rastgele değişimdir. Çaprazlama ve mutasyon sonucu yeni çözümler oluşturulmaktadır. Her bir nesilde, her bir birey için uygunluk fonksiyonu hesaplanmaktadır. Hesaplamalar sonucunda süreç iki şekilde sonlandırılmaktadır; hedeflenen sayıdaki nesile ulaşıldığında veya önceden hedeflenen bir performans sınırına ulaşıldığında (Experian Credit Score, 2021: https://www.experian.co.uk/blogs/latest-thinkinge/fraud-prevention/5-of-the-most-remarkable-instances-in-the-history-of-fraud).

1.2.4. Hile Tespit Teknolojileri

Günümüzde hile tahmininde ağırlıklı olarak yukarıda açıklanan veri madenciliği metotları kullanılmaktadır. Bu metotların farklı kombinasyonları farklı sektörlerde hile

tespitinde kullanılmaktadır. Aşağıda literatürde kredi kartı hile analiz yaklaşımları açıklanmıştır. Diğer alanlardaki çalışmalar genel çerçevede belirtilmiştir.

1.2.4.1. Kredi Kartı Hile Analiz Yaklaşımları

Kredi kartı hareketlerine ilişkin hile tahmin çalışmaları kredi kartı veri tabanları üzerinden yapılmaktadır (Makki, 2019). Bu veri tabanlarında hesap numarası, kart tipi, alım tipi, lokasyon, müşteri ismi, ticari işletme kodu, yapılan işlemin tutarı gibi bilgiler yer almaktadır. Bu bilgiler araştırmacılar tarafından hileli işlemlerin tespitinde kullanılmaktadır. Bolton ve Hand (2002) çalışmalarında hile tespit araçlarını ve hile çalışmalarının yoğunlaştığı alanlardaki literatür hakkında bilgi vermişlerdir (Weston ve diğ., 2008).

Weston ve diğ. (2008) şüpheli işlemlerin ve aykırı değerlerin tespiti için eş grup analizini gerçek kredi kartı işlem veri seti üzerinde kullanmışlardır. Bu yöntem kullanılarak eş gruptan belirgin şekilde sapan işlemlerin varlığı tespit edilmiş ve bu işlemler kırmızı bayraklı işlemler olarak ayrıştırılmıştır. Zamanlama eşleşmesi, eş grupların kalitesi, işlemlere kırmızı bayraklar ile bağlamanın zamanı açıklanmıştır. (Ramakalyani ve Umadevi, 2012) genetik algoritma yöntemini hileli kart hareketlerinin tahmininde uygulamışlardır (Bentley ve diğ., 2000).

Bentley ve diğ. (2000) çalışmalarında şüpheli ve şüpheli olmayan kredi kartı işlemlerini sınıflandırmak için evrimsel bulanık sistem yöntemini uygulamışlardır (Srivastava ve diğ., 2008).

Srivastava ve diğ. (2008) yaptıkları çalışmada Kart sahibinin normal davranışı ile Gizli Markov Modelini eğitmişlerdir. Yeterince yüksek olasılıkla eğitilen model tarafından kabul edilmeyen kredi kartı işlemi hileli olarak kabul edilir. Aynı zamanda hileli olmayan işlemlerin modelleme ile ret edilmediğinden emin olmak istemişlerdir.

Esakkiraj ve Chidambaram (2013) çalışmalarında gizli Markov modelini kullanarak kullanıcı işleminin hileli veya normal olduğunu tespit eden bir modelleme yapmışlardır. Kullanılan yöntemde faydalanıcının son birkaç kredi kartı işlemine dayanarak sistem eğitilmekte ve her yeni işlem geçiş ve gözlem olasılığı ile ölçülmektedir. İşlemin gözlem olasılığına bağlı olarak ilgili işlem kabul edilmekte veya ret edilmektedir. Yapılan

modellemenin avantajı hileli işlemin gerçekleştikten sonra değil işlem sırasında tespitini sağlamasıdır ve böyle bir uyarı sisteme verildiğinde para transferi durdurulmaktadır.

Modellemeye göre eğer gelen bir kredi kartı işlemi yüksek olasılıkla ret ediliyorsa hileli işlem olarak kabul edilmektedir. Aynı zamanda hile olmayan normal işlemlerin ret edilmediği bu modellemede hileli işlemler garantilenmeye çalışılmaktadır. Diğer Markov modelini kullanan yazarların ifade ettiği gibi modelin avantajı hileli işlemleri gerçekleştikten sonra değil gerçekleşme aşamasında tespit etmesidir (Brabazon ve diğ., 2010).

Brabazon ve diğ. (2010) yaptıkları çalışmada yüksek tutarlı ödeme akışları günümüzde on-line olarak gerçekleştiğinden ve bankacılık sisteminin online işlemler üzerinden yürütüldüğünde kredi kartı hilelerine karşı verimli sistemlerin tasarlanması zorunluluğundan bahsetmektedirler. Hile tespit teknolojileri geliştikçe dolandırıcılarda aynı hızla sisteme adapte olmakta ve kendilerini güncellemektedir.

Bu yüzden geçmiş örneklere dayanan sistem öğrenmesi yeni hile şablonlarına karşı hile tespit sistemlerini tehlikeye karşı açıkta bırakmaktadır. Yapay bağışıklık sistemi adı verilen yöntemle bir takım standart dışı işlemlerde hesaplamaya dahil edilebilmektedir. Bu çalışmada kredi kartı ödemelerinin etkinliği online perakende işlem veri seti üzerinden yapay bağışıklık yöntemi ile test edilmiştir. Çalışmada üç yapay bağışıklık sistemi algoritması veri seti üzerinde uygulanmış ve sonuçlar lojistik regresyon modeli ile karşılaştırılmıştır. Sonuçlar yapay bağışıklık sisteminin hile tespit sistemlerine eklenmesinin bir potansiyeli olduğunu göstermekle birlikte bu alanda daha fazla çalışmaya ihtiyaç duyulduğunu göstermektedir.

Wong, Ray, Stephens ve Lewis (2012) çalışmalarında bazı biyolojik olayların karışık, gerçek hayat sorunlarını çözmek için ipuçları içermesinden yola çıkmışlardır. Birçok araştırmacı, sayısal zekâ ve onların uygulamalarının karışık problemlere uyarlanması için sinir ağları ve genetik algoritma teknikleri üzerinde çalışmaktadırlar. E-ticaret hizmetleri ve ağlarının gelişimindeki en önemli konulardan biri güvenlik yönetimi problemidir. Yakın zamandaki olaylar bilişim suçları faillerinin oldukça gelişmiş yöntemler kullandıklarını göstermektedir. Bu durum e-ticaret ve ağlarının güvenliğinin sağlanmasında geleneksel olan yöntemlerin dışına çıkılmasını gerektirmektedir. Biyolojik yöntemler bunlardan biridir. Yapay bağışıklık sistemi algoritması bu çalışmada araştırmacılar tarafından kullanılmıştır. Yapay

bağışıklık sistemi karışık doğal biyolojik saldırılardan insan hayatını kurtaran insan bağışıklık sistemini taklit etmektedir. Bu çalışmada kredi kartı dolandırıcılıklarının tespitinde yapay bağışıklık sistemi kullanılmaktadır. Bu teknik bu makalede güvenlik yönetiminin bir alanı olan kredi kartı yolsuzluklarına ilişkin bir vaka olarak çalışılmış iken bu yöntemin daha geniş bir mecrada e-ticaret alanında da uygulamasının mümkün olabileceği düşünülmektedir.

Sanchez ve diğ. (2009) hileli kart işlemlerini tespitte birleşme kurallarını uygulamışlardır. Birleşme kuralları veri madenciliğinde kullanılan en iyi yöntemlerden biri olarak kabul edilmektedir. Bu makalede araştırmacılar yasal olmayan işlemlerdeki davranış kalıbını elde etmek için gerekli veriyi veri setinden çıkarabilmek için birleşme kurallarını kullanmışlardır. Bu yöntem Şili'deki en önemli perakende firmalarından birine ait veri seti üzerinde denenmiştir.

Sahin ve Duman (2010) araştırmalarında bilgi teknolojilerindeki gelişmelere paralel olarak dolandırıcılığın dünyanın her yerinde yüksek finansal kayıplara yol açarak arttığına işaret etmişlerdir. Her ne kadar CHIP ve PIN gibi hile önleme mekanizmaları kredi kartı sistemlerinde kullanılsa da bu mekanizmalar sanal POS makinaları üzerinden gerçekleştirilen usulsüz işlemleri önlemekte yetersiz kalmaktadır. Bu çalışmalarında karar ağacı ve destek vektör makineleri yöntemlerini hile tespitinde kullanmışlardır. Çalışmaları gerçek bir veri seti üzerinde bu iki yöntemi karşılaştırmış ve çıktıları karar ağacı yönteminin daha iyi sonuç verdiği yönünde olmuştur.

Sahin ve diğ. (2013) online kredi kartı dolandırıcılığı ile ilgili diğer çalışmalarında, bu tip yolsuzların tespitinde araştırmacılar bu çalışmalarında yeni bir maliyet duyarlı bir karar ağacı yöntemi kullanmışlardır. Bu yöntem ile yanlış sınıflandırma maliyetlerini en aza indirilmesi için her terminal dışı düğümde bölümleme özelliği seçilmiştir. Çalışmada bu yöntemin performansı diğer bilinen geleneksel yöntemlerle gerçek kredi kartı veri seti üzerinden karşılaştırılmıştır. Sonuçlar araştırmacılar tarafından önerilen bu yöntemin kesinlik ve gerçek pozitiflik oranı gibi oranlar bazında ve araştırmacılar tarafından yeni tanımlanmış maliyet hassas ölçü bazında diğer yöntemlere göre iyi sonuçlar verdiği yönündedir.

Bahnsen, Stojanovic ve Aouada (2013) maliyet hassas karar ağacı yöntemini çalışmalarında kullanmışlardır. Günümüzde kullanılan kredi kartı algoritmaları kredi kartı

yolsuzluklarının gerçek maliyetini hesaplamada yetersiz kalmaktadır. Bu çalışmada dolandırıcılığın gerçek mali kayıp ve kazancını ölçen yeni bir karşılaştırma ölçütü üzerinde durulmaktadır. Araştırmacılar tarafından önerilen yeni ölçü; Bayes minimum riski üzerinden hesaplanan maliyet hassas bir yöntemle hesaplanmaktadır. Bu yöntem benzer algoritmalarla karşılaştırılmış ve maliyette %23 oranında bir iyileşme sağladığı görülmüştür. Çalışmada büyük bir Avrupalı kart işleme firmasının işlem veri seti kullanılımıştır.

Pasarica (2014) çalışmasında hileli işlemlerin gerçekleştirilmesinde gösterilen gelişmiş metotların varlığı, yüksek doğruluk oranı ile hileli işlemleri diğer işlemlerden ayıran gözetimli öğrenme algoritmasından elde edilen bir sınıflandırma yöntemine ihtiyaç duyulduğunu belirtmektedir. Bu çalışmada, gauss kernel fonksiyonu ile destek vektör makine sınıflandırması hile tahmininde önerilmektedir. Destek vektör makine yöntemi iki hipotez üzerine dayanmaktadır; marj optimizasyonu ve kernel gösterimi. Bu çalışmada doğrusal olmayan kernel kullanılan destek vektör yöntemi kullanılmıştır. Çalışmada, hile kalıplarını tespit etmek için en iyi yöntem olarak özellikler arasındaki benzerlikleri ölçen Gauss Kernel fonksiyonu önerilmektedir.

Ganji ve Mannem (2012) çalışmalarında bankaların ve finansal kuruluşların erken hileli işlem tespit sistemlerini bir süredir kullanmakta olduğunu ifade etmişlerdir. Ödeme sistemlerini yaşanabilir tutabilmek için gelişmiş hile tespit sistemleri gerekli hale gelmiştir. Veri madenciliğinde uç değer tespiti var olan algoritmaların önemli bir fonksiyonu haline gelmiştir. Bu yöntemler istatistiksel, uzaklık bazlı, yoğunluk bazlı, sapma bazlı yöntemler olarak bölümlenebilir. Bu çalışmada bir veri akışı algoritması kullanılarak kredi kartı hile tespiti yapılmıştır. Algoritma ters k en yakın komşu yöntemine dayalı olarak oluşturulmuştur. Algoritmanın avantajı veri setini sadece bir kez taramasının yeterli olmasıdır. Geleneksel yöntemlerin veri setini defalarca taradığı göz önüne alındığında bu yöntemin veri akışının olduğu bir ortama daha uygun olduğu araştırmacılarca öne sürülmektedir.

Maes ve diğ. (2002), makalelerinde makine öğrenmesi yöntemlerinin kredi kartı hile tespitinde kullanılmasını incelenmektedir. Çalışmada iki makine öğrenmesi yöntemi gerçek finansal veri üzerine uyarlanmıştır; yapay sinir ağları yöntemi ve Bayes inanç ağları yöntemi. İkinci yöntem ile makaledeki çalışmada daha iyi sonuçlar elde edildiği belirtilmektedir.

Whitrow ve diğ. (2009), çalışmalarında gözetimli hile sınıflandırması için işlem verisinin ön işleme tabi tutulması problemini incelenmektedirler. Verinin çok boyutlu ve işlemlerin heterojen olmasından dolayı bir hile tespit sisteminde işlemlerin tamamını kullanmak verimli olmamaktadır. Bu yüzden işlem birleştirmesi için bir çerçeve değerlendirilmiş ve bunun verimliliği işlem seviyesinde farklı sınıflandırma metotları ve gerçekçi maliyet bazlı performans ölçütü kullanılarak ölçülmüştür. Bu metotlar iki vaka bazında uygulanmıştır. İşlem birleştirmesinin tüm durumlarda değil ancak birçok durumda avantajlı olduğu görülmüştür. Birleştirme sürecinin uzunluğunun performans üzerinde önemli etkisi olduğu görülmüştür. Birleştirmenin rastgele orman yöntemi ile sınıflandırma yapıldığından daha etkili sonuçlar verdiği gözlemlenmiştir. Rastgele orman yönteminin sınıflandırma destek vektör makineleri, en yakın komşu, lojistik regresyon yöntemlerine göre daha iyi performans sergilediği görülmüştür. Birleştirme sınıflandırılmış veriye ihtiyaç duymaması ve topluluk sapmasının etkilerine karşı daha dayanıklı olma avantajlarına da sahiptir.

Mahmoudi ve Duman (2015) makalelerinde oldukça karışık algoritmalarla yazılan birçok gözetimli öğrenme yöntemlerinin literatürde yer aldığını belirtmektedirler. Karışık algoritmalı modeller her veri seti üzerinde iyi sonuçlar vermemektedir. Daha basit algoritmaların daha sağlıklı sonuçlar verdiği görülebilmektedir. Doğrusal ayrıştırıcı fonksiyonları daha az karışık sınıflandırıcılar olmalarına rağmen, kredi kartı hile tespiti gibi çok boyutlu problemlerde kullanılabilmektedir. Araştırmacılar bu çalışmalarında Fischer ayrıştırıcı fonksiyonunu kredi kartı hile tespitinde kullanmışlardır. Yanlış negatiflerin maliyetinin yanlış pozitiflerin maliyetinden çok daha fazla olmasından dolayı en önemli örneklere eğilimli sınıflandırma yöntemleri geliştirmenin önemini belirtmişlerdir. Değiştirilmiş Fischer ayırıcı fonksiyonu geleneksel fonksiyona göre önemli örneklere daha hassas olacak şekilde çalışmada uyarlanmıştır. Bu şekilde hileli olan/hileli olmayan sınıflandırmasından elde edilecek fayda maksimize edilmiştir.

1.2.4.2. Diğer Alanlardaki Çalışmalar

Hile ve usulsüzlükle ile ilgili kredi kartı hile tespiti dışında başka alanlarda da benzeri çalışmalar yapılmaktadır. Çalışmaların yoğunlaştığı alanlar; telekomünikasyon sektöründe hile tespiti, finansal tablolar hile analizi, hisse senedi hile tahmin yaklaşımları, sigorta

sektöründe hile analizi, kara para aklamayı önleme sistemleri, bilgisayar ihlal tahmin sistemleri, gerçek zamanlı hile tahmin sistemleri ve hile tahmininde makine öğrenmesi yaklaşımları olarak özetlenebilir (Makki, 2019).

1.2.4.3. Büyük Denetim Firmalarına Uygulanan Veri Madenciliği Uygulamaları

Akademik alanda hızla ilerleyen veri madenciliği çalışmalarına paralel olarak, denetim firmalarında denetim aracı olarak veri madenciliği araçlarını denetim sırasında kullanmaktadırlar. Aşağıda Dört büyük olarak da adlandırılan KPMG, PwC, Ernst&Young ve Deloitte denetim firmalarının kullandığı bazı veri madenciliği araçları açıklanmaktadır.

PwC analitik bir denetim araç olarak makine öğrenmesi ve yapay zekâ yöntemlerini kullanmaktadır. Örnek olarak Proaktif (vaka gerçekleşmeden önce) suistimal tespitini aşağıdaki basamakları takip ederek gerçekleştirmektedirler;

Veri toplama: Standart veya hileli işlemler, kurumsal BT sistemlerine kaydedilir. Her boyutta veri kümesinin tanımlanır, toplanır ve oluşturulur (PwC, 2021: https://www.pwc.com/us/en/services/consulting/cybersecurity-risk-regulatory/fraud-financial-crime.html).

Analitik Uygulamalar: Veriler temizlenir ve düzenlenir, işletmenin yapısına uygun bilgiler oluşturmak için özel senaryolar uygulanır.

Görselleştirme: Karar verme sürecine yardımcı olmak için verilerden anlamlı ve etkileşimli görseller oluşturulur.

Anomalilerin tespiti: Yüksek riskli ve şüpheli işlemlerin tanımlanması makine öğrenmesi ve yapay zekâ yöntemleri aracılığı gerçekleştirilir (PwC, 2021: https://www.pwc.com/veri-analitigi-ve-adli-bilisim-cozumleri).

Suistimalin tespit edilmesi durumunda alınabilecek önlemlerin ve eylem planının değerlendirilmesi ve iyileştirilmesi

Diğer taraftan, KPMG Kara Para Aklama ve Terörizmin Finansmanı tespitlerinde yapay zekâ ve makine öğrenmesi yöntemlerini kullanmaktadır. Müşterini Tanı uygulamaları olan K3PID'yı üçüncü taraf tarama maliyetini düşürmek amacı ile kullanırken ve makine öğrenmesi yöntemini işlem izleme sistemlerinin etkinliğini ve verimliliğini optimize etmekte kullanmaktadır. Böylelikle yanlış pozitif uyarıları azaltarak işlemi inceleme sürecini daha verimli hale getirilmektedir (KPMG, 2021: https://home.kpmg/id/en/home/services/advisory/forensic-services/anti-money-laundering-and-trade-sanctions-services.html).

Deloitte anormallik tespiti ve kural tabanlı yöntemleri, 20 yılı aşkın bir süredir dolandırıcılık, yolsuzluk ve suistimalle mücadelede yaygın olarak kullanmaktadır. Kullanılan dolandırıcılık analitiği, dolandırıcılık ve/veya rüşvete dayalı olanlar gibi olası uygunsuz işlemleri, işlemler tamamlanmadan önce veya gerçekleştikten sonra tespit etmeye yardımcı olmak için kullanılmaktadır. Dolandırıcılık analitiği süreci, ilgili verilerin toplanıp depolanmasını ve kalıp, tutarsızlık ve anormallikler tespiti için veri madenciliği sürecini içerir. Daha sonra bulgular, bir şirketin potansiyel tehditleri oluşmadan önce yönetmesine ve proaktif bir dolandırıcılık ve rüşvet tespit ortamı geliştirmesine olanak tanıyan kırmızı bayrak sistemine dönüştürülmektedir. Analitik araçlar kural tabanlı test yöntemlerini geliştirmenin yanı sıra, performansı ölçmek ve standardize etmek içinde kullanılmaktadır. Analitik teknolojisi tarafından yönlendirilen denetimsiz veya kural tabanlı olmayan analizler, geleneksel yaklaşımların gözden kaçırdığı gizli ve yeni kalıpları, eğilimleri, hileli şemaları ve senaryoları ortaya çıkarmakta kullanılmaktadır (Deloitte, 2022: https://www2.deloitte.com/tr/en/pages/deloitte-analytics/articles/fraud- analytics.html).

Ernst&Young usulsüzlük incelemeleri konusundaki verileri analiz ederek, sosyal medya, modelleme, metin madenciliği, ağ analizleri, makine öğrenmesi ve yapay zekâ modelleri kurgulayarak usulsüzlük, kaynakların gereksiz kullanımı, suistimal veya finansal suçları tespit edici veya önleyici denetim hizmeti vermektedir. Ek olarak kendi bünyesinde geliştirdiği yazılımlar veya farklı yazılım şirketleriyle geliştirdiği ortaklıklar çerçevesinde, şirketlerde kural bazlı veya makine öğrenmesi ve yapay zekâ modellerine dayalı suistimal tespiti yapılarını kurgulayarak ve uyarlayarak, anlık suistimal tespiti ve durdurucu altyapıları kullanmaktadır (EY, 2021: https://www.ey.com/tr_tr/forensic-integrity-services/adli-teknoloji-ve-kesif-hizmetleri).

BÖLÜM II.

MAKALE ÖZETİ VE VERİNİN TESTİ

Gerek akademik araştırmalarda ve gerek denetim firmalarının güncel uygulamalarında suistimal gerçekleşmeden önce oluşturulan kırmızı bayrak yöntemi yaygın yer almaktadır. Bu çerçevede bu bölümde kredi kartı hile analiz yaklaşımı ile genetik algoritma yöntemini baz alan bir kırmızı bayrak tespit uygulaması referans makale bazında incelenmiştir (Ramakalyani ve Umadevi, 2012).

Finansal kuruluşlar aracılığı ile gerçekleştirilen kredi kartı işlemleri arasında şüpheli olanların tespitinin sağlanması ve kırmızı bayrakların bazı işlemler için oluşturulması bazı yöntemler kullanılarak mümkün olabilmekte ve bu sayede geniş bir veri setinde olası hileli işlemler için bir ön uyarı mekanizması oluşturulabilmektedir. Bu bölümde kullanılacak olan genetik algoritma yöntemi, birinci bölümde de açıklanan kırmızı bayrakların oluşturulmasında kullanılan belli başlı yöntemlerden biri olan evrimsel algoritmalardandır.

Bu bölümde izlenecek olacak anlatım sıralaması ekteki şekildedir. İlk bölümde referans makalenin içeriği açıklanmaktadır. İkinci bölümde makalede kullanılan beş kuralın yazıldığı java kodu Matlab'a çevrilerek ulaşılan sonuçlar ve yapılan yöntem değişikliği anlatılmaktadır.

2.1. Genetik Algoritma Yönteminin Tercih Edilme Nedeni

Hile tespitinde birçok veri madenciliği yöntemi kullanılmaktadır. Genetik algoritma yöntemi kullanılan birçok yöntemden bir tanesidir. Genetik algoritmanın diğer yöntemlere göre bazı avantajları bulunmaktadır. Bu avantajları nedeni ile bu çalışmada tercih edilmiştir. Bu avantajlar ekteki şekilde sıralanabilir;

- Yöntemin göreceli olarak kullanımı diğer yöntemlere göre daha kolaydır ve birçok tekrara izin vermektedir.

- Genetik algoritma tanımlanan uygunluk fonksiyonu üzerinden çalışmaktadır. Diğer bir deyişle önceden tanımlanmış bir bilgiyi yerine uygunluk fonksiyonun çıktısı olan

veriyi kullanmaktadır. Genetik algoritmada belirleyici değil olasılıksal geçiş kuralları geçerli olması uygulamada esneklik sağlamaktadır (Goldberg, 1989).

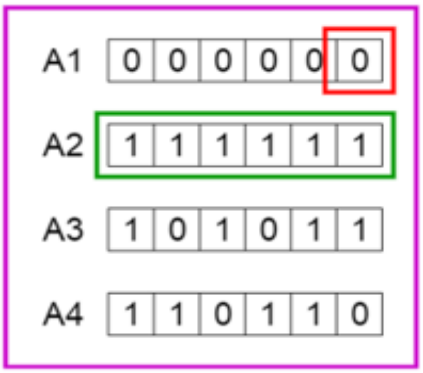
- Genetik algoritma aynı anda birçok sayı dizini paralel olarak aynı anda incelenebilir çünkü genetik algoritma, parametreler ile değil kodlanmış parametre setleri ile çalışmaktadır (Korucu, 2021: https://www.academia.edu/4715561/Genetik_Algoritmada_Tek_ve_Çok_Noktalı_Çaprazlama_Tekniklerinin_Doğrusal_Anten_Dizisi_Sentez_Probleminde_İncelenmesi).
  Bu sayede sürekli olmayan ve birçok yerel optimum noktaya sahip kompleks uygunluk fonksiyonuna sahip problemlerin çözümünde genetik algoritma iyi sonuçlar vermektedir (Buckles ve Petry, 1992). Diğer bir deyişle genetik algoritma tek bir noktadan değil, popülasyonun tanımlamış birçok noktasından arama yapmaktadır. Bu sayede genetik algoritma geleneksel optimizasyon yöntemleri için oldukça zor olarak kabul edilen çok değişkenli optimizasyon problemlerinin çözümümde tercih edilmektedir.

2.2. Genetik Algoritma Uyarlaması

2.2.1. Genetik Algoritmanın Çalışma Mantığı

Genetik algoritma Charles Darwin'in doğal seleksiyon teorisinden esinlenmiştir. Bu yöntemde gelecek nesil için en uygun adaylar doğal seleksiyon yolu ile seçilmektedir (Mallawaarachci, 2021). Genetik algoritma yönteminin bazı temel kavramları bulunmaktadır. Bu kavramlar izleyen tablolarda görselleştirilmiştir.

Şekil 10 Temel Genetik Algoritma Kavramları



Kaynak: Introduction to Genetic Algorithms — Including Example Code, Towards Data Science https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3

A1: Gen

A2: Kromozom

A3-A4: Topluluk

Doğal Seleksiyon süreci en uygun adayların seçilmesi ile başlar ve süreç sırasında ebeveynlerin özelliklerini genetik olarak alan yavrular üretilir ve yeni nesile dahil edilir. Ebeveynlerin daha formda olma durumlarında yavrular ebeveynlerinden daha iyi olacaktır ve hayatta kalma olasılıkları daha yüksek olacaktır. Bu süreç tekrar eder ve süreç sonunda en formda olan bireylerin olduğu nesil bulunur. Bu işleyiş şekli farklı araştırma konularına da uyarlanabilir. Bir problemin birden fazla çözümü bulunabilir ve içlerinden en iyisi genetik algoritma metodu ile seçilir. Genetik algoritma uygulamada beş basamakta uygulanmaktadır;

1. İlk topluluk

2. Uygunluk fonksiyonu

3. Seçim

4. Çaprazlama

5. Mutasyon

2.2.1.1. İlk Topluluk

Uygulama ilk topluluğun seçimi ile başlamaktadır. Bir grup birey topluluğu oluşturur. Problem çözümlerinde topluluk sorulan sorunun çözümünü içeren alternatifler serisidir. Her bir birey gen serisi ile tanımlanmaktadır. Gen serisi bir seri değişkenden oluşmaktadır. Genlerden bir dizi oluşturularak bir kromozom(çözüm) oluşturulmaktadır. Genetik algoritmada bir bireyin genleri bir dizi oluşturarak tanımlanmaktadır. Tanımlamada iki elemanlı değerler kullanılmaktadır (1 ve 0). Bu duruma genlerin kromozom içerisine kodlaması denilmektedir.

2.2.1.2. Uygunluk Fonksiyonu

Uygunluk fonksiyonu bir bireyin ne kadar uygun olduğunu tanımlayan kurallar dizisidir. Bir bireyin diğer bireyler ile rekabet yeteneğini tanımlar. Uygunluk fonksiyonu ile her bir birey için uygunluk puanı hesaplanır. Bir bireyin yeniden eşleşme için seçimi bireyler arası uygunluk puanı sıralamasına göre belirlenir.

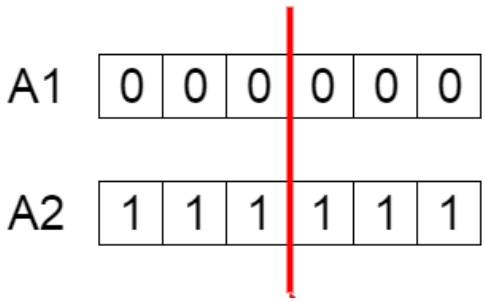
2.2.1.3. Seçim

Seçimde temel olan en uygun bireylerin seçilerek bir sonraki jenerasyona genlerinin iletilmesidir. İki birey (ebeveynler) uygunluk puanlarına göre seçilir. Daha yüksek uygunluk puanı olan bireylerin eşleşme için seçilme olasılıkları daha yüksektir.

2.2.1.4. Çaprazlama

Çaprazlama genetik algoritmadaki en belirgin safhadır. Her bir çift ebeveynin eşleşmesi için çaprazlama noktası genler arasından rastlantısal olarak seçilir. Ekte görsel olarak çaprazlamanın üçüncü noktadan başladığı bir örnek paylaşılmaktadır;

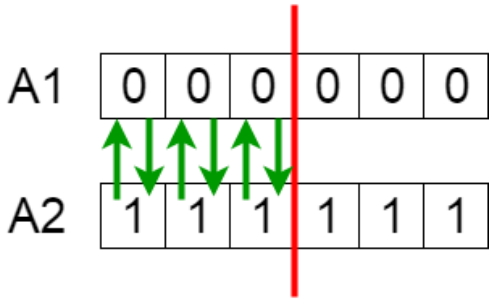
Şekil 11 Çaprazlama Noktası



Kaynak: Introduction to Genetic Algorithms — Including Example Code, Towards Data Science. https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3

Yavrular çaprazlama noktasına ulaşıncaya kadar ebeveynlerin genlerinin değiş tokuş edilmesi süreci sonunda oluşturulmaktadır.

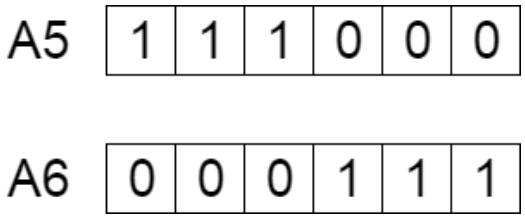
Şekil 12 Genlerin ebeveynler arasında değiş tokuş edilmesi



Kaynak: Introduction to Genetic Algorithms — Including Example Code, Towards Data Science https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3

Sonrasında yeni yavrular topluluğa dahil edilmektedir.

Şekil 13 Yeni Yavruların Oluşması

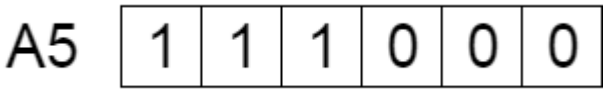


Kaynak: Introduction to Genetic Algorithms — Including Example Code, Towards Data Science. https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3

2.2.1.5. Mutasyon

Yeni yavrular oluştuğunda bazılarının genleri düşük olasılık ile mutasyona tabi kalır. Bu yöntem ile bit dizinindeki bazı bitlerin yeri değişebilmektedir.

Şekil 14 Mutasyon Öncesi



Kaynak: Introduction to Genetic Algorithms — Including Example Code, Towards Data Science. https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3

Mutasyondan sonra;

Şekil 15 Mutasyon Sonrası

Kaynak: Introduction to Genetic Algorithms — Including Example Code, Towards Data Science.https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3

Mutasyon topluluk içerisindeki çeşitliliği koruyabilmek ve zamanından önce yakınsamayı önlemek amacı ile uygulanmaktadır.

Algoritma topluluk yakınsadığında sonlanmaktadır. Önceki nesilden belirgin olarak farklı olan nesiller oluşturulamadığında ve probleme çözüm üretildiği zaman genetik algoritma sonlanmaktadır.

Toplulukların sabit bir büyüklüğü bulunmaktadır. Yeni kuşaklar oluşturulduğunda, en düşük uygunluk skoru olan bireyler elenmekte ve yeni yavrulara yer açılmaktadır. Süreç bir önceki kuşaktan daha iyi bireylere sahip kuşak yaratılıncaya kadar tekrar devam etmektedir. Topluluk yakınsayana, en iyi çözüm bulunana kadar işlem tekrar etmektedir.

2.3. Genetik Algoritma Uygulaması

Genetik algoritma en iyi sonuç bulunana kadar devam eden ve önceden belirlenmiş bir kuşak sayısınca tekrar edilir (Ramakalyani ve Umadevi, 2012). Daha iyi bir performans almak amacı ile bir problem üzerinden işletilen parametrik olarak işletilir. Referans makalede, parametreler ve genetik algoritmanın kurulumu hileli bir işlem oluşturmak amacı ile kurgulanmıştır.

Bu parametrelerin neler olduğu ilerleyen bölümlerde detaylandırılmıştır. Parametreler üzerinden kritik değerler hesaplanmaktadır. Çalışma dört safhaya ayrılabilir;

Birinci Safha: Kredi kartı işlemlerinin verisi içeri alınır, her işlemin n sayıda özelliği bulunmaktadır ve veri standardize edilir.

İkinci Safha: Kritik değerler hesaplanır, kritik değerler oluşturulan kurallar için hesaplanır. Kural sayısı beş adettir.

Üçüncü Safha: Kritik değerler belirli bir kuşak sayısından sonra bulunur. Kritik değerler; kritik, izlenebilir ve olağan olarak üç grupta kategorize edilir.

Dördüncü Safha: Dördüncü safhada, algoritma kullanılarak hileli işlemler yaratılır. Bu basamak genetik algoritmanın uygulanabilirliğini analiz etmek için tasarlanmıştır. Bu basamakta, hileli kredi kartı işlemlerinin tespiti için kritik değerlerin hesaplanmasını baz alınan hile tespit yöntemi ve prosedürleri önerilir.

İlk bölümde de açıklandığı üzere genetik algoritma basamaklar halinde uygulanmaktadır. Öncelikle başlangıç topluluğu diğer topluluklar arasından rastgele seçilir. Ardından her bir topluluk için uygunluk değeri hesaplanır ve topluluklar sıralanır. Yüksek uygunluk değerleri ve en iyi bireyleri seçmek için seçim yapılır. İlk veri yaratıldıktan sonra turnuva yöntemi ile uygunluk puanları hesaplanmaktadır. Yeniden bireyleri içeren eşleşme havuzu oluşturulur ve var olan topluluğun seçim süreci gerçekleşir. Ardından her bir topluluk için uygunluk değeri hesaplanır ve topluluklar sıralanır.

Şüpheli seçim, rastgele seçim, rulet tekerlek seçimi, turnuva seçimi gibi birçok seçim yöntemi bulunmaktadır Turnuva yöntemi farklı gruplardan en iyi bireyleri seçtiğinden referans makalede tercih edilmiştir. Rastgele olarak var olan topluluktan t sayıda birey seçilmektedir ve bir turnuva düzenlenmektedir. Gruptaki en iyi birey turnuvayı kazanmaktadır. Turnuva yöntemi ile maksimum sayıda turnuva sayısına ulaşılmakta ve yeniden eşleşme için eşleşme havuzuna konulmaktadır. Bu süreç ortalama topluluk boyutuna erişilinceye kadar birçok kez tekrar edilmektedir. Turnuvanın büyüklüğü seçimin gücünü belirleyen önemli bir faktördür. Turnuva büyüdükçe, seçim süreci güçlenmektedir.

Daha sonra tek nokta yöntemi ile çaprazlama yapılmakta ve elitist yöntemi ile mutasyon uygulanmaktadır. Genetik operatörler (çaprazlama ve mutasyon) aracılığı ile ikinci bir çözüm topluğu oluşturulmaktadır. Bu süreçler ilk topluluktan farklı gelecek nesil topluluklarının oluşturmak için kullanılmaktadır. Elitist seçim yöntemi ile bir sonraki nesile geçen en iyi bireyler seçilmektedir. Topluluğun bir sonraki nesline geçen bireyler en yüksek uygunluk değerine göre seçilir. Yeni topluluk yaratılır ve maksimum sayıda topluluğa ulaşıncaya kadar

aynı süreç devam eder. Bu yöntem ile ortalama uygunluk puanının arttığı görülmektedir çünkü bir sonraki nesile sadece en iyi bireyler geçmektedir.

Çıkış sonucuna erişilinceye kadar yaratım sürecine devam edilmektedir. Bilinen bazı ortak çıkış şartları şunlardır;

● Minimum kriterleri sağlayan çözüm,

● Belirli bir sayıda nesile ulaşmak

● Tahsis edilen bütçeye ulaşmak

● Çıkış durumunda sıralanan en iyi uygunluk çözümüne ulaşılır ve ulaşılan platoda izleyen tekrarlar daha iyi sonuçlar üretmemektedir.

● Manuel tespitler

● Yukarıda belirtilen bulguların tamamı

2.4. Veri Seti ve Değişkenler

Referans makalede kullanılan veri setindeki değişkenler ektedir. Veri setinde 12 adet değişken kullanılmıştır;

CardID: Kart numarası

Auth: Kimlik numarası

Cur.BB: Kredi kartına ait güncel bakiye

CU: Kredi kartının toplam kullanılma sayısı

Avg.BB: Kredi kartına ait ortalama hesap bakiyesi

Times of Overdraft: Limit üzeri kredi kartından yapılan işlem sayısı

Credit card age: Kredi kartının gün bazında yaşı (kullanım süresi)

Credit Card Used Today: Bugün (veya seçilen gün) kredi kartı ile gerçekleştirilen işlem sayısı

Location: İlgili kredi kartın ait işlemlerinin gerçekleştiği konum (yer) sayısı

Location Today: Bugün (veya seçilen gün) gerçekleşen kredi kartı işlemlerinin konum (yer) sayısı

Overdraft Today: Bugün (veya seçilen gün) limit üstü işlemin gerçekleşmemesi veya gerçekleşmesi (0,1)

AmountSpentToday: Bugün (veya seçilen gün) kredi kartından gerçekleştirilen harcama tutarı

2.5. Veri Seti ve Hile Kuralları

Referans makalede veri setinin yapısına uygun olarak 5 adet hile kuralı oluşturulmuştur. Bu kurallar veri seti üzerinde genetik algoritma yöntemine tabi tutularak kurallar çerçevesinde şüpheli olan işlemler tespit edilmiştir.

2.5.1. Kural 1: Kredi Kartı Kullanma Sıklığı

Oluşturulan birinci kural belirli bir eşik üzerinde gerçekleşen işlem sıklığının şüpheli bir işlem olma olasılığına işaret eder.

$$Kredi\ Kartı\ Kullanma\ Sıklığı = \frac{Kredi\ Kartının\ Toplam\ Kullanım\ Sayısı}{Kredi\ Kartının\ Yaşı\ (Kullanım\ Süresi)} \qquad (2.1)$$

Varsayım 1: Eğer kredi kartının kullanılma sıklığı 0.2'den küçük ise, işlem şüpheli işlem değildir. Hile kuralı uygulanmaz. Bu durumda;

Kritik Değer=Kredi Kartı Kullanma Sıklığı

Varsayım 2: Eğer kredi kartının kullanılma sıklığı 0.2'den büyük ise, işlemin şüphelidir ve hile kuralı uygulanır.

Hileli işlemin gerçekleşme kuralı: Eğer, bugün (veya seçilen gün) kredi kartından gerçekleştirilen işlem sayısı kredi kartı kullanma sıklığının 5 katından büyük ise işlemin hileli işlem olma ihtimali bulunur ve kritik değer aşağıdaki şekilde hesaplanır;

Kritik değer=Bugün (veya seçilen gün) kredi kartı ile gerçekleştirilen işlem sayısı*Kredi kartı kullanma sıklığı

Eğer, bugün (veya seçilen gün) kredi kartından gerçekleştirilen işlem sayısı kredi kartı kullanma sıklığının 5 katından büyük değil ise işlemin hileli işlem olma olasılığı yoktur ve Kritik Değer= Kredi Kartı Kullanma Sıklığına eşittir.

2.5.2. Kural 2: Konuma Bağlı Kredi Kartı Kullanımı

İkinci kurala göre işlemler ile işlemlerin gerçekleştiği konum (yer) sayısı arasında ilişki vardır.

Hile Kuralı: Eğer kredi kartı işlemlerinin bugün (veya seçilen gün) gerçekleştiği konum (yer) sayısı 5'ten az ise, hile kuralı uygulanmaz

Kritik değer= 0,01

Eğer ilgili kredi kartından gerçekleştirilen işlemlerinin bugün (veya seçilen gün) gerçekleştiği konum (yer) sayısı 5'e eşit veya 5'ten büyük ise, kredi kartından gerçekleştirilen işlemlerin hileli olma olasılığı vardır.

$$Kritik\ Değer = \frac{loc}{CUT} \tag{2.2}$$

Bu durumda kritik değer kredi kartı işlemlerinin gerçekleştiği konum (yer) sayısının bugün (veya seçilen gün) kredi kartı ile gerçekleştirilen işlem sayısına bölünmesi ile hesaplanır.

2.5.3. Kural 3: Limit Üstünde İlgili Kredi Kartından Yapılan Toplam İşlem Sayısı

Referans makaledeki üçüncü kurala göre belirli bir limit üzerinde yapılan işlemlerin şüpheli işlemlere işaret edebilir.

Kredi kartının toplam kullanılma sayısına göre limit üzerinde gerçekleşen toplam işlem sayısı ekteki formül ile hesaplanmaktadır.

$$CU_{OD} = \frac{OD}{CU} \tag{2.3}$$

OD=Limit üzeri kredi kartından yapılan işlem sayısı

CU=Kredi kartının toplam kullanılma sayısı

CU_{od}=Kredi kartının toplam kullanılma sayısına göre limit üzerinde gerçekleşen toplam işlem sayısı

Varsayım 1: Kredi kartının toplam kullanılma sayısına göre limit üzerinde gerçekleşen toplam sayısı 0,02' den küçük ise şüpheli işlem olasılığı yoktur ve üçüncü hile kuralı uygulanmaz.

Hile kuralı: Kredi kartının toplam kullanılma sayısına göre limit üzerinde gerçekleşen toplam sayısı 0,02' ye eşit veya büyük ise işlem şüphelidir ve hile kuralı işletilir;

Kural olarak bugün (veya seçilen gün) limit üstü işlem gerçekleşip gerçekleşmediği kontrol edilir. Eğer gerçekleşmiş ise kritik değer bugün (veya seçilen gün) limit üstü gerçekleşmesi ile kredi kartının toplam kullanılma sayısına göre limit üzerinde gerçekleşen toplam sayısı çarpılarak hesaplanır

$$Kritik\ Değer = ODT * CU_{OD} \tag{2.4}$$

2.5.4. Kural 4: Kredi Kartı Bakiyesi Kuralı

Dördüncü kurala göre kredi kartına ait güncel ve ortalama bakiye oranlamasının belirli bir oran üzerinde olması şüpheli işleme işaret eder.

Varsayım 1: İki bakiyenin birbirine oranına defter bakiyesi adı verilir. Defter Bakiyesi ekteki iki bakiyenin birbirine oranıdır;

$$B_b = \frac{Güncel\ Bakiye}{Ortalama\ Bakiye} \quad (2.5)$$

$B_b \leq 0.25$ ise şüpheli işlem yoktur ve hile kuralı uygulanmaz.

Bu durumda Kritik değer= BB olarak kabul edilir.

Varsayım: $BB=B_b*2$'dir. (defter bakiyesinin iki katı)

Hile Kuralı: Eğer $B_b>0.25$ ise işlemlerin şüpheli olma olasılığı vardır ve hile kuralı işletilir;

Kritik değer= B_b* BB olarak hesaplanır

2.5.5. Kural 5: Günlük Harcama Kuralı

Beşinci kurala göre bugün (veya seçilen gün) kredi kartından gerçekleştirilen harcama tutarının gün bazındaki harcama tutarının üzerinde olması şüpheli işleme işaret eder.

Kural ekteki değişkenler kullanılarak hesaplanmaktadır;

Kredi kartının ay bazında yaşı (kullanım süresi) $=\frac{CC\ Age}{30}$

CC Age: Kredi kartının gün bazında yaşı

Avg.BB: Kredi kartının ortalama hesap bakiyesi

Kredi kartı limiti: 100.000 rupi (Makalede öngörülen limittir)

Bal=Kredi kartının cari bakiyesi

Bal= 100.000- Avg.BB

Tot= Kredi kartından yapılan aylık toplam harcama

Tot= Kredi kartının ay bazında kullanım süresi* Kredi kartının cari bakiyesi

Ds=Günlük harcama tutarı

Ds=Tot*kredi kartının ay bazında kullanım süresi

AmtT= Bugun (veya seçilen gün) kredi kartından gerçekleştirilen harcama tutarı

Hile Kuralı: Eğer Bugun (veya seçilen gün) kredi kartından gerçekleştirilen harcama tutarı günlük harcama tutarının 10 katına eşit veya büyük ise ise işlemin hileli işlem olma olasılığı vardır.

AmtT≥10*Ds 'dir.

Kritik değer: AmtT/(10*Ds) formulü ile hesaplanır.

Eğer Bugun (veya seçilen gün) kredi kartından gerçekleştirilen harcama tutarı günlük harcama tutarının 10 katından küçük ise;

AmtT<10*Ds ise;

Kritik değer :0,01

2.6. Referans Makale ve Matlab ile Ulaşılan Sonuçların Karşılaştırılması

Makalede çalıştırılan beş kural sonucunda ulaşılan şüpheli işlemler şunlardır,

Tablo 1 Şüpheli İşlemler Tablosu-1

	Kural1	Kural2	Kural3	Kural4	Kural5			
Kural İsmi/Kart Numarası	Kredi Kartı Kullanma Sıklığı	Konuma Bağlı Kredi Kartı Kullanımı	Limit üstünde Kredi Kartından Yapılan Toplam İşlem	Kredi Kartı Bakiyesi Kuralı	Günlük Harcama Kuralı	Kritik Değer	Kırmızı Bayrak sayısı	Kırmızı Bayrak Kategorisi
11111	0,10	0,01	0,31	0,33	0,01	0,00	0	
11112	0,15	0,01	0,50	0,45	0,01	0,00	0	
11113	0,19	0,01	0,14	0,33	0,01	0,14	1	
11114	0,26	0,79	0,32	1,67	0,01	0,00	0	
11115	1,21	0,43	0,20	0,49	0,01	**2,13**	**3**	**İzlenebilir**
11116	0,16	0,01	0,37	1,20	0,01	0,37	0	
11117	0,16	0,01	0,28	0,50	0,01	0,00	0	
11118	0,12	0,01	0,29	0,66	0,01	0,00	0	
11119	0,16	0,01	0,24	0,27	0,01	0,00	0	
11120	3,51	0,40	0,16	0,50	1,10	**5,17**	**4**	**Kritik**
11121	0,10	0,01	0,23	1,28	0,01	0,00	0	
11122	0,15	0,18	0,29	0,77	0,01	0,18	1	
11123	0,18	0,50	0,31	0,53	0,01	0,00	0	
11124	0,63	0,01	0,29	1,16	0,01	0,63	1	
11125	2,41	0,01	0,14	0,47	1,27	**4,29**	**4**	**Kritik**
11126	0,18	0,17	0,28	1,40	0,01	0,17	1	
11127	0,19	0,01	0,18	0,72	0,01	0,18	1	
11128	0,17	0,01	0,26	1,80	0,01	0,00	0	
11129	0,32	0,01	0,21	0,47	0,01	0,00	0	
11130	3,31	0,46	0,26	0,33	1,08	**4,85**	**3**	**Kritik**

İşletilen kurallar sonucu 1120,1130,1125 nolu kartların kırmızı bayrak kategorisinin kritik derecede olduğu tespit edilmiştir.1115 nolu kartta izlenebilir seviyede hileli işlem olasılığı bulunmaktadır.

Matlab'da genetik algoritma kodlaması sonucuna elde edilen sonuçlar ektedir, görüldüğü üzere Matlab kodlaması ile yazarın java kodlamasındaki çıktıları birebir aynıdır.

Tablo 2 Şüpheli İşlemler Tablosu-2

	Kural1	Kural2	Kural3	Kural4	Kural5			
Kural İsmi/ KartNumarası	**Kredi Kartı Kullanma Sıklığı**	**Konuma Bağlı Kredi Kartı Kullanımı**	**Limit üstünde Kredi Kartından Yapılan Toplam İşlem Sayısı**	**Kredi Kartı Bakiyesi Kuralı**	**Günlük Harcama Kuralı**	**Kritik Değer**	**Kırmızı Bayrak sayısı**	**Kırmızı Bayrak Kategorisi**
11111	0,10	0,01	0,31	0,33	0,01	0,00	0	
11112	0,15	0,01	0,50	0,45	0,01	0,00	0	
11113	0,19	0,01	0,14	0,33	0,01	0,14	1	
11114	0,26	0,01	0,32	1,67	0,01	0,00	0	
11115	1,21	0,43	0,20	0,49	0,01	**2,13**	**3**	**İzlenebilir**
11116	0,16	0,01	0,37	1,20	0,01	0,37	0	
11117	0,16	0,01	0,28	0,50	0,01	0,00	0	
11118	0,12	0,01	0,29	0,66	0,01	0,00	0	
11119	0,16	0,01	0,24	0,27	0,01	0,00	0	
11120	3,51	0,40	0,16	0,50	1,10	**5,17**	**4**	**Kritik**
11121	0,10	0,01	0,23	1,28	0,01	0,00	0	
11122	0,15	0,18	0,29	0,77	0,01	0,18	1	
11123	0,18	0,01	0,31	0,53	0,01	0,00	0	
11124	0,63	0,01	0,29	1,16	0,01	0,63	1	
11125	2,41	0,01	0,14	0,47	1,27	**4,29**	**4**	**Kritik**
11126	0,19	0,17	0,28	1,40	0,01	0,17	1	
11127	0,19	0,01	0,18	0,72	0,01	0,18	1	
11128	0,17	0,01	0,26	1,80	0,01	0,00	0	
11129	0,32	0,01	0,21	0,47	0,01	0,00	0	
11130	3,31	0,46	0,26	0,33	1,08	**4,85**	**3**	**Kritik**

2.7. Kodun Matlab'a Çevrilmesi Sonucunda Oluşan Kritik Değerlerin Karşılaştırılması

Kodun Matlab'a çevrilmesi sonucunda her bir işlem için hesaplanmış değerler aşağıda verilmiştir;

In CC ID : 11115-Usage CCfreq Fraud is found with value:1.208531

In CC ID : 11120-Usage CCfreq Fraud is found with value:3.513011

In CC ID : 11124-Usage CCfreq Fraud is found with value:0.632653

In CC ID : 11125-Usage CCfreq Fraud is found with value:2.405797

In CC ID : 11130-Usage CCfreq Fraud is found with value:3.313433

In CC ID : 11115-Usage Location Fraud is found with value:0.428571

In CC ID : 11120-Usage Location Fraud is found with value:0.400000

In CC ID : 11122-Usage Location Fraud is found with value:0.181818

In CC ID : 11126-Usage Location Fraud is found with value:0.166667

In CC ID : 11113-Usage Overdraft Fraud is found with value:0.142857

In CC ID : 11120-Usage Overdraft Fraud is found with value:0.158730

In CC ID : 11125-Usage Overdraft Fraud is found with value:0.144578

n CC ID : 11127-Usage Overdraft Fraud is found with value:0.176471

In CC ID : 11115-Usage BookBalance Fraud is found with value:0.491803

In CC ID : 11125-Usage BookBalance Fraud is found with value:0.472727

In CC ID : 11120-Usage Daily Spending Fraud is found with value:1.100000

In CC ID : 11125-Usage Daily Spending Fraud is found with value:1.266667

In CC ID : 11130-Usage Daily Spending Fraud is found with value:1.076923

pop =

0.4500	4.0000	0.1600	1.8000	0.1500
3.5130	0.4000	0.1600	0.5000	1.1000
3.5130	1.0000	0.1600	0.5000	0.1500
3.5130	1.0000	0.1600	0.5000	0.1500
0.1500	0.3333	0.1587	0.4000	0.1500
0.1500	4.0000	0.1587	0.4000	0.1500
0.1500	0.4000	0.1587	0.4000	1.1000
0.9000	0.4000	0.1587	0.4000	0.1500
0.4500	4.0000	0.1600	1.8000	0.1500
0.4500	4.0000	0.1600	1.8000	0.1500
0.1500	0.4000	0.1600	0.4000	0.1500
0.1500	0.1111	0.1600	0.4000	0.1500
0.1500	0.5000	0.1600	0.5000	1.1000
0.1500	0.5000	0.1600	0.5000	0.1500

0.7500 0.4000 0.1600 0.4000 0.1500

0.7500 0.2000 0.1600 0.4000 0.1500

0.7500 0.4000 0.1587 0.4000 0.1500

0.7500 0.4000 0.1587 0.4000 0.1500

0.3000 0.4000 0.1600 0.4000 0.1500

0.7500 0.4000 0.1600 0.4000 0.1500

fvalcriti =5.3230

fvalmonit =2.0087

fvalordin =1.6600

Ulaşılan çıktılar makale çıktıları ile uyumludur 1120,1125,1130 nolu kredi kartlarında kritik seviyede 1115 nolu kredi kartında izlenebilir seviyede hileli işlem olasılığına işaret etmektedir.

2.8. Alternatif Metod Çıktısı

Referans makalede kritik izlenebilir ve normal hile eşik değerleri hesaplanırken rastgele ve ve varsayım olarak 5,10,15. satırlardaki değerler olarak seçilmiştir. Alternatif bir metot olarak ortalamanın 2 sigma ilerisi kritik seviyede hileli işlem olarak ortalamanın bir sigma ilerisi gözlenebilir değer ve ortalama değer normal hile eşik değeri olarak kabul edilip veri setine uyarlandığında ekteki sonuçlara ulaşılmaktadır;

In CC ID : 11115-Usage CCfreq Fraud is found with value:1.208531

In CC ID : 11120-Usage CCfreq Fraud is found with value:3.513011

In CC ID : 11124-Usage CCfreq Fraud is found with value:0.632653

In CC ID : 11125-Usage CCfreq Fraud is found with value:2.405797

In CC ID : 11130-Usage CCfreq Fraud is found with value:3.313433

In CC ID : 11115-Usage Location Fraud is found with value:0.428571

In CC ID : 11120-Usage Location Fraud is found with value:0.400000

In CC ID : 11122-Usage Location Fraud is found with value:0.181818

In CC ID : 11126-Usage Location Fraud is found with value:0.166667

In CC ID : 11113-Usage Overdraft Fraud is found with value:0.142857

In CC ID : 11120-Usage Overdraft Fraud is found with value:0.158730

In CC ID : 11125-Usage Overdraft Fraud is found with value:0.144578

In CC ID : 11127-Usage Overdraft Fraud is found with value:0.176471

In CC ID : 11115-Usage BookBalance Fraud is found with value:0.491803

In CC ID : 11125-Usage BookBalance Fraud is found with value:0.472727

In CC ID : 11120-Usage Daily Spending Fraud is found with value:1.100000

In CC ID : 11125-Usage Daily Spending Fraud is found with value:1.266667

In CC ID : 11130-Usage Daily Spending Fraud is found with value:1.076923

pop =

3.5130	1.0000	0.1600	0.5000	1.1000
3.5130	1.0000	0.1587	0.5000	1.1000
3.5130	1.0000	0.1600	0.4000	1.0769
3.5130	0.2000	0.1600	0.4000	1.0769
0.7500	0.6000	0.1587	0.4000	0.1500
0.7500	0.5714	0.1587	0.4000	0.1500
3.5130	1.0000	0.1587	0.4000	0.1500
0.9000	1.0000	0.1587	0.4000	0.1500
3.5130	1.0000	0.1600	0.5000	1.1000
3.5130	0.3333	0.1600	0.4000	1.1000
0.4500	1.0000	0.1600	0.5000	1.1000
0.4500	1.0000	0.1600	0.5000	1.1000
0.9000	0.4000	0.1600	0.5000	1.1000
0.9000	0.4000	0.1600	0.5000	0.1500
0.9000	0.4000	0.1600	0.4000	0.1500
0.9000	0.3333	0.1600	0.4000	0.1500
3.5130	0.4000	0.1600	0.4000	0.1500

3.5130 0.4000 0.1600 0.4000 0.1500

0.9000 1.0000 0.1600 0.5000 1.1000

0.9000 4.0000 0.1600 0.4000 1.1000

fvalcriti =7.5787

fvalmonit =5.8582

fvalordin =4.1376

Alternatif metodun uygulanması sonucu elde edilen çıktılar referans makaledeki kritik değerlerden daha yüksektir.

BÖLÜM III.

GENETİK ALGORİTMANIN YENİ VERİ SETİNE UYARLANMASI

3.1. Yeni Veri Seti Özellikleri

Makalede kullanılan veri setinden sonra algoritma bu bölümde daha büyük ve yeni bir veri setine uyarlanmıştır. Buradaki amaç makalede uygulanan kurallar ve algoritmanın yeni ve daha büyük bir veri setine uyarlanabilirliğini test etmektir. Uyarlanan veri seti anonim bir kaynaktan alınmıştır (Fraud Detection Data Set, 2021: https://www.kaggle.com/alukosayoenoch/datascientist).

Veri setinin büyüklüğü 800 KB'in üzerindedir, veri setinde aşağıda detayları açıklanan 9 adet değişken bulunmaktadır;

Tablo 3 Değişken Tablosu

Değişken Adı	**Değişken Açıklaması**
accountID	Hesap Numarası
transactionAmount	İşlem Tutarı
transactionCurrencyCode	İşlem Para Birim Kodu
transactionDate	İşlem Tarihi
transactionTime	İşlem Saati
localHour	Yerel Saat
transactionDeviceId	İşlemin Yapıldığı Cihaz Numarası
transactionIPaddress	İşlemin Yapıldığı IP adresi

Veri setinin seçiminde referans makaledeki veri setine benzer yapıda bir veri seti olması dikkate alınmıştır. Bu durum hem algoritmanın daha büyük veri setine uyarlandığı verdiği çıktıları analizinde hem de benzer kuralların veri setine uygulanmasında kolaylık sağlamıştır. Veri seti ekteki dönüşüm tablosunda özetlendiği şekilde makalenin veri formatına çevrilmiştir.

Tablo 4 Değişken Dönüşüm Tablosu

Verinin Orijinal Kolon Sıralaması	**Verinin Dönüştürülmüş Kolon Sıralaması**	**Referans Makale Değişken Kolon Sıralaması**
transactionID	index	CardID
accountID	authid	Auth
transactionAmount	max_amt	Cur.BB
transactionCurrencyCode	avgamt	CU
transactionDate	odn	Avg.BB
transactionTime	age	OD
localHour	transnT	CCAge
transactionDeviceId	loc	CUT
transactionIPaddress	locT	Loc
	odnT	LocT
	amtT	ODT
	maxsamt	AmtT

3.2. Yeni Kural ve Özellikleri

Yeni kurallar ikinci bölümde referans makalede oluşturulmuş beş adet kuralın yeni veri setine ekteki şekilde uyarlanması ile oluşturulmuştur.

3.2.1. Kural 1

Birinci kural işlem sıklığının şüpheli işleme işaret etme olasılığını gösterir. Kredi kartının kullanılma süresi ya da işlem yapılan tarihler arasındaki farkın hesaplanması ile bulunur.

age:(maxdate-mindate) işlem yapılan tarih aralığı hesaplanır.

transn: Her bir hesaba ait yapılan toplam işlem sayısı hesaplanır;

transnT: En çok işlem yapılan gündeki toplam işlem sayısı

Her bir hesap için için işlem sıklığı hesaplanır;

$$ccFreq = \frac{transn}{age} \quad (3.1)$$

Eğer hesapta gerçekleşen işlem sıklığı 4'e küçük veya eşit ise, işlem şüpheli işlem değildir ve hile kuralı işletilmez. Bu durumda kritik değer işlem sıklığına eşittir.

Eğer ccFreq≤4;

Kritik değer: ccFreq

Eğer hesapta gerçekleşen işlem sıklığı 4'ten büyük ise hesapta gerçekleşen işlemlerin şüpheli olma ihtimali bulunmaktadır ve kritik değer hesapta en çok işlem yapılan gündeki işlem sayısı ve toplam işlem sayısının çarpımı ile bulunur.

Eğer ccFreq>4;

transT>ccFreq

TransT*trans

Referans makaleye göre yeni veri setine uyarlama yapılırken şu değişiklikler yapılmıştır;

- Kredi kartının kullanım sıklığı toplam kullanılan kart sayısının kredi kartının yaşına bölünmesi ise hesaplanmaktadır.

- Kredi kartının yaşı yeni veri setinde her bir hesapta en yüksek tarihli günden en düşük tarihli günün çıkarılması olarak değiştirilmiştir.

- Kredi Kartının toplam kullanılma sayısı ilgili hesapta yapılan toplam işlem sayısı olarak değiştirilmiştir.

- Önceki veride kullanılan kredi kartı işlemi sayısı olan CU, Transn her bir hesaptaki tüm işlemlerin sayısı olarak yeni veri setine uyarlanmıştır.

- CUT bugün veya seçilen gündeki kredi kartı kullanum sayısını verirken yeni veri setine TransT olarak uyarlanmıştır. TransT yeni veri setinde en çok işlem yapılan gün olarak seçilmiş ve yeni veri setine uyarlanmıştır.

Referans değerler yeni veri setinin yapısına göre 0,2 yerine 4 olarak değiştirilmiş ve 5 olan katsayı sıfırlanmıştır.

3.2.2. Kural 2

Kurala göre belirli bir gün içerisinde işlemlerin gerçekleştiği lokasyonların sayısının eşik değerini aşması şüpheli işleme işaret eder.

Loc: Her bir hesaba ait işlemlerin gerçekleştiği lokasyon sayısı yeni veri setinde Transaction IP address (İşlemim gerçekleştiği IPadresi) değişkeni ile ilişkilendirilmiştir.

Loc: Hesaba ait işlemlerin gerçekleştiği lokasyon sayısı

LocT: Hesaba ait en çok işlem gerçekleşen günde hesaba ait gerçekleşen işlemlerin lokasyon sayısı

Kurala göre, işlemlerin gerçekleştiği lokasyon sayısı 2'ten küçük veya eşit ise lokasyon çeşitliliği şüpheli işleme işaret etmez ve kritik değer 0,01 olarak varsayılır.

Eğer Loc≤2; kritik değer: 0,01

Eğer Loc>2;

LocT>0.5*Loc

Kritik değer =LocT*loc;

Eğer işlemlerin gerçekleştiği toplam lokasyon sayısı 2'den büyük ise ve en çok işlem görülen günde gerçekleşen işlemlerin lokasyon sayısı hesaba ait lokasyon sayısının 0,5 katından büyük ise şüpheli işlemdir ve kritik değer hesaba ait en çok işlem yapılan gündeki lokasyon sayısı ile toplam lokasyon sayısının çarpımıdır.

Yeni veri setinde kritik değer hesaplaması en çok işlem yapılan gündeki lokasyon sayısı ile toplam lokasyon sayısının çarpımı olarak değiştirilmiş ve aynı zamanda referans makaleye göre katsayılar değiştirilmiştir.

3.2.3. Kural 3

Üçüncü kurala göre belirli bir limit üzerinde gerçekleşen dikkat çekici işlemlerin sayısı şüpheli işlemlere işaret eder.

OD: Her hesapta limit tutarının üzerinde gerçekleşen işlem adedi

CU: Her hesapta gerçekleşen toplam işlem adedi

ODT: En çok işlem yapılan günde hesapta gerçekleşen işlemlerin limit tutarının üzerinde olması veya olmaması

(0,1) = 0; limit üzerinde değil

1; limit üzerinde

Hesapta gerçekleşen limit üstü işlem sayısının toplam işlem sayısına oranı;

$$\text{Eğer } \frac{OD}{CU} \leq 2 \tag{3.2}$$

Kritik değer;

$$\frac{OD}{CU} \tag{3.3}$$

Hesapta gerçekleşen limit üstü işlem sayısının toplam işlem sayısına oranı;

$$\text{Eğer } \frac{OD}{CU} \geq 2 \tag{3.4}$$

Kritik değer;

$$ODT * \frac{OD}{CU} \tag{3.5}$$

Yeni veri setinde eski veri setine göre şu değişiklikler yapılmıştır;

- Referans makale veri setinde kullanılan kredi kartı işlemi sayısı olan CU değişkeni Transn değişkeni ile değiştirilmiş ve Transn her bir hesaptaki tüm işlemlerin sayısı kuralı olarak yeni veri setine uyarlanmıştır.

- Her hesaba ilişkin limit üstü işlem sayısı hesaplanmış ve kural olarak veri setinde her hesaba ait en çok işlem yapılan gün seçilmiştir.

- Limit üstü işlem tutarı her bir hesabın ortalama işlem tutarının iki standart sapma oranında üzeri olarak varsayılmıştır.

- Veri setinin özelliklerine göre katsayı ve karşılaştırma oranları değiştirilmiştir.

3.2.4. Kural 4

Dördüncü kurala göre maksimum işlem tutarının toplam işlem tutarına oranının belirli bir eşik değerin üzerinde olması şüpheli işlemlere işaret eder.

CurrentBB: Her bir hesap için maksimum işlem tutarı

AverageBB: Her bir hesap için işlem tutar ortalaması

$$bb = \frac{Current\ BB}{Average\ BB} \quad (3.6)$$

BB=bb*2

Eğer bb>2;

İşlemin şüpheli olma olasılığı bulunmaktadır ve kritik değer bb'nin iki katı olan BB'ye eşittir.

Eğer bb≤2;

İşlem kurala göre şüpheli işlem değildir ve kritik değer bb'ye eşittir.

Yeni veri setine uyarlama aşamasında yapılan değişiklikler şunlardır;

- İlk veri setinde Curremt BB güncel bakiyeye eşit iken yeni veri setine her hesap için en çok işlem gerçekleşen günde gerçekleşen maksimum işlem tutarı olarak uyarlanmıştır.

- Average BB ilk veri setinde ortalama bakiyeye eşit iken bu veri setinde her hesapta en yüksek işlem gerçekleşen gündeki işlemlerin tutar ortalamasına dönüştürülmüştür.

- Referans makale veri setinde 0,25 olarak kabul edilen bb oranı yeni veri setinin yapısına göre 2 olarak değiştirilmiştir.

Yeni kuralın yukarıda belirtilen çerçevede yeniden tanımlanmasından sonra kuralda gerçekleşen maksimum işlem tutar ile ortalama tutar arasındaki ilişki tanımlanmış ve belirli bir oranın üzerinde gerçekleşen maksimum işlem tutarının şüpheli işleme işaret etme olasılığına işaret edilmiştir.

3.2.5. Kural 5

Bu kuralda her bir hesap için aynı günde, aynı miktarda en fazla kaç tane işlem yapıldığı hesaplanmaktadır.

İlgili günde gerçekleşen maksimum işlem tutarı 4'ten büyük ise, bu işlemlerin şüpheli işlem olma olasılığı bulunmaktadır.

Eğer maxsamt>4; bu durum şüpheli işleme işaret edebilir.

Bu durumda kritik değer maksimum işlem sayısının toplam işlem sayısına oranıdır.

$$Kritik\ Değer = \frac{maxamt}{transn} \qquad (3.7)$$

Kurala göre eğer maksamt≤ 4; işlem şüpheli değildir ve kritik değer 0,01'dir.

Veri setindeki tüm tutarlar USD para birimindedir. Veri setindeki veriler 2013 yılına ait olduğundan farklı para cinsindeki işlem tutarları her bir para birimi için 31.12.2013 tarihli döviz kurundan USD'ye çevrilmiştir (X-Rates, 2021: https://www.x-rates.com/historical/?from=USD&amount=1&date=2013-12-31).

Beşinci kural yeni veri setinde yeniden yorumlanmıştır. Referans makaledeki veri setinde hesaplanan günlük harcama tutarı toplam harcama tutarı ile karşılaştırılmakta (bölüm 2.4.5) ve belirli bir oranın üzerindeki işlemlerin şüpheli işlem olma durumuna işaret edilmekte iken kuralın yeni veri setinde günlük işlem tutarı yerine maksimum işlem sayısı ile toplam işlem sayısı değişkenleri kullanılmış ve belirli bir oranın üzerindeki işlemler şüpheli işlemler olarak varsayılmıştır.

3.3. Genetik Algoritmanın Yeni Veri Setine Uyarlanması

Veri setine yeniden uyarlanmış kurallar uygulandığında ekteki sonuçlara ulaşılmıştır. Birinci kuralın çalıştırılması sonucunda veri setinde şüpheli iki hesap tespit edilmiştir;

In CC ID : 3538-Usage CCfreq Fraud is found with value:873.600000 and

cc:874.958214

In CC ID : 6108-Usage CCfreq Fraud is found with value:333.500000 and

cc:335.303103

İkinci kuralın çalıştırılması sonucunda şüpheli dört hesap tespit edilmiştir;

In CC ID : 1760-Usage Location Fraud is found with value:6.000000 and

cc:7.753133

In CC ID : 2861-Usage Location Fraud is found with value:9.000000 and

cc:15.127266

In CC ID : 2864-Usage Location Fraud is found with value:9.000000 and

cc:65.980899

In CC ID : 6484-Usage Location Fraud is found with value:9.000000 and

cc:13.010000

Üçüncü kuralın çalıştırılması sonucu şüpheli 4 adet hesap tespit edilmiştir.

In CC ID : 1836-Usage BookBalance Fraud is found with value:5.372937 and

cc:5.605058

In CC ID : 3123-Usage BookBalance Fraud is found with value:4.166667 and

cc:6.886667

In CC ID : 3499-Usage BookBalance Fraud is found with value:4.707299 and

cc:5.260632

In CC ID : 3881-Usage BookBalance Fraud is found with value:4.527395 and

cc:4.913696

Dördüncü kural çıktı vermemiştir.

Beşinci kuralın çalıştırılması sonucunda şüpheli 13 adet hesap tespit edilmiştir;

In CC ID : 75-Usage Daily Spending Fraud is found with value:1.000000 and

cc:9.010000

In CC ID : 1023-Usage Daily Spending Fraud is found with value:0.625000 and

cc:5.635000

In CC ID : 1322-Usage Daily Spending Fraud is found with value:1.000000 and

cc:8.010000

In CC ID : 1323-Usage Daily Spending Fraud is found with value:1.000000 and

cc:31.010000

In CC ID : 2498-Usage Daily Spending Fraud is found with value:1.000000 and

cc:17.010000

In CC ID : 2864-Usage Daily Spending Fraud is found with value:0.963636 and

cc:65.980899

In CC ID : 2919-Usage Daily Spending Fraud is found with value:1.000000 and

cc:25.010000

In CC ID : 2921-Usage Daily Spending Fraud is found with value:1.000000 and

cc:7.010000

In CC ID : 3484-Usage Daily Spending Fraud is found with value:1.000000 and

cc:7.010000

In CC ID : 3538-Usage Daily Spending Fraud is found with value:0.348214 and

cc:874.958214

In CC ID : 3607-Usage Daily Spending Fraud is found with value:1.000000 and

cc:7.010000

In CC ID : 4731-Usage Daily Spending Fraud is found with value:0.833333 and

cc:2.593333

In CC ID : 6108-Usage Daily Spending Fraud is found with value:0.793103 and

cc:335.303103

Genetik algoritma optimizasyonu iki hesaba işaret etmektedir;

Credit Card with ID 3538 is detected as fraud with 2 occurances and its critical values

is 874.958214

Credit Card with ID 6108 is detected as fraud with 2 occurances and its critical values

is 335.303103

3.4. Sonuçların Yorumlanması

Çalışmada iki ayrı veri setinde çalışılmış ve referans makaleye göre ekteki sonuçlara ulaşılmıştır. Ulaşılan sonuçların farklı olmasının temel nedeni kurallarda yeni veri setinde anlamlı sonuçlar elde etmek için yapılan değişiklikler ve ikinci veri setinin birinci veri setine göre daha büyük olmasıdır.

Tablo 5. Karşılaştırma Tablosu

	Çıktı Sayısı	
Kural	**Referans Makale**	**Uyarlanmış Veri Seti**
1	5	2
2	4	4
3	4	0
4	2	4
5	3	13
Bulunan kritik çıktı sayısı	**4**	**2**

Referans makaleye göre yeni veri setine uyarlama yapılırken yapılan değişiklikleri özetlendiğinde;

Kural 1'de referans makaleye göre yeni veri setine uyarlama yapılırken yapılan değişiklikler şunlardır;

- Kredi kartının kullanım sıklığı toplam kullanılan kart sayısının kredi kartının yaşına bölünmesi ise hesaplanmaktadır.

- Kredi kartının yaşı yeni veri setinde her bir hesapta en yüksek tarihli günden en düşük tarihli günün çıkarılması olarak değiştirilmiştir.

- Kredi Kartının toplam kullanılma sayısı ilgili hesapta yapılan toplam işlem sayısı olarak değiştirilmiştir.

- Önceki veride kullanılan kredi kartı işlemi sayısı olan CU, Transn her bir hesaptaki tüm işlemlerin sayısı olarak yeni veri setine uyarlanmıştır.

- CUT bugün veya seçilen gündeki kredi kartı kullanım sayısını verirken yeni veri setine TransT olarak uyarlanmıştır. TransT yeni veri setinde en çok işlem yapılan gün olarak seçilmiş ve yeni veri setine uyarlanmıştır.

- Referans değerler yeni veri setinin yapısına göre 0,2 yerine 4 olarak değiştirilmiş ve 5 olan katsayı sıfırlanmıştır.

Kural 2'de yeni veri setinde lokasyon işlem IP adresi ile ilişkilendirilmiştir. Ek olarak kritik değer hesaplaması en çok işlem yapılan gündeki lokasyon sayısı ile toplam lokasyon sayısının çarpımı olarak değiştirilmiş ve aynı zamanda referans makaleye göre katsayılar değiştirilmiştir.

Kural 3'de referans makale veri setinde kullanılan kredi kartı işlemi sayısı olan CU değişkeni Transn değişkeni olarak değiştirilmiş ve Transn her bir hesaptaki tüm işlemlerin sayısı kuralı olarak yeni veri setine uyarlanmıştır.

Her hesaba ilişkin limit üstü işlem sayısı hesaplanmış ve kural olarak veri setinde her hesaba ait en çok işlem yapılan gün seçilmiştir.

Limit üstü işlem tutarı her bir hesabın ortalama işlem tutarının iki standart sapma oranında üzeri olarak varsayılmıştır.

Veri setinin özelliklerine göre katsayı ve karşılaştırma oranları değiştirilmiştir.

Kural 4'de kullanılan referans veri setinde Curremt BB güncel bakiyeye eşit iken yeni veri setine her hesap için en çok işlem gerçekleşen günde gerçekleşen maksimum işlem tutarı olarak uyarlanmıştır.

Average BB ilk veri setinde ortalama bakiyeye eşit iken bu veri setinde her hesapta en yüksek işlem gerçekleşen gündeki işlemlerin tutar ortalamasına dönüştürülmüştür.

İlk veri setinde 0,25 olarak kabul edilen bb oranı veri setinde veri yapısına uygun olarak 2 olarak değiştirilmiştir.

Yeni kuralın yukarıda belirtilen çerçevede yeniden tanımlanmasından sonra kuralda gerçekleşen maksimum işlem tutar ile ortalama tutar arasındaki ilişki tanımlanmış ve belirli bir oranın üzerinde gerçekleşen maksimum işlem tutarının şüpheli işleme işaret etme olasılığına işaret edilmiştir.

Son olarak beşinci kural yeni veri setinde yeniden yorumlanmıştır. Referans makaledeki veri setinde hesaplanan günlük harcama tutarı toplam harcama tutarı ile karşılaştırılmakta ve belirli bir oranın üzerindeki işlemlerin şüpheli işlem olma durumuna işaret edilmekte iken kural yeni veri setinin yapısı gereği maksimum işlem tutarı ile toplam işlem tutarının oranını kullanarak belirli bir oranın üzerindeki işlemler şüpheli işlemler olarak varsayılmıştır.

3.5. Çalışmada Geliştirilebilecek Bölümler ve Uyarlanabileceği Yeni Alanlar

Seçilen veri seti gerçekleşmiş olan hileli işlem verisini içermemektedir. Bu açıdan yazılan kural çıktıları gerçek değerler ile test edilememiştir. Yine aynı neden ile veri setinin veri seti test ve eğitim verisi olarak ayrılamamış, veri setine öğrenme kabiliyeti kazandırılamamıştır. Aynı kapsamda gerçek veri ve test veri arasında Tip 1 ve Tip 2 hata oranları hesaplanamamıştır. Çalışmanın geliştirilebilecek bölümü benzeri kuralların

gerçekleşmiş olan hileli işlemlerden oluşan bir veri seti ile belirtilen hata oranlarının hesaplanarak kuralların güvenilirliği test edilmesidir.

Yine farklı bir açılım olarak kredi kartı işlemlerinde hile tespit alanını inceleyen bu çalışma kuralların yeniden uyarlanması ile birinci bölümde diğer alanlar olarak belirtilen birçok alana uyarlanabilir. Örneğin, hasılat alanında yapılan çalışmalarda bir destek aracı olarak kullanılabilir, yazılan kurallar aracılığı ile hasılat bilgilerinde normal olmayan durumları belirleyebilir anormallikleri veya standartların dışında kalan değerleri ortaya çıkarabilir.

BÖLÜM IV.

SONUÇ

Günümüzde finansal işlemler teknolojisinin geldiği noktada işlem hızı ve kapasitesi yüksek bir seviyeye ulaşmıştır. Finansal işlemlerde elde edilen hıza ve işlem rahatlığına paralel olarak yapılan işlemlerin güvenilirliğini sağlamak aynı derecede önem kazanmıştır. Teknolojiye bağlı dolandırıcılık ve yolsuzluklar işlemlerin artışı ile birlikte hız kazanmıştır. Gerek verinin büyüklüğü gerek usulsüzlük işlemlerinin giderek daha karmaşık ve tespiti zor bir şekilde gerçekleştirilmeye başlanması, bu işlemlerin tespiti için yeni yaklaşımlara ihtiyaç duyulduğu ortaya çıkmıştır.

Usulsüz işlem gerçekleştikten sonra yapılan tespitler kaybı azaltmamakta ve işlem gerçekleştiği için bu tip usulsüz işlemlerde geriye dönüş zor olmaktadır. İhtiyaç duyulan yaklaşım proaktif bir yaklaşım ile henüz usulsüz işlem gerçekleşmeden işlemin tespitidir. Bu noktada kırmızı bayrak adı verilen önceden hile tespiti konusu önem kazanmaktadır. Günümüz finansal işlemler teknolojisinde veri setinin içerisine belirli hile kuralları yazılmakta ve işlem gerçekleşmeden önce sistem bu kurallar aracılığı ile sinyal vermekte ve kırmızı bayraklar oluşturulmaktadır. Şüpheli işlemlerin onayı işlem gerçekleşmeden önce durdurulmakta ve böylelikle olası yüksek kayıpların önüne geçilmektedir.

Tezde yapılan çalışmada referans makaledeki veri setinde kodlama yeni bir dilde yapılmış, alternatif bir metot denenmiş ardından hile kuralları daha büyük bir veri setine uyarlanmış ve yeni veri setinde hileli işlemleri işaret eden bir önceden uyarı mekanizması oluşturularak bazı şüpheli işlemler tespit edilmiştir.

Diğer taraftan veri setinde hileli olan olmayan işlem ayrımı olmadığından oluşturulan hile kurallarının güvenilirliği veri setinin yapısından dolayı tam olarak test edilememiştir.

Sonuç olarak gelinen nokta, işletme, kurum veya kuruluşların iç veya dış denetçilerinin işlem veri setlerinin veri yapısını inceleyerek ve verinin özelliklerine göre hile kuralları oluşturarak bazı şüpheli işlemleri gerçekleşmeden tespit edebilecekleri bir mekanizma oluşturabilecekleri yönündedir. Bu uygulama proaktif hile denetim teknikleri ile örtüşerek

denetçinin bazı işlemlere yoğunlaşmasına yardımcı olan bir denetim aracı olarak kullanılabilir.

KAYNAKÇA

A short history of Evolutionary Algorithms. *Article Data Science*. https://www.mydatamodels.com/a-short-history-of-evolutionary-algorithms/. Erişim Tarihi: 12.03.2021.

Abbott, L. J., Parker, S., & Peters, G. F. (2004). Audit committee characteristics and restatements. *Auditing: A journal of practice & theory*,*23*(1), 69-87.

ACFE (2018). *Global Study on Occupational Fraud and Abuse Report to the Nations*. https://www.acfe.com/report-to-the-nations/2018/default.aspx . Erişim tarihi: 18.12.2021.

ACFE (2020). *Global Study on Occupational Fraud and Abuse Report to the Nations*. https://www.acfe.com/report-to-the-nations/2020/ . Erişim tarihi: 16.12.2021.

EY (2020). *Adli Teknoloji ve Keşif Hizmetleri (2020).* https://www.ey.com/tr_tr/forensic-integrity-services/adli-teknoloji-ve-kesif-hizmetleri. Erişim tarihi: 18.12.2021.

Albrecht, W. S., & C. O. Albrecht. (2003)**.** *Fraud Examination and Prevention*. Mason, OH: Thompson-Southwestern.

Albrecht, W. S., Albrecht, C. C., & Albrecht, C. O. (2004). Fraud and Corporate Executives: Agency, Stewardship and Broken Trust. *Journal of Forensic Accounting*, 5, 109–130.

Albrecht, W. S., Romney, M. B., Cherrington, D. J., Payne, I. R., Roe, A. J., & Romney, M. B. (1986). Red-flagging management fraud: A validation. *Advances in Accounting*, *3*(3), 323-333.

Asare, S. K., & Wright, A. M. (2004). The effectiveness of alternative risk assessment and program planning tools in a fraud setting.*Contemporary Accounting Research*,*21*(2), 325-352.

Bahnsen, A. C., Stojanovic, A., Aouada, D., & Ottersten, B. (2013). Cost sensitive credit card fraud detection using Bayes minimum risk. In *2013 12th international conference on machine learning and applications* (Vol. 1, pp. 333-338). IEEE.

Beasley, M. S. (1996). An empirical analysis of the relation between the board of director composition and financial statement fraud.*Accounting review*, 443-465.

Bedard, J. C., & Johnstone, K. M. (2004). Earnings manipulation risk, corporate governance risk, and auditors' planning and pricing decisions. *The Accounting Review*, *79*(2), 277-304.

Bell, T. B., & Carcello, J. V. (2000). A decision aid for assessing the likelihood of fraudulent financial reporting. *Auditing: A Journal of Practice & Theory*, *19*(1), 169-184.

Beneish, M. D. (1999a). Incentives and penalties related to earnings overstatements that violate GAAP. *The Accounting Review*,*74*(4), 425-457.

Beneish, M. D. (1999b). The detection of earnings manipulation. *Financial Analysts Journal*, *55*(5), 24-36.

Bentley, P. J., Kim, J. W., Jung, G. H., & Choi, J. U. (2000). Fuzzy darwinian detection of credit card fraud. In *Proceedings of the Korea Information Processing Society Conference* (pp. 277-280). Korea Information Processing Society.

Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical science*,*17*(3), 235-255.

Brabazon, A., Cahill, J., Keenan, P., & Walsh, D. (2010). Identifying online credit card fraud using artificial immune systems. In *IEEE Congress on Evolutionary Computation* (pp. 1-7). IEEE.

Buckles, B.P. & Petry, F. E., (1992). *An overview of genetic algorithms and their applications, in Genetic Algorithms*. California: IEEE Computer Society Pres. ,5-11.

Burns, N., & B. Kedia. (2006). The impact of performance-based compensation on misreporting. *Journal of Financial Economics*, *79*(1), 35–67.

Cambridge Dictionary, https://dictionary.cambridge.org/tr/ . Erişim Tarihi: 08.12.2021.

Coenen, Tracy L. (2008). *Essentials of Corporate Fraud*, John Wiley and Sons.

Cottrell, D. M., & Albrecht, W. S. (1994). Recognizing the symptoms of employee fraud. *Healthcare financial management: journal of the Healthcare Financial Management Association*, *48*(5), 18-22.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory*, *13*(1), 21-27.

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1996). Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC. *Contemporary accounting research*, *13*(1), 1-36.

Doğan, S., & Kayakıran, D. (2017). İşletmelerde hile denetiminin önemi. *Maliye ve Finans Yazıları*, 108, 167-187.

Doost, R. K. (1990). Accounting Irregularities and Computer Fraud, *The National Public Accountant*, (May) 36-9.

Efendi, J., Srivastava, A., & Swanson, E. P. (2007). Why do corporate managers misstate financial statements? The role of option compensation and other factors. *Journal of financial economics*, *85*(3), 667-708.

Erickson, M., Hanlon, M., & Maydew, E. L. (2006). Is there a link between executive equity incentives and accounting fraud?. *Journal of accounting research*, *44*(1), 113-143.

Esakkiraj, S., & Chidambaram, S. (2013). A predictive approach for fraud detection using hidden Markov model. *International Journal of Engineering Research & Technology*, *2*(1), 1-7.

Anonim (2021). *Experian Credit Score*. https://www.experian.co.uk/blogs/latest-thinkinge/fraud-prevention/5-of-the-most-remarkable-instances-in-the-history-of-fraud/.Erişim Tarihi: 10.01.2021

Fanning, K., Cogger, K. O., & Srivastava, R. (1995). Detection of management fraud: a neural network approach. *Intelligent Systems in Accounting, Finance and Management*, *4*(2), 113-126.

Farber, D. B. (2005). Restoring trust after fraud: Does corporate governance matter?.*The accounting review*,*80*(2), 539-561.

Deloitte (2021). *Fraud Analytics.* https://www2.deloitte.com/tr/en/pages/deloitte-analytics/articles/fraud-analytics.html. Erişim Tarihi,12.02.2022.

Pwc (2021). *Fraud and Economic Crime*. https://www.pwc.com/us/en/services/consulting/cybersecurity-risk-regulatory/fraud-financial-crime.html. Erişim tarihi: 12.11.2021.

Anonim (2021). *Fraud Detection Data Set*. https://www.kaggle.com/alukosayoenoch/datascientist. Erişim tarihi: 13.11.2021.

Free, C. (2015). Looking through the fraud triangle: A review and call for new directions. *Meditari Accountancy Research*, *23* (2), 175-196.

Ganji, V. R., & Mannem, S. N. P. (2012). Credit card fraud detection using anti-k nearest neighbor algorithm. *International Journal on Computer Science and Engineering*, *4*(6), 1035-1039.

Gillett, P. R., & Uddin, N. (2005). CFO intentions of fraudulent financial reporting. *Auditing: A Journal of Practice & Theory*, *24*(1), 55-75.

Glover, S. M., Prawitt, D. F., Schultz Jr, J. J., & Zimbelman, M. F. (2003). A test of changes in auditors' fraud- related planning judgments since the issuance of SAS No. 82. *Auditing: A Journal of Practice & Theory*, *22*(2), 237-251.

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning.* Addison-Wesley.

Green, B. P., & Choi, J. H. (1997). Assessing the risk of management fraud through neural network technology. *Auditing*, *16*, 14-28.

Güneş, Ş. (2014). İşletmelerde Hile Riskinin Önlenmesi ve Hastane İşletmelerinde Uygulama. (*Yayımlanmamış yüksek lisans tezi). Okan Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul*.

Hamilton, E. L. (2016). Evaluating the intentionality of identified misstatements: How perspective can help auditors in distinguishing errors from fraud. *Auditing: A Journal of Practice & Theory*, *35*(4), 57-78.

Hernandez, J. R., & Groot, T. (2007). How Trust Underpins Auditor Fraud Risk Assessments. *Free University of Amsterdam Working paper***.**

Hogan, C. E., Rezaee, Z., Riley Jr, R. A., & Velury, U. K. (2008). Financial statement fraud: Insights from the academic literature. *Auditing: A Journal of Practice & Theory*, *27*(2), 231-252.

How Average Nearest Neighbor Works, ArcGIS Pro 2.7, https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-average-nearest-neighbor-distance-spatial-st.htm. Erişim Tarihi: 17.06.2020.

Introduction to Genetic Algorithms — Including Example Code, Towards Data Science. https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3 . Erişim Tarihi: 12.11.2021.

Intuitive explanation of Minimum Covariance Determinant (MCD), https://stats.stackexchange.com/questions/475636/intuitive-explanation-of-minimum-covariance. Erişim Tarihi: 15.03.2020.

Korucu, V. (2021). *Genetik Algoritmada Tek ve Çok Noktalı Çaprazlama Tekniklerinin Anten Dizisi Sentez Probleminde İncelenmesi*.https://www.academia.edu/4715561/Genetik_Algoritmada_Tek_ve_Çok_Noktalı_Çaprazlama_Tekniklerinin_Doğrusal_Anten_Dizisi_Sentez_Probleminde_İncelenmesi, (Erişim Tarihi:23.05.2021)

Kotu, V., & Deshpande, B. (2015). Anomaly Detection in Predictive Analytics and Data Mining, Elsevier Inc.

KPMG (2021). Anti-Money Laundering and Trade Sanctions Services, https://home.kpmg/id/en/home/services/advisory/forensic-services/anti-money-laundering-and-trade-sanctions-services.html. Erişim tarihi: 12.11.2021.

Kriegel, H., Schubert, M. & Zimek, A. (2008). Angle-Based Outlier Detection in High-dimensional Data**.** *In Proceedings of the International Conference on Knowledge Discovery & Data Mining (KDD'08)*, Las Vegas, NV, (pp. 444-452).

Lie, E. (2005). On the Timing of CEO Stock Option Awards. *Management Science***,** 51(5): 802–812.

Loebbecke, J. K., Eining, M. M., & Willingham, J. J. (1989). Auditors experience with material irregularities-frequency, nature, and detectability. *Auditing-A Journal of Practice & Theory*, *9*(1), 1-28.

Mackevičius, J., & Giriūnas, L. (2013). Transformational research of the fraud triangle. *Ekonomika*, *92*(4), 150-163.

MacQueen, J. (1967, June). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability* (Vol. 1, No. 14, pp. 281-297).

Maes, S., Tuyls, K., Vanschoenwinkel, B., & Manderick, B. (2002). Credit card fraud detection using Bayesian and neural networks. In *Proceedings of the 1st international naiso congress on neuro fuzzy Technologies* (Vol. 7).

Mahmoudi, N., & Duman, E. (2015). Detecting credit card fraud by modified Fisher discriminant analysis. *Expert Systems with Applications*, *42*(5), 2510-2516.

Makki, S. (2019). *An efficient classification model for analyzing skewed data to detect frauds in the financial sector* (Doctoral dissertation, Université de Lyon; Université libanaise).

Marks, J. (2009). Playing Offense in a High-Risk Environment. *New York: Crowe Horwath*.

McDaniel, L., Martin, R. D., & Maines, L. A. (2002). Evaluating financial reporting quality: The effects of financial expertise vs. financial literacy. *The accounting review*, *77*(s-1), 139-167.

McMullen, D. A., & Raghunandan, K. (1996). Enhancing audit committee effectiveness. *Journal of Accountancy*,*182*(2), 79.

Moyes, G. D., Din, H. F. M., & Omar, N. H. (2009). The effectiveness of the auditing standards to detect fraudulent financial reporting activities in financial statement audits in Malaysia. *International Business & Economics Research Journal (IBER)*, *8*(9).

Mzila, P. and Dube, E. (2013). The Effect of Destination Linked Feature Selection In Real-Time Network Intrusion Detection. *In Proceedings of the ICIMP 2013: The Eighth International Conference on Internet Monitoring and Protection*.

Nelson, M. W., Elliott, J. A., & Tarpley, R. L. (2002). Evidence from auditors about managers' and auditors' earnings management decisions. *The accounting review*, *77*(s-1), 175-202.

Pasarica, A. (2014). Card fraud detection using learning machines. *Bull. Polytech. Inst. Jassy, Fac. Cybern., Statist. Econ. Inform., Bucharest, Romania, Tech. Rep*, 29-45.

Patterson, E., & Noel, J. (2003). Audit strategies and multiple fraud opportunities of misreporting and defalcation. *Contemporary Accounting Research*, *20*(3), 519-549.

Paulauskas, N., & Baskys, A. (2019). Application of histogram-based outlier scores to detect computer network anomalies. *Electronics*, *8*(11), 1251.

Pincus, K. V. (1989). The efficacy of a red flags questionnaire for assessing the possibility of fraud. *Accounting, Organizations and Society*, *14*(1-2), 153-163.

RamaKalyani, K., & UmaDevi, D. (2012). Fraud detection of credit card payment system by genetic algorithm. *International Journal of Scientific & Engineering Research*,*3*(7), 1-6.

Ramamoorti, S., Morrison III, D. E., Koletar, J. W., & Pope, K. R. (2013).*ABC's of behavioral forensics: applying psychology to financial fraud prevention and detection*. John Wiley & Sons.

Rezaee, Z. (2004). Restoring public trust in the accounting profession by developing anti-fraud education, programs, and auditing. *Managerial Auditing Journal*.

Romney, M. B., Albrecht, W. S., & Cherrington, D. J. (1980). Auditors and the detection of fraud. *Journal of Accountancy*, *149*(5), 63-69.

Rosner, R. L. (2003). Earnings manipulation in failing firms. *Contemporary accounting research*, *20*(2), 361-408.

Sahin, Y., & Duman, E. (2010). Detecting credit card fraud by decision trees and support vector machines. In *World Congress on Engineering 2012. July 4-6, 2012. London, UK.*(Vol. 2188, pp. 442-447). International Association of Engineers.

Sahin, Y., Bulkan, S., & Duman, E. (2013). A cost-sensitive decision tree approach for fraud detection.*Expert Systems with Applications*,*40*(15), 5916-5923.

Sánchez, D., Vila, M. A., Cerda, L., & Serrano, J. M. (2009). Association rules applied to credit card fraud detection. *Expert systems with applications*, *36*(2), 3630-3640.

Srivastava, A., Kundu, A., Sural, S., & Majumdar, A. (2008). Credit card fraud detection using hidden Markov model. *IEEE Transactions on dependable and secure computing*, *5*(1), 37-48.

Summers, S. L., & Sweeney, J. T. (1998). Fraudulently misstated financial statements and insider trading: An empirical analysis. *Accounting Review*, 131-146.

Sutherland, E.H. (1940). White-collar criminality, *American Sociological Review*, 5(1), 1–12

Sutton, C., Sindelar, M. and McCallum, A. (2005). Feature Bagging: Preventing Weight Undertraining in Structured Discriminative Learning, *CIIR Technical Report IR-402*, University of Masschusetts.

Taşcı, E., & Onan, A. (2016). K-en yakın komşu algoritması parametrelerinin sınıflandırma performansı üzerine etkisinin incelenmesi. *Akademik Bilişim*, *1*(1), 4-18.

Türedi, H. (2020). *Hile Denetimi.* İstanbul Ticaret Üniversitesi İşletme Fakültesi

Üstünel, M. (2018). K-Ortalamalar Algoritmasına Dayalı Kümeleme Analizi Sistemi ve Perakendecilik Sektöründe Uygulaması, Yıldız Teknik Üniversitesi. *Fen Bilimleri Enstitüsü, İstanbul*.

Vardar, M. K. (2019). Hilenin önlenmesi ve ortaya çıkarılması: muhasebe meslek mensuplarının görüşleri üzerine nitel bir araştırma. *İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü*.

Vassiljev, M., & Alver, L. (2016). Concept and periodisation of fraud models: Theoretical review. In *5th International Conference on Accounting, Auditing, and Taxation. Atlantis Press.*

Pwc (2021). *Veri Analitiği ve Adli Bilişim Çözümleri*, https://www.pwc.com/veri-analitigi-ve-adli-bilisim-cozumleri. Erişim tarihi: 13.10.2021.

Wells, A. (1990). Panic disorder in association with relaxation induced anxiety: An attentional training approach to treatment. *Behavior therapy*, *21*(3), 273-280.

Weston, D. J., Hand, D. J., Adams, N. M., Whitrow, C., & Juszczak, P. (2008). Plastic card fraud detection using peer group analysis. *Advances in Data Analysis and Classification*, *2*(1), 45-62.

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 discovery*, *18*(1), 30-55.

Wilks, T. J., & Zimbelman, M. F. (2004a). Using game theory and strategic reasoning concepts to prevent and detect fraud. *Accounting horizons*, *18*(3), 173-184.

Wolfe, D. T. & Hermanson, D.R. (2004). The Fraud Diamond: Considering the Four Elements of Fraud, *CPA Journal*, 74,38–42.

Wong, N., Ray, P., Stephens, G., & Lewis, L. (2012). Artificial immune systems for the detection of credit card fraud: an architecture, prototype and preliminary results. *Information Systems Journal*, *22*(1), 53-76.

Anonim (2021). *X-rates*. https://www.x-rates.com/historical/?from=USD&amount=1&date=2013-12-31. Erişim tarihi: 31.12.2021.

Yücel, E. (2013). Effectiveness Of Red Flags in Detecting Fraudulent Financial Reporting: An Application In Turkey. *Journal of Accounting & Finance*, (60).

Zimbelman, M. F. (1997). The effects of SAS No. 82 on auditors' attention to fraud risk factors and audit planning decisions. *Journal of Accounting Research*, *35*, 75-97.