T.C.

ANKARA SOSYAL BİLİMLER ÜNİVERSİTESİ

SOSYAL BİLİMLER ENSTİTÜSÜ

MAKİNE ÖĞRENMESİ ALGORİTMALARI İLE SUİSTİMALLERİN BELİRLENMESİ

Yüksek Lisans Tezi

Harun KUTLUAY

Denetim ve Risk Yönetimi Ana Bilim Dalı

Tezli Yüksek Lisans Programı

Ocak 2023

T.C.

ANKARA SOSYAL BİLİMLER ÜNİVERSİTESİ

SOSYAL BİLİMLER ENSTİTÜSÜ

MAKİNE ÖĞRENMESİ ALGORİTMALARI İLE SUİSTİMALLERİN BELİRLENMESİ

Yüksek Lisans Tezi

Harun KUTLUAY

Denetim ve Risk Yönetimi Ana Bilim Dalı

Tezli Yüksek Lisans Programı

Tez Danışmanı

Dr. Öğr. Üyesi Furkan UYSAL

Ocak 2023

TEŞEKKÜR

Bu çalışmanın hazırlanması sürecinde bilgi ve deneyimleriyle beni yönlendiren, desteklerini esirgemeyen ve her konuda yardımcı olmaya çalışan değerli danışman hocam Dr. Öğr. Üyesi Furkan UYSAL'a ve Dr. Öğr. Üyesi Esra SATICI'ya şükranlarımı sunarım.

Tanıdığım günden bu yana her zaman desteğini üzerimde hissettiğim ve göstermiş olduğu sabır ve anlayışı için değerli eşim Nur'a, varlığı ile beni motive eden oğlum Asaf'a, emek, fedakârlık ve dualarıyla bugünlere gelmeme vesile olan ve manevi desteklerini her daim hissettiğim kıymetli anneme ve babama ve tez çalışması boyunca motivasyonumu diri tutan, içten yorumlarıyla katkısını esirgemeyen dostum Mustafa Faruk KOCA başta olmak üzere, yardımlarını eksik etmeyen tüm arkadaşlarıma teşekkürü bir borç bilirim.

İÇİNDEKİLER

ÖZET

MAKİNE ÖĞRENMESİ ALGORİTMALARI İLE SUİSTİMALLERİN BELİRLENMESİ

Kurumlar için tehdit unsuru olarak görülen suistimaller, kurumların yalnızca itibarını zedeleyip, güvensizlik ortamı oluşturmakla kalmayıp, faaliyetlerinin sonlandırılmasına neden olacak kadar etkili olabilmektedir. Suistimaller, yönetilmesi gereken en önemli kurumsal risklerden biri olmakla birlikte, bu riskin gerçekleşmesi durumunda ise sosyal ve ekonomik açıdan birçok olumsuz sonuç ile karşılaşılabilmektedir. Bu bağlamda, böyle bir riski yönetebilmek ve beraberinde sürdürülebilirliği sağlamak için, suistimallerin önlenmesi ve büyük boyutlara ulaşamadan ortaya çıkarılmasına yönelik, başta bankacılık ve finans sektöründe olmak üzere tüm sektörlerde faaliyet gösteren organizasyonların, nitelikli insan kaynağının yanı sıra veri analitiğine dayalı teknolojiye yatırım yapmaları kaçınılmaz olmuştur.

Bu tez çalışmasının amacı, bankacılık ve finans sektöründeki suistimallerin ortaya çıkarılması konusunda, çalışanların demografik, mesleki ve finansal bilgileri üzerinden makine öğrenmesi algoritmalarından beşi aracılığıyla suistimal yapıp yapmadığına yönelik öngörü modelleri geliştirmek ve çalışanların profilleri üzerinden suistimallerin belirlenmesine yönelik literatürdeki eksikliği gidermektir.

Bu çalışmada yöntem olarak, bir bankada 2017 ile 2021 yılları arasında suistimal yapan ve yapmayan çalışanlara ilişkin istatistiksel olarak anlamlı 8 değişken kullanılarak, Lojistik Regresyon, Rassal Orman, Karar Ağacı, Naive Bayes ve Destek Vektör Makinesi sınıflandırma algoritmaları ile çalışanların suistimal gerçekleştirip gerçekleştirmediği test edilmiş ve bu algoritmaların performans değerleri karşılaştırılmıştır.

Çalışma sonucunda, Yaş, Cinsiyet, Eğitim Düzeyi, Birimde Çalışma Süresi, İzin Kullanım Oranı, KKB Skoru, Kredi Başvuru Adedi ve Kredi Kartı Doluluk Oranı değişkenleri kullanılarak oluşturulan Lojistik Regresyon modelinin en yüksek doğruluk oranı ve F-Skor değerine sahip olduğu belirlenmiştir.

Anahtar Kelimeler: Suistimal, Denetim, Makine Öğrenmesi, Lojistik Regresyon

ABSTRACT

FRAUD DETECTION BY MACHINE LEARNING ALGORITHMS

Known as a threat to organizations, fraud, not only destroys their reputation by creating an untrustable environment, but also causes bankruptcy. Fraud as an enterprise risk to be managed can have many social and economic effects in the case of not handled. It is inevitable for organizations in a range of sectors, especially in the banking and finance sector, to invest in technology based on data analytics besides qualified human resources to prevent and detect frauds before reaching substantial damages, and to ensure organizational sustainability.

The study aims to develop a predictive model through a machine learning algorithm based on demographic, professional, and financial features for detecting frauds in the banking and finance industry, and to extend the literature in this field by proposing a machine learning-based model.

As a research methodology, whether employees committed fraud was tested by classification algorithms, Logistic Regression, Random Forest, Decision Tree, Naive Bayes, and Support Vector Machine with the features of the bank employees worked between 2017 and 2021, and the performance metrics of these algorithms were compared.

It is concluded that the Logistic Regression model using the features of Age, Gender, Education Level, Working Time in the Unit, Annual Leave Rate, Personal Credit Rating, Number of Loan Applications, and Credit Card Utilization Rate has the highest accuracy rate and F-Score value.

Keywords: Fraud, Audit, Machnine Learning, Logistic Regression

ŞEKİL LİSTESİ

TABLO LİSTESİ

BÖLÜM 1: GİRİŞ

Süreçlerin ve iş modellerinin değişimi ve teknolojik gelişmeler ile birlikte suistimalciler de bunlara paralel olarak daha sistemsel ve anlaşılması zor suistimaller gerçekleştirmekte ve olayların tespit edilme süreleri de bunlara bağlı olarak uzamaktadır. Böyle bir ortamda geleneksel yöntemler ile suistimallerin ortaya çıkarılması güçleşmekte (Yılmaz vd, 2018), dolayısıyla usulsüzlüklerden kaynaklanan kayıp tutarı da her geçen gün artmaktadır. Kurumların, ekseriyetle karşılaştıkları varlıkların kötüye kullanılması, yolsuzluk ve finansal tablo usulsüzlükleri gibi suistimallerden kaynaklı kayıplarını minimize edebilmeleri ve organizasyonlarına katma değer sağlayabilmeleri için risk odaklı ve bilgi teknolojileri destekli sürekli denetim yaklaşımlarını benimseyerek faaliyetlerini sürdürmeleri kaçınılmaz olmuştur.

Bankacılık ve finans sektöründe çalışanlar tarafından gerçekleştirilen suistimaller, genellikle zimmet, güveni kötüye kullanma, banka veya kredi kartlarını kötüye kullanma ve dolandırıcılık eylemleri şeklinde karşımıza çıkmaktadır (Durkaya, 2008; Kaban, 2018). Söz konusu eylemler sonucunda maddi ve/veya itibar kaybı gibi manevi zararlar ile karşılaşabilen kurumlar, suistimallerin erkenden belirlenip zararın büyümemesine yönelik kontrollere ve suistimallerin tespit ve inceleme sürecinden sorumlu olan birimlerine her geçen gün daha fazla önem vermekte ve bu alanlarda kullanılmak üzere teknoloji destekli uygulamalara daha fazla yatırım yapmaktadır.

Bu tez çalışmasında, çalışanların demografik, mesleki ve finansal bilgilerine yönelik belirlenen değişkenler ile 5 farklı makine öğrenmesi algoritması üzerinden suistimal yapıp yapmadığı konusunda başarılı bir şekilde öngörü yapabilen modeller önerilmiştir. Çalışmanın sonucunda, modellerin doğruluk oranı ile F-Skor değerleri kıyaslanmış ve en iyi sınıflandırma algoritmasının Lojistik Regresyon olduğu sonucuna ulaşılmıştır.

1.1 Çalışmanın Amacı ve Önemi

Bu çalışmada, bankacılık ve finans sektöründe operasyonel risk adı altında değerlendirilen, yöneticiler ve çalışanlar tarafından gerçekleştirilen suistimallerin ortaya çıkarılmasına ilişkin teftiş kurulu, iç denetim veya suistimal inceleme birimleri tarafından hâlihazırda kullanılmakta olan yöntemlere ilave olarak, gerçekleşmekte olan suistimal vakalarının daha ileriye gitmeden belirlenebilmesine ve böylece suistimallerden kaynaklı

kayıpların minimize edilebilmesine olanak sağlayacak, çalışanların demografik, mesleki ve finansal bilgileri üzerinden suistimal gerçekleştirip gerçekleştirmediğine yönelik makine öğrenmesi algoritmaları ile sınıflandırmaya dayalı modeller oluşturulması ve bu algoritmaların performans değerlerinin karşılaştırılarak en başarılı modelin ortaya konması ve bu çalışmanın kurumlarda rehber edinilerek uygulanması ve geliştirilmesi, yanı sıra, çalışan profilleri üzerinden suistimallerin belirlenmesine yönelik literatürdeki eksikliğin giderilmesi amaçlanmıştır.

Bu çalışmanın diğer bir amacı ise, bankacılık ve finans alanında, çalışan suistimallerinin ortaya çıkarılmasına yönelik yapay zekâ veya makine öğrenmesi çalışması planlayan araştırmacılara makine öğrenmesi ve yöntemleri, sınıflandırma algoritmaları, uygulama süreci ve performanslarının değerlendirilmesi hakkında bilgi aktarmaktır.

Makine öğrenmesi olarak bilinen bir yöntemin yeni bir alana uygulanarak suistimallerin tespitinde kullanılması, yanı sıra literatürde, suistimalci profiline yönelik yapılan çalışmalarda kullanılan yaş, cinsiyet, eğitim düzeyi, pozisyon ve çalışma süresi gibi değişkenlere ilave farklı demografik, mesleki ve finansal değişkenlerin de eklenmesi ile toplam 23 değişken üzerinden suistimalci profilinin belirlenmesi çalışmamızın önemini ortaya koymaktadır.

1.2 Çalışmanın Kapsamı ve Yöntemi

Çalışmamızın kapsamını, ülkemizde faaliyet gösteren bir bankada 2017 ile 2021 yılları arasında, suistimal gerçekleştiren 60 çalışan ile rastgele örneklem yöntemiyle seçilen suistimal gerçekleştirmeyen 75 çalışan ve bu çalışanlara ilişkin belirlenen 23 adet değişken üzerinden temin edilen veri seti oluşturmaktadır. Öncelikle, bazı değişkenler üzerinden aşırı öğrenmeye sebebiyet verilmemesi ve daha sağlıklı sonuçlar elde edebilmek için SPSS Statistics uygulaması aracılığıyla Lojistik Regresyon analizi gerçekleştirilmiş, değişkenlerin istatistiksel olarak anlamlı olup olmadıkları test edilmiş ve tüm değişkenlerin modele dahil edilmesi yerine istatistiksel olarak anlamlı olanları ile açık kaynak kodlu KNIME veri analitiği platformu aracılığıyla model oluşturma sürecine geçilmiştir. Sınıflar arasındaki dengesiz veri dağılımını gidermek amacıyla SMOTE tekniği ile sentetik örnekler oluşturulmuş ve veri ön işleme süreci sonrası veriler, k-katlı çapraz doğrulama ve tabakalı örnekleme yöntemi ile eğitim ve test verisi olarak ayrılmıştır. Akabinde, makine öğrenmesi sınıflandırma algoritmalarından Lojistik Regresyon, Rassal Orman, Karar Ağacı, Naive Bayes ve Destek Vektör Makinesi ile

çalışanların suistimal gerçekleştirip gerçekleştirmediğine yönelik öngörü modelleri geliştirilmiş ve algoritmaların performans sonuçları karşılaştırılmıştır.

1.3 Çalışmanın Kısıtları

Suistimal vakaları ve bunlara ilişkin yönetilen soruşturma faaliyetleri, itibar riski açısından kurumlar için kritik öneme sahip olduğundan çalışmaya konu finansal kurumun adı gizli tutulmuştur. Suistimaller neticesinde söz konusu kurumun uğramış olduğu zarar ise çalışmamızın konusunu oluşturmadığı gibi modele katkısının da olmayacağı düşüncesiyle talep edilmemiştir. Diğer yandan, araştırmaya konu çalışanların demografik, mesleki ve finansal durumlarına ilişkin verileri kişisel veri içerdiğinden, personelin kimliklerine ilişkin alanlar maskelenmiş bir şekilde temin edilmiştir.

Çalışmamızda kullanılan veri seti, finansal kurumun vermiş olduğu bilgiler ile sınırlı olduğundan, söz konusu verilerin eksik ve/veya hatalı olması muhtemeldir ve bu durum model sonuçlarında da sapmalar olmasına sebebiyet verebilir. Bununla birlikte, devam eden ancak tespit edilememesinden dolayı bir grup çalışanın suistimal gerçekleştirmediği şeklinde sınıflandırılmış olması da olasıdır.

Her ne kadar sınıflandırma modelleri, yüksek başarı oranı ile çalışanın suistimal gerçekleştirip gerçekleştirmediğine yönelik bir öngörü yapmış olsa da, söz konusu personelin kesinlikle eylemi gerçekleştirdiği veya kesinlikle gerçekleştirmediği sonucuna ulaşılamaz ve farklı yönlerden de incelenmesi gerekir.

1.4 Literatür Taraması

Literatürde pek çok sayıda, suistimal, suistimal türleri, suistimal üçgeni, suistimale iten sebepler, suistimallerin önlenmesine ve tespit edilmesine yönelik kullanılan araçlar ve suistimal denetimi üzerine yapılmış çalışmalar bulunduğu gibi (Sandhu, 2019; Azam, 2018; Sánchez vd., 2018; Bozlak, 2019; Kaban, Gül & Eryılmaz, 2018; Doğan & Kayakıran, 2017; Ataman & Aydın, 2017; Elhakan, 2017; Erol 2016; Mui & Mailley, 2015; Yıldız & Başkan, 2014; Dorminey vd., 2012; Emir, 2008), suistimallerin ortaya çıkarılmasında veri analitiği, veri madenciliği ve makine öğrenmesi gibi yöntemlerin kullanılmasına yönelik de çok sayıda çalışma bulunmaktadır (Güven, 2021; Tatlıtürk, 2021; Geren, 2020; Çamlıfidan, 2019; Yılmaz, 2019; Hazım,2018; Kaban, 2018; Aksoy, 2018; Soileau vd., 2015; West & Bhattacharya, 2016;

Kandemir & Kandemir, 2012; Terzi, 2012; Ngai vd, 2010; Çatıkkaş & Çalış, 2010). Diğer yandan suistimal faillerinin genel özelliklerinin ortaya konulmasına yönelik yapılan araştırmaların ise, anket sonuçlarının analizi ile sınırlı olduğu görülmüş olup (ACFE, 2018; ACFE, 2020; ACFE, 2022; KPMG Türkiye &TEİD, 2020; Yılmaz, 2019; Dönmez & Bağışlar, 2017; Yıldırım & Turgut 2016; Durkaya, 2008), yapılan literatür taramasında rastlanılan çalışmalardan bazılarının detayına aşağıda yer verilmiştir.

Uluslararası Suistimal İnceleme Uzmanları Birliği (ACFE), dünya genelinde gerçekleşen suistimaller üzerinde araştırma yapmak amacıyla sertifikalı suistimal denetçilerine yönelik anket çalışmaları yürütmekte olup, araştırma sonuçlarını düzenlemiş olduğu bir rapor ile yayımlamaktadır ve en son rapor 2022 yılına ilişkindir. Söz konusu raporda, 133 ülkeden sertifikalı suistimal inceleme uzmanlarından derlenen 26 farklı sektöre yönelik 8'i Türkiye'den olmak üzere toplamda 2110 vaka analiz edilmiştir. Söz konusu vakalar ışığında, toplam zararın 3.6 milyar doları aştığı, bir şirketin her yıl gelirinin ortalama %5'ini suistimal sebebiyle kaybettiği, suistimalin başlamasından tespit edilmesine kadar geçen sürenin ortalama 12 ayı bulduğu ve tespit etme süresi uzadıkça zararın da arttığı, suistimalcinin pozisyonu yükseldikçe zararın da aynı şekilde arttığı, vakaların tespit edilmesinde en etkili yöntemlerin sırasıyla ihbar (%42), iç denetim (%16) ve yönetim incelemeleri (%12) olduğu belirtilirken, vakaların gerçekleştiği anda suistimal ile mücadeleye yönelik %45'inde veri izleme/analizi kontrolünün bulunduğu, bu yöntemin %56 daha hızlı tespit etme yeteneğine sahip olduğu ve kayıpları %47 oranında azalttığı ve gerçekleşen vakaların en çok bankacılık ve finans sektöründe meydana geldiği bilgisine yer verilmiştir. Usulsüzlükleri gerçekleştiren faillerin profili analiz edildiğinde ise %73'ünün erkek; %68'inin 31 ile 50 yaşları arasında; %47'sinin lisans mezunu; %37'sinin çalışan, %39'unun ise yönetici pozisyonunda bulunduğu; %47'sinin hizmet süresinin 1 ile 5 yıl arasında, %25'inin ise 6 ile 10 yıl arasında olduğu ve hizmet süresi arttıkça kayıpların da aynı şekilde arttığı sonucuna ulaşılmıştır.

KPMG Türkiye ile Etik ve İtibar Derneği (TEİD), 2020 yılında ülkemizdeki suistimallerin ve suistimalcilerin profiline yönelik farklı sektörlerden 157 çalışana 28 sorudan oluşan anket düzenlemiş ve 59 adet vakaya ilişkin sonuçlarını bir suistimalcinin profili 2020 raporunda yayımlamıştır. Söz konusu raporda, çok uluslu şirketlerin %49'unun, yerel şirketlerin ise %22'sinin son iki yıl içerisinde en az bir suistimal ile karşı karşıya geldiği; suistimalcilerin %92'sinin erkek ve yaklaşık %90'ının ise 26-55 yaş aralığında olduğu; vakaların %49'unun yönetici pozisyonun altında çalışan personel tarafından gerçekleştirildiği; kayıpların %31'inin 100.000.-TL ve daha az tutarda, yaklaşık %36'sının ise 500.000.-TL ve

üzerinde olduğu ve suistimalcinin pozisyonu yükseldikçe zararın da aynı şekilde arttığı; vakaların daha çok varlıkların kötüye kullanılması, zimmet ve satın alma usulsüzlükleri şeklinde ortaya çıktığı; iç kontrollerin zayıf, gözetimlerin etkinsiz ve yetkilerin sınırsız olması hususunun, suistimallere sebep olan üç önemli faktör olduğu; vakaların tespitinde ilk sırada yönetim incelemelerinin yer aldığı ve bunu iç denetim, ihbar hatları ve resmi olmayan ihbar kanallarının takip ettiği; diğer yandan vakaların %57'sinin iç denetim, %31'inin hukuk veya uyum birimleri, %12'sinin ise bağımsız denetçi veya danışman tarafından incelendiği, %7'sinin ise hiçbir şekilde incelenmediği; suistimallerin %62'sinin 1 yıl veya daha az, %31'inin 2 ile 5 yıl arasında, %7'sinin ise 6 ile 10 yıl arasında bir sürede tespit edilebildiği; söz konusu şirkette suistimalcilerin %2'sinin 1 yıldan az, %44'ünün 1 ile 4 yıl arasında, %54'ünün ise 5 yıl ve üstü bir süredir çalıştığı; faillerin arkadaş canlısı, ikna kabiliyeti yüksek, hırslı ve üstünlük duygusuna sahip olduğu görülürken, suistimale iten en güçlü motivasyonun ise kişisel menfaat sağlamaya dayalı olduğu sonucuna ulaşılmıştır.

Yılmaz (2019), çalışmasında suistimallerin tespit edilmesi ve önlenmesinde başta veri analitiği olmak üzere kullanılan yöntemlerin etkinliğini değerlendirmiş ve suistimallerin işletmeler üzerindeki etkilerini ele almıştır. Yanı sıra çalışmada, Türkiye'deki işletmelerde gerçekleştirilen suistimallere ve faillere ilişkin bir araştırma yapmak amacıyla 288 kişiye yönelik anket çalışması yürütülmüştür. Anket sonuçlarına göre; işletmelerin suistimal dolayısıyla kayıplarının yıllık gelirlerine oranının %3,51 olduğu, kurumlarda suistimalleri önlemeye yönelik kontrollerin başında davranış ve etik kurallarının geldiği, kayıpların %34'ünün 250.000.-TL ile 1.000.000.-TL arasında olduğu, gerek kayıp tutarı gerek gerçekleştirilen suistimallerin sayısı bakımından bankacılık sektörünün birinci sırada yer aldığı, suistimallerin ortaya çıkarılış biçimlerinde en etkili dört yöntemin sırasıyla ihbar hattı (%43,2), iç denetim (%21,4), sistem üzerinden yapılan kontroller (%6,8) ile veri analizinin (%6,3) olduğu, katılımcıların %35'inin işletmesinde suistimallerin önlenmesi, denetime destek olunması ve ortaya çıkarılması faaliyetlerinde veri analitiği tekniğinin kullanıldığı, kırmızı bayrak olarak nitelendirilen faillere ait davranışların başında gelirine oranla yüksek yaşam tarzına sahip olmaları ile kurum paydaşlarıyla samimi ilişkiler içerinde olmalarının geldiği, suistimal türlerinden harcamalar veya giderler üzerinde yapılan hileler ile zimmete para geçirilmesi eylemlerinin diğerlerine oranla daha sık gerçekleştirildiği sonucuna ulaşılmıştır.

Dönmez ve Bağışlar (2017), çalışmalarında hata, hile ve hile üçgenine dair kavramsal bilgilere yer vermiş, bankacılık sektöründe karşılaşılan hileleri gerçekleştiren kişilerin özelliklerini ortaya koymak amacıyla bankaların teftiş kurulu ve iç kontrol birimleri bünyesinde

çalışan 50 kişiye yönelik anket çalışması uygulamıştır. Ankete katılan kişilerin çoğunun müfettiş olduğu ve tamamımın suistimal içeren işlemlerle karşılaştığı, vakaların genelde şube satış ve gişe işlemlerinde yoğunlaştığı, suistimalleri gerçekleştiren kişilerin %66'sının 31 ile 40 yaşları arasında olduğu, olayların yarısından fazlasının alt unvanlara sahip personel tarafından gerçekleştirildiği, karşılaşılan suistimal vakalarının %30'unda kayıp tutarının 5 milyon TL ve üzerinde olduğu, %76'sının çalışma süresinin en fazla 10 yıl olduğu, davranışsal özellikleri açısından %28'inin çevresiyle uyum içerisinde olduğu sonucuna ulaşılmış olup, ayrıca suistimallerin önlenmesine yönelik bir takım önerilerde bulunulmuştur.

Durkaya (2008), Türkiye'deki bankacılık sektöründe karşılaşılan suistimallerin önlenmesi ve tespit edilmesine ilişkin iç denetim birimleri tarafından gerçekleştirilen faaliyetlere ve alınan aksiyonlara değinerek, sistem içerisindeki önemine vurgu yapmış, suistimallerin engellenmesi ve belirlenmesinde iç denetim faaliyetlerinin daha etkin yürütülmesine yönelik önerilerde bulunmuştur. Çalışmada iç denetim, risk odaklı iç denetim yaklaşımı, suistimal, iç denetim ve suistimale ilişkin regülasyonlar ve bankacılıkta karşılaşılan suistimal çeşitleri gibi hususlar kavramsal boyutta incelenerek sektörde karşılaşılan örneklere yer verilmiştir. Yanı sıra, iki bankaya ait 1998 ve 2008 yılları arasında gerçekleşen 52 adet suistimal vakası incelenerek, olayları gerçekleştiren faillere ilişkin yaş, cinsiyet, medeni durum, eğitim seviyesi, çalışma süresi ve öncesinde farklı bir bankada çalışıp çalışmadığı gibi demografik özellikler üzerinden bir profil analizi yapılmıştır. Çalışmada, suistimal faillerinin yaşlarının 26 ile 37 arasında, %65'inin erkek, %55'inin evli, %59'unun lise mezunu, toplam banka çalışma sürelerinin 1 ile 4 yıl arasında, suistimal yaptıkları birimdeki çalışma sürelerinin 4 ile 36 ay arasında, %92'sinin çalışmaya başlamadan önce bankacılık deneyiminin olduğu sonucuna ulaşılmıştır.

Yıldırım ve Turgut (2016), çalışmalarında finansal tablolar üzerinde gerçekleştirilen suistimaller olarak bilinen yönetim hilesi üzerinde durmuştur. Bu kapsamda, 2010 ile 2014 yılları arasında Erzurum Adliyesinde açılmış olan 171 adet yönetim suistimali dava dosyası üzerinden, suistimaller ile olayları gerçekleştiren 345 yöneticiye ilişkin elde edilen bilgiler istatistiksel olarak analiz edilmiştir. Yapılan araştırmada, suistimali gerçekleştiren yöneticilerin %95,7'si erkek; %73,3'ü 10 yıldan daha fazla iş tecrübesine sahip; %20'si 41 ile 45 yaş arasında; yöneticiler tarafından gerçekleştirilen suistimallerin, hissedarlar tarafından yapılanlara göre sayıca az, fakat tutar olarak daha yüksek olduğu sonucuna ulaşılmıştır.

Ataman ve Aydın (2017), çalışmalarında suistimal aşamaları, suistimal denetimi, suistimal denetçisinin görevlerine ilişkin bilgilere yer vererek, suistimal denetçilerine yönelik hazırlanan anket çalışması ile işletmelerde gerçekleşen suistimal türlerini, alınan önlemleri, hilelerin ortaya çıkarılış biçimlerini ele almıştır. Araştırmada, 73 anket katılımcısından %83'ünün suistimal ile karşılaştığını, işletmelerde karşılaşılan suistimal türlerinin fatura ve finansal tablolarda gerçekleştirilen hileler ile yolsuzluk (rüşvet, illegal hediyeler) üzerinde yoğunlaştığı, suistimale iten nedenlerin daha çok mali baskılar ile kontrol yetersizliğinden kaynaklanan fırsatların olduğu, suistimallerin tespit edilmesinde en etkili yöntemlerin ise sırasıyla iç denetim, ihbar hattı ve iç kontroller olduğu sonucuna ulaşılmıştır.

Kaban, Gül ve Eryılmaz (2018) tarafından yapılan çalışmada, risk yönetiminde kullanılan yöntemlerden suistimal belirtilerini gösteren kırmızı bayraklar ile risk değerlendirmesine ilişkin tanıtıcı bilgilere yer verilmiştir. Yanı sıra bankacılık sektöründe karşılaşılan zimmet ve kredi kartlarının kötüye kullanılması eylemleri ele alınarak örneklendirmeler yapılmış, suistimal olasılıkları ile etki düzeyleri belirlenmiş ve kırmızı bayraklar kişisel, yapısal ve operasyonel bazda tasnif edilerek hile üçgeninin baskı ve fırsat faktörü ile ilişkilendirilmiştir.

Çatıkkaş ve Çalış (2010), çalışmalarında suistimal kaynaklı kayıpları azaltmada kullanılabilecek proaktif yaklaşımlar olarak bilinen veri analitiği yöntemlerini detaylı bir şekilde ele almış, tümdengelim ve tümevarım metotlarından veri madenciliği ile Benford Yasası hakkında açıklamalara yer vererek satın alma faaliyetinde karşılaşılabilecek suistimallere örnekler vermiştir.

Kandemir ve Kandemir (2012), çalışmalarında muhasebe hilelerinin önlenmesi ve tespit edilmesinde kullanılan yöntemleri incelemiştir. Suistimaller ile mücadelede, geleneksel yöntemlerin yetersiz olduğuna ve veri madenciliği gibi çağdaş yöntemler ile birlikte kullanılması gerektiğine değinmiş, her ne kadar birlikte kullanımları maliyetleri artırsa da faydasının çok daha fazla olacağı sonucuna ulaşmışlardır.

Terzi (2012), büyük veriye sahip kurumlarda veri madenciliği metodu ile anormal işlemlerin ortaya çıkarılması konusunu ele almış, muhasebe denetim sürecinin hangi aşamasında hangi veri madenciliği yönteminin kullanılabileceğine ilişkin bilgiler vermiştir. Ayrıca veri madenciliği türlerinden olan karar ağacı yöntemine ve yöntemin avantaj ve dezavantajlarına ilişkin açıklamalarda bulunarak, bir örnek üzerinden karar ağacı algoritmasını görselleştirmiştir.

Yıldız ve Başkan (2014), muhasebe suistimallerinin önlenmesinde kullanılan bağımsız denetim, iç denetim, iç kontrol, tesadüfi, ihbar hattı ve özel suistimal inceleme birimleri gibi araçların ne derecede kullanıldığını ve en etkili olanını ortaya koymak amacıyla Borsa İstanbul kapsamındaki şirketler üzerinden inceleme yapmışlardır. Yapılan incelemelere göre, finansal tablo suistimalleri, varlıkların kötüye kullanılması ve haksız edinimi önlemede kullanılan yöntemlerin başında iç denetim, iç kontrol ve bağımsız denetimin geldiği, diğer yöntemlerin ise bunlara kıyasla daha az kullanıldığını belirtmişlerdir.

Çalış, Kayapınar ve Çetinyokuş (2014), çalışmalarında veri madenciliği, süreç ve modellerine değinerek, sınıflama ve regresyon modellerinden karar ağacı algoritmaları hakkında açıklamalarda bulunmuştur. Araştırmada, bilgisayar ve internet güvenliğine yönelik rastgele 300 kişiye anket çalışması uygulanarak; cinsiyet, yaş, eğitim durumu ve internet kullanım süresi gibi demografik özellikler üzerinden çıkarım yapılmıştır. Değişkenler için dört farklı karar ağacı algoritması uygulanmış, her biri için doğruluk oranları hesaplanmış olup, diğerlerine göre daha yüksek değere sahip olan C5.0 algoritmasının kullanılmasına karar verilmiştir.

Çamlıfidan (2019), araç kasko sigortalarındaki sahte hasarları tespit etmek amacıyla, 4159'u suistimal olmak üzere 45190 adet dosyada 20 adet ve istatistiksel olarak anlamlı olan 15 adet bağımsız değişken üzerinden makine öğrenmesi algoritmalarından Karar Ağacı, Lojistik Regresyon, K En Yakın Komşu ve Yapay Sinir Ağları ile model geliştirmiştir. Performans sonuçları değerlendirildiğinde, suistimal tespitinde 20 değişkenli modelde %91,2 doğruluk oranı ile Yapay Sinir Ağları algoritması diğerlerine göre daha iyi sonuç verirken, 15 değişkenli modelde ise %87 doğruluk oranı ile Karar Ağacı algoritması diğerlerine kıyasla daha iyi performans sergilemiştir.

Çalışmanın izleyen ikinci bölümünde, suistimal, suistimal üçgeni, bankacılık sektöründe karşılaşılan suistimaller ve bunlara yönelik düzenlenen mevzuatlar, suistimal tespit etme yöntemleri, makine öğrenmesi ve çalışmaya konu makine öğrenmesi yöntemleri hakkında genel bilgiler verilmiştir.

Üçüncü bölümde, uygulamaya yönelik değişkenler tanıtılmış ve değişkenlere ilişkin veriler analiz edilerek grafikler ile aktarılmıştır. Akabinde değişken seçiminde izlenen adımlara yer verilerek, makine öğrenmesi sınıflandırma algoritmaları ile modeller geliştirilmiş ve elde edilen performans sonuçları karşılaştırılmıştır.

Sonuç ve öneriler ile tez çalışmasının katkısına ise dördüncü bölümde değinilmiştir.

BÖLÜM 2: GENEL KAVRAMLAR

Çalışmanın bu bölümünde, suistimal ve suistimale ilişkin kavramlara, suistimal türlerine, suistimal tespit yöntemlerine, bankacılık sektöründe karşılaşılan suistimal eylemlerine, makine öğrenmesine ve çalışmaya konu makine öğrenmesi yöntemlerine değinilmiştir.

2.1 Suistimal

Yapılan araştırmalara göre, suistimal vakalarının başlamasından tespit edilmesine kadar geçen sürenin ortalama 12 ayı bulduğu, hatta etik hattı bulunmayan organizasyonlarda bu sürelerin ortalama 18 aya kadar uzadığı, araştırmaya konu 2110 adet vakada toplam kaybın 3.6 milyar doları aştığı ve vaka başına ortalama zararın 117.000 dolar olduğu, bununla birlikte zararın boyutunun da tespit edilme sürelerine paralel olarak arttığı ortaya konmaktadır (ACFE, 2022). Dolayısıyla kurumların buna yönelik riskleri belirleyip, yönetebilmeleri ve bundan kaynaklı kayıplarını azaltabilmeleri için suistimal kavramının doğru bir şekilde tanımlanması ve anlaşılması gerekmektedir.

2.1.1 Suistimal Kavramı

Günlük hayatta hile, sahtekârlık, güveni kötüye kullanma, zimmet, dolandırıcılık gibi birçok karşılığı bulunan suistimal (İngilizcede fraud) kavramının, farklı alanlarda yapılan tanımlamalarına ve açıklamalarına aşağıda yer verilmiştir.

Türk Dil Kurumu sözlüğünde suistimal, *"görev, yetki vb.ni kötüye kullanma"* şeklinde tanımlanırken, zaman zaman suistimal ifadesi yerine kullanılan hile terimi ise, *"birini aldatmak, yanıltmak için yapılan düzen, dolap, oyun, ayak oyunu, alavere dalavere, desise, entrika"* şeklinde tanımlanmıştır.[1]

Uluslararası Suistimal İnceleme Uzmanları Birliği (ACFE) suistimali, bir kazanç elde etmek için aldatmaya dayanan herhangi bir faaliyet olarak tanımlarken,[2] Uluslararası İç Denetim Enstitüsü (IIA) ise, *"hile, sahtekârlık, emniyeti kötüye kullanma ile nitelendirilebilecek hukuk dışı fiillerdir, suistimaller para, mal veya hizmet sağlamak, hizmet*

[1] Türk Dil Kurumu, https://sozluk.gov.tr/, (Erişim Tarihi: 20.09.2022)
[2] ACFE, https://www.acfe.com/fraud-101.aspx, (Erişim: 20.09.2022)

kaybından veya ödeme yapmaktan kaçınmak veya şahsıyla veya işle ilgili bir avantaj elde etmek amaçlarıyla çeşitli taraflar ve kurumlar tarafından gerçekleştirilebilir" şeklinde tanımlamıştır.[3]

Diğer taraftan, suistimal eylemi olarak değerlendirilen güveni kötüye kullanma, dolandırıcılık, zimmet, hırsızlık, banka ve kredi kartlarının kötüye kullanılması gibi fiiller ise suç kapsamında değerlendirilerek tanımına, unsurlarına ve karşılığı cezalarına kanunlarda yer verilmiştir.

Her ne kadar suistimal yerine birçok kavram kullanılıyor olsa da, bütün kelimelerin temelinde aldatma, dolandırma, haksız kazanç sağlama, kendisi veya başkası lehine bir menfaat (para, mal veya hizmet sağlama) temin etme ya da başkasının aleyhine olacak şekilde tasarrufta bulunma gibi hususlar yer almaktadır.

Görüldüğü üzere suistimal kavramı, oldukça kapsamlı bir tanıma sahip olmakla birlikte, hile, güveni/görevi kötüye kullanma, sahtekârlık, usulsüzlük, dolandırıcılık veya zimmet gibi farklı terimler ile de ifade edilebilmektedir. Araştırma her ne kadar genel olarak bankacılık sektöründe karşılaşılan güveni kötüye kullanma, zimmet, banka ve kredi kartlarının kötüye kullanılması gibi eylemler üzerinden yürütülmüş olsa da, kavram olarak suistimal ifadesi tercih edilmiştir. Diğer yandan, haksız bir menfaat elde etmek amacıyla kasıtlı yapılan suistimal eylemleri, yönetimden, üst yönetimden sorumlu olanlardan, çalışanlardan veya üçüncü taraflardan bir veya birden fazla kişi tarafından gerçekleştirilse de (Bağımsız Denetim Standardı 240, 11a), çalışmamızın konusunu bankacılık sektöründe yönetici ve çalışanlar tarafından gerçekleştirilen kurum içi suistimaller oluşturmaktadır.

2.1.2 Suistimal Türleri

Kurum ve kuruluşlar, faaliyet gösterdiği alana, hacmine, organizasyon yapısına, çalışan sayısına ve yapılan işin ne derecede teknolojiye bağlı olup olmadığına göre değişmekle birlikte zaman zaman farklı türlerde suistimallere maruz kalabilmektedirler.

[3]IIA, https://www.theiia.org/globalassets/documents/certifications/the-iia-official-glossary/official-iia-glossary-turkish.pdf, (Erişim: 20.09.2022)

ACFE, suistimali varlıkların kötüye kullanımı, finansal tablo suistimalleri ve yolsuzluk olmak üzere temel olarak üç farklı kategoride ele almaktadır. 2022 yılı raporuna göre; araştırmaya konu vakaların %86'sı varlıkların kötüye kullanılması şeklinde gerçekleşmiş olsa da ortalama kayıp tutarı ($100.000) açısından en düşük suistimal türüdür. Finansal tablolar üzerinde gerçekleştirilen suistimaller ise %9 ile en az rastlanan vaka türü olmakla birlikte, ortalama kayıp tutarı ($593.000) ile işletmeye en fazla zarar veren suistimal çeşididir. Üçüncü kategori olan yolsuzluk ise gerek karşılaşılan vaka sıklığı (%50) gerek ortalama tutar ($150.000) açısından diğer iki türün arasında yer almaktadır (ACFE, 2022).

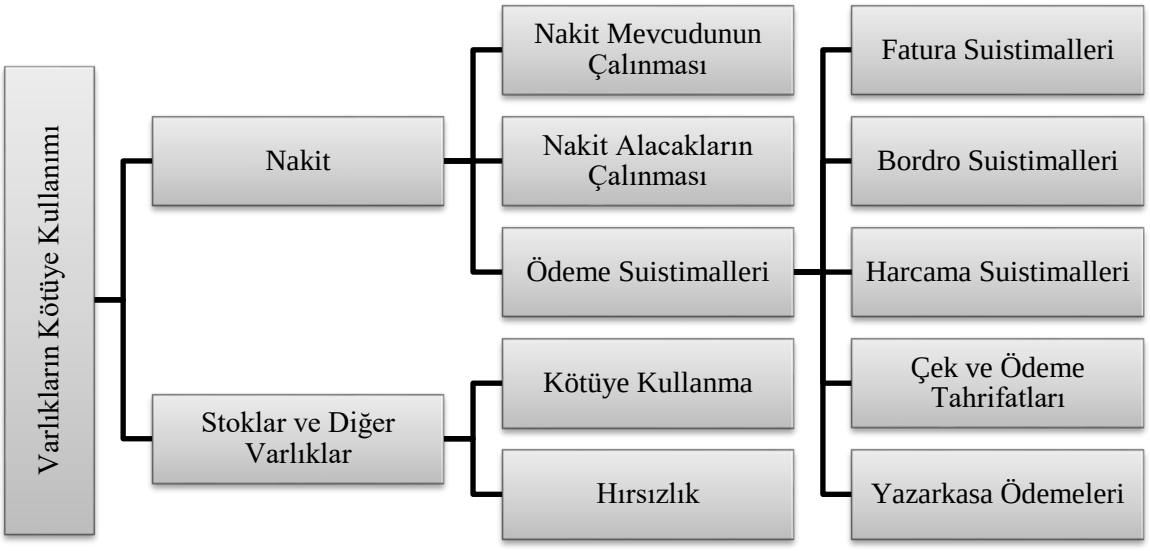
Şekil 2.1: Suistimal Türleri (ACFE, 2022)

2.1.2.1 Varlıkların Kötüye Kullanımı

Yöneticiler, çalışanlar ya da üçüncü kişiler tarafından, kişisel finansal problemlerinin veya kuruma olan sadakatsizliklerinin sonucu olarak işletmeye ait nakdin veya envanterinin çalınması veya işletmenin almadığı mal veya hizmet için harcama yapması şeklinde ortaya çıkan varlıkların kötüye kullanımı (Golden, Skala, & Clayton, 2006, 133), yolsuzluk ve finansal tablo suistimaline nazaran en sık karşılaşılan vaka türüdür (ACFE, 2022).

Varlıkların kötüye kullanımı; perakende sektöründe, işletmeye ait bir malın çalınması, fatura tahrifatı veya işletmenin borç ve alacaklarının suistimali vb. şeklinde karşımıza çıkabilirken, finans sektöründe ise bu durum çalışanın banka kasasından veya müşteri hesabından zimmetine para geçirmesi ya da kendisi veya bir başkası lehine mevzuata aykırı olduğunu bile bile kredi kullandırması suretiyle vuku bulabilmektedir. Görüldüğü üzere varlıkların kötüye kullanımı adı altında çeşitli yöntemler ile suistimaller işlenmektedir.

ACFE, varlıkların kötüye kullanımını, nakit ile stoklar ve diğer varlıklar olmak üzere iki ana kategoriye, söz konusu iki kategoriyi de kendi içerisinde alt gruplara ayırmıştır. Varlıkların kötüye kullanımı içerisinde, gerek sayıca en çok rastlanan gerek en yüksek ortalama kayba neden olan vaka türü fatura suistimali olup, karşılaşma sıklığı ve vermiş olduğu ortalama zarar dikkate alındığında, en düşük riski barındıran suistimal türü ise yazarkasa ödemeleri ile işlenendir (ACFE, 2022).



Şekil 2.2: Varlıkların Kötüye Kullanımı Suistimali (ACFE, 2022)

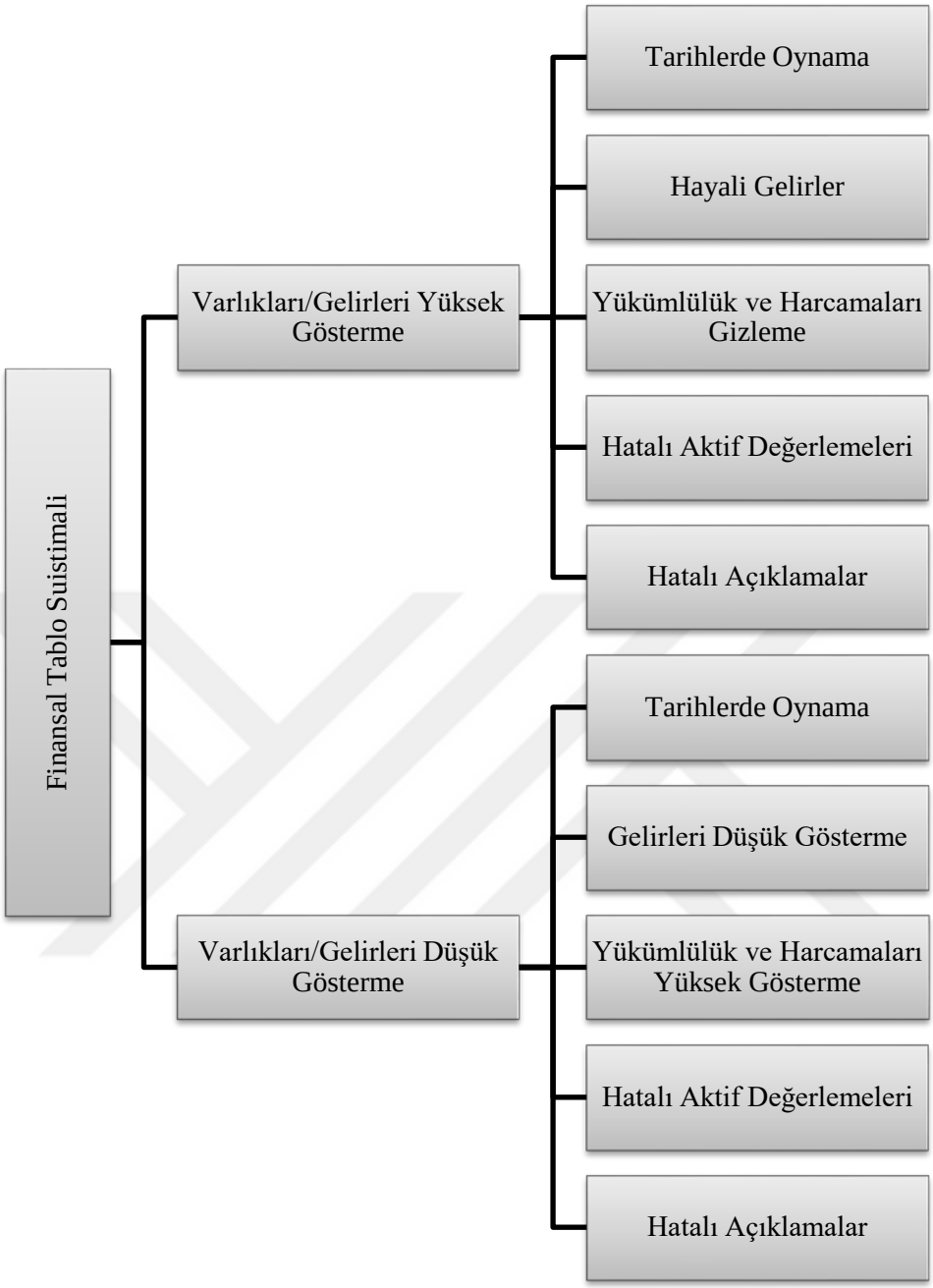
2.1.2.2 Finansal Tablo Suistimali

Finansal tablo suistimalleri, kurum çalışanı tarafından kasti olarak mali tablolarda yer alan hesap kayıtlarının veya rakamlara ilişkin açıklamaların gerçeğe aykırı olacak şekilde değiştirilmesi ya da kritik öneme sahip bilgilere yer verilmemesi neticesinde ilişkili tarafların aldatılması ile ortaya çıkan suistimal türüdür (Demir, Özdarak, & Sebilcioğlu, 2022, 13).

Kurum varlıklarının ya da yükümlülüklerinin olduğundan az veya çok gösterilmesi veya bu tür aldatıcı eylemlere konu belgeler üzerinde tahrifat gerçekleştirilmesi sonucunda vergi alacağından dolayı yalnızca devletin değil bunun yanı sıra şirket ortaklarının ve çalışanlarının, kullandırdığı/kullandıracağı krediden dolayı bankaların, hâlihazırda ticaret yapan ya da yapmayı planlayan şirketlerin veya söz konusu şirkete yatırım yapmayı düşünen bireysel yatırımcılar gibi kadar birçok farklı kesimin bu durumdan etkilenmesi ve kararlarını manipüle edilmiş bilgi ve belgelere göre vermesi söz konusudur.

Şirkete ait finansal tabloların doğru ve eksiksiz bir şekilde hazırlanıp sunulması sorumluluğu üst yönetimde (Kaya & Uzay, 722) olduğu için, finansal tablo suistimalleri varlıkların kötüye kullanımından farklı olarak genellikle üst yönetim tarafından yapılmaktadır (Çalıyurt & Tezgel 2015, 38).

Her ne kadar diğerlerine kıyasla sayıca en az karşılaşılan vaka türü olsa da işletmeye en fazla zararı dokunan finansal tablo suistimali, ACFE tarafından varlıkları/gelirleri yüksek gösterme ve varlıkları/gelirleri düşük gösterme olmak üzere iki ana kategoriye, söz konusu iki kategori de kendi içerisinde alt gruplara ayrılmıştır (ACFE, 2022).



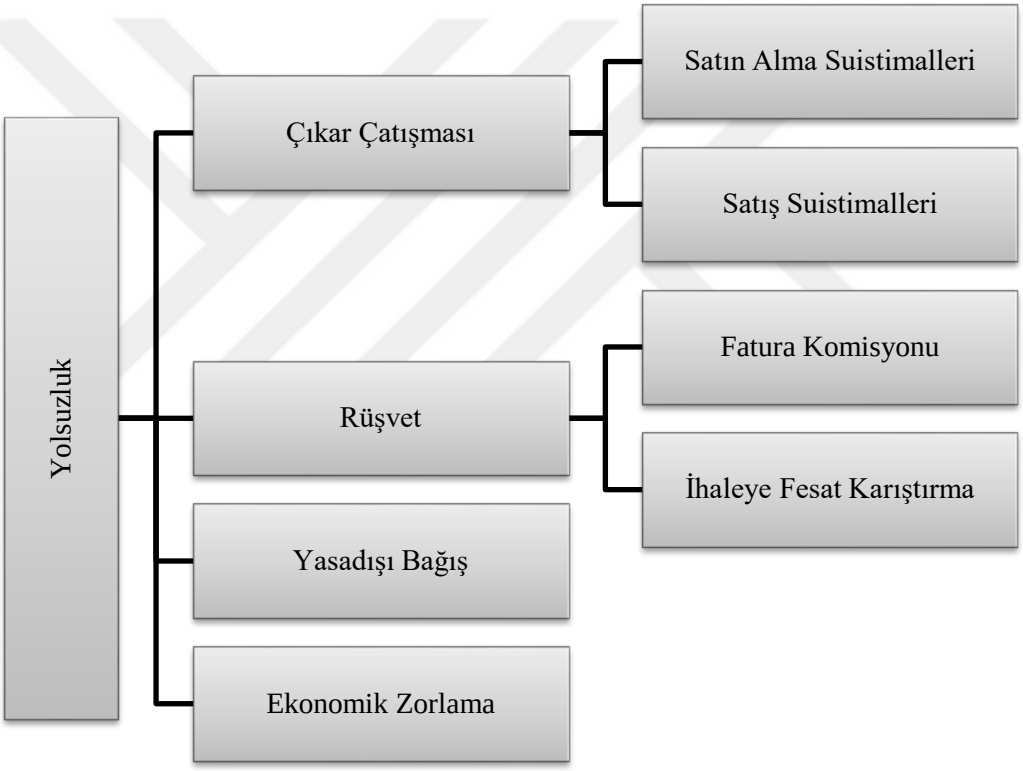
Şekil 2.3: Finansal Tablo Suistimali (ACFE, 2022)

2.1.2.3 Yolsuzluk

Yolsuzluk, çalışanların bireysel veya kolektif bir şekilde (Timofeyev, 2015, 630) kendisi veya bir başkası lehine doğrudan veya dolaylı (ACFE, 2022) çıkar elde etmek amacıyla güveni veya görevini kötüye kullanma eylemidir (Vona, 2008, 120). Bu etik dışı kazanım, bir ihaleyi kazanmak veya yasal bir zorunluluktan kaçmak için kabul edilen bir rüşvet veya yasadışı bir bağış olabilir (Roy, 2001, 16). Yolsuzluk eylemini gerçekleştiren kişi, kendisinin,

tedarikçilerden bir avantaj elde etmesinin ötesinde işletmenin satın almış olduğu mal ve hizmetler için de daha fazla ödeme yapmasına sebebiyet vermektedir (Golden, Skalak, & Clayton, 2006, 158).

Büyük işletmelerde (çalışan sayısı 100'den fazla) daha yaygın olan yolsuzluk, karşılaşma sıklığı (%50) ve işletmeye vermiş olduğu zararın büyüklüğü ($150.000) açısından varlıkların kötüye kullanımı ile finansal tablo suistimalleri arasında konumlanmakta olup, kendi içerisinde çıkar çatışması, rüşvet, yasadışı bağış ve ekonomik zorlama olmak üzere dört ana kategoride gruplandırılmıştır (ACFE, 2022).

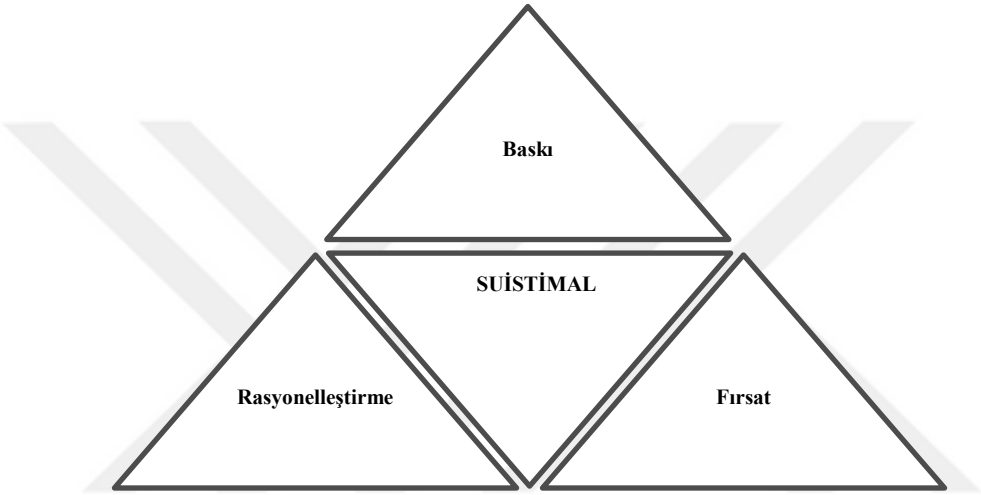


Şekil 2.4: Yolsuzluk (ACFE, 2022)

2.2 Suistimal Üçgeni

Bir suistimalin gerçekleştirilmesinin altında yatan psikolojik sebepleri, en iyi şekilde, suç konusunda birçok kitabı ve kriminoloji alanında uzman biri olan Donald R. Cressey açıklamıştır. Cressey, insanların neden suç işlediği üzerine çalışmalarını yoğunlaştırmış ve

bunun arkasında üç kritik etmenin olduğu sonucuna ulaşmıştır. Söz konusu üç unsurun ise baskı (pressure), fırsat (opportunity) ve rasyonelleştirme (rationalization) olduğunu belirtmiş ve failin suç işleme dürtüsünün oluşabilmesi için bu üç unsurun birlikte vuku bulması gerektiğini ileri sürmüştür (Sánchez vd., 2018). Zamanla Cressey'in ortaya koymuş olduğu bu bulguların üzerine araştırmalar ilerletilmiş ve suistimal üçgeni (fraud triangle), suistimal ölçeği (fraud scale) ve suistimal elması (fraud diamond) gibi modeller türetilmiştir (Coenen, 2008, 10; Dorminey vd., 2012; Albrecht, 2014). Her ne kadar farklı düşünce ve görüşler ileri sürülmüş olsa da suistimal üçgeni, bu alanda popülerliğini devam ettirmektedir (Sandhu, 2016).



Şekil 2.5: Suistimal Üçgeni (Sánchez ve diğ, 2018)

2.2.1 Baskı

Suistimal üçgeninin ilk unsuru olan baskı, suistimal eyleminin gerçekleştirilmesinin arkasındaki motivasyon ve dürtüye ilişkin bir düşüncedir (Sánchez vd., 2018). Diğer taraftan, kişinin çok istediği bir şeye, imkânsızlıktan dolayı sahip olamamasından kaynaklı bir finansal ihtiyaç da denilebilir (Coenen, 2008, 11). Dolayısıyla suistimalin ilk ateşleyicisi olan baskı unsuru, bazı çalışmalarda teşvik veya motivasyon (motivation) olarak da geçmektedir.

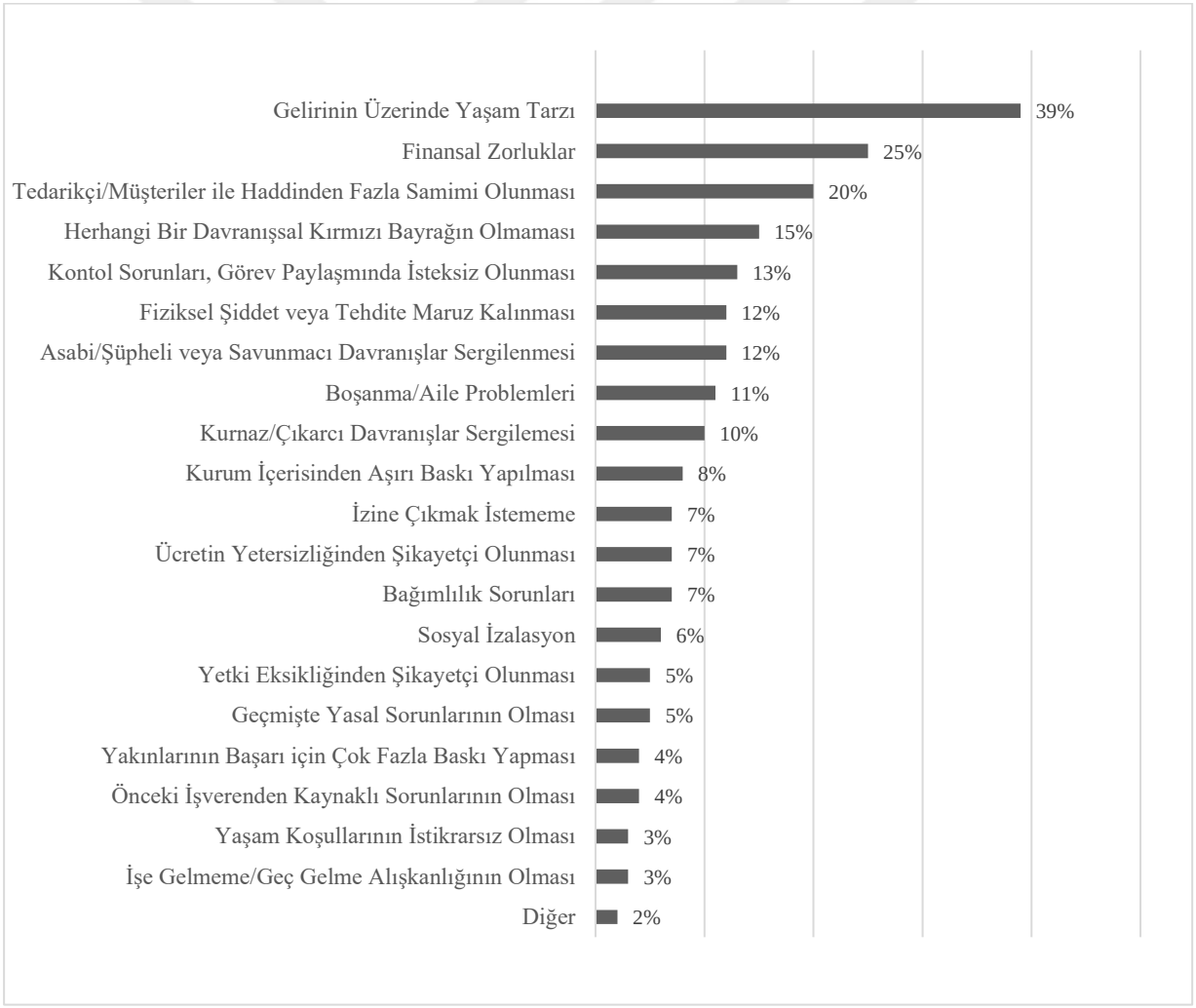
Baskılar, organizasyonun veya işverenin politika ve tutumundan kaynaklanabileceği gibi, çalışanın özel hayatına yönelik içerisinde bulunduğu finansal veya psikolojik durumlardan da kaynaklanabilmektedir.

Organizasyon kaynaklı baskılar, işletmelerin içinde bulunduğu sektöre göre değişmekle birlikte genel olarak pazarlama, satış veya karlılığa yönelik rakamsal hedeflere ulaşılmasından, piyasa değerinin yükseltilmesine ilişkin paydaş beklentisinden, çalışma koşullarından,

yıldırmaya maruz kalınmasından, performansa dayalı ücret/prim veya terfi politikasından kaynaklı baskılar şeklinde karşımıza çıkabilmektedir.

Çalışanın özel hayatına yönelik baskılar ise, yüksek hayat standartlarında yaşama hevesinden, gelirine oranla daha yüksek seviyelerde borçlanma eğiliminden, yetersiz gelire sahip olunduğu düşüncesinden, borçlanma limitlerinin sınırda olmasından, kumar ve bahis gibi yasa dışı kötü alışkanlıklardan, tehdit, fiziksel şiddet veya şantaja maruz kalınmasından, işsiz kalma korkusundan, kişilik bozukluğu ya da kleptomani gibi hastalık kaynaklı olabilmektedir.

ACFE tarafından 2022 yılında yapılan araştırmaya göre, suistimal gerçekleştiren çalışanlarda, gelirinin üzerinde yaşam tarzı, finansal zorluklar ile tedarikçiler veya müşteriler ile haddinden fazla samimi olunması gibi baskıyı oluşturan unsurların davranışsal kırmızı bayraklar arasında ilk üç sırayı aldığı görülmektedir (ACFE, 2022).



Şekil 2.6: Çalışanlarda Görülen Davranışsal Özellikler (ACFE, 2022)

2.2.2 Fırsat

Suistimalin gerçekleştirilebilmesi için yetersiz veya etkili olmayan iç kontroller, denetim ve gözetim eksikliği, iş ve işlemlerin karmaşıklığı, etik kuralların olmaması veya görev ayrımı ile yetki ve rollerin belirsizliği gibi fırsatların var olması ve failin de olayı gerçekleştirmesi sonucunda herhangi bir ceza ile karşılaşmayacağına veya karşılaşsa bile etkisinin önemsiz olacağına inanması gerekir (Golden vd., 2016). Fırsatlar kontrol ortamının olmadığı veya yetersiz olduğu durumlarda ortaya çıkar ki, bu suistimali yapan kişi için yakalanma olasılığı düşük olduğunda çok daha çekici olabilmektedir (Mui & Mailley, 2015). Ayrıca, diğer çalışanlara veya yöneticilere nazaran işe erken gelinmesi ve işten geç çıkılması, terfi/atama ya da işten ayrılmanın reddedilmesi durumları da suistimali gerçekleştirmek için gerekli olan fırsat ortamını oluşturabilir (Sandhu, 2016).

Kurumsal yapılar büyüyüp daha karmaşık hale geldikçe ve üst yönetime daha fazla personel ile birçok birimin sorumluluğu verildikçe, bununla birlikte çalışanlara da daha fazla varlığa erişim yetkisi verilmektedir ve bu da suistimallerin daha kolay işlenmesi için ortamı uygun hale getirmektedir (Coenen, 2008, 12). Diğer yandan, günümüzde bazı iş modelleri, görevler, işlemler veya varlıklar ise doğası gereği fırsatlar barındırabilmekte ve bunlar diğerlerine kıyasla manipüle edilmeye ya da amacı dışında kullanıma daha elverişli olabilmektedir (Golden vd., 2016).

2.2.3 Rasyonelleştirme

Suistimal üçgeninin son unsuru olan rasyonelleştirme, failin, gerçekleştirmiş olduğu suistimalin bir suç olmadığını, etik veya vicdani olarak kabul edilebilir bir durum olduğunu meşrulaştırma çabasıdır ve bu kişiler genellikle kendilerini kurallara uyan ve iyi bireyler olarak görürler (Kramer, 2015). Olayın arkasında yatan asıl nedenleri gizleyerek veya inkâr ederek, davranış ve tutumlarını toplumsal normlara göre daha makul olacak şekilde bir mantığa büründürmeye ve kendilerini bu duruma inandırmaya çalışırlar.

Bu kişiler, herhangi bir varlığı zimmetlerine geçirmelerinin aslında haklarının olduğunu, yöneticilerinin veya çalışma arkadaşlarının kurallara uymama durumlarının kendileri için de geçerli olabileceğini düşünürler. Diğer yandan, finansal olarak zorluk çeken ya da gelirinin üzerinde yaşam tarzı olan ve bunların üzerinde baskı oluşturması ile kimi birey, çaldığı para veya para yerine çeken varlıkları, durumu düzeldiği vakit iade edeceğini ileri sürerek, bu

durumu kendince akla uygun hale getirmekte ve vicdanını rahatlatmaktadır. Bazıları ise, gerçekleştirmiş olduğu suistimal eylemlerinin hiçbir şekilde suç unsuru ve gayriahlaki bir durum olmadığını düşünmektedir (Emir, 2008).

2.3 Bankacılık Sektöründe Karşılaşılan Suistimaller

Bankacılık suçları, genel olarak 5411 sayılı Bankacılık Kanununda düzenlenmekle birlikte, bazı bankacılık suçları ise 5237 sayılı Türk Ceza Kanunu, 5464 sayılı Banka Kartları ve Kredi Kartları Kanunu, 5941 sayılı Çek Kanunu, 1211 sayılı Türkiye Cumhuriyet Merkez Bankası Kanunu gibi diğer kanunlarda da yer alabilmektedir.

Bankacılık sektöründe en sık karşılaşılan suistimal türlerinden güveni kötüye kullanma, dolandırıcılık, banka veya kredi kartlarının kötüye kullanılması ve zimmet suçları, aynı zamanda çalışmamıza konu suistimal türlerini oluşturduğu için yalnızca bu suçların tanımına yer verilmiştir.

Güveni kötüye kullanma, 5237 Sayılı Türk Ceza Kanunu'nun 155. Maddesinde düzenlenmiş olup, söz konusu düzenlemeye göre güveni kötüye kullanma; *"başkasına ait olup da, muhafaza etmek veya belirli bir şekilde kullanmak üzere zilyedliği kendisine devredilmiş olan mal üzerinde, kendisinin veya başkasının yararına olarak, zilyedliğin devri amacı dışında tasarrufta bulunmaktır"* (5237 Sayılı TCK, m.155).

Dolandırıcılık suçu, 5237 Sayılı Türk Ceza Kanunu'nun 157. Maddesine göre; *"hileli davranışlarla bir kimseyi aldatıp, onun veya başkasının zararına olarak, kendisine veya başkasına bir yarar sağlamaktır"* (5237 Sayılı TCK, m.157).

Banka veya kredi kartlarının kötüye kullanılması suçu, 5237 Sayılı Türk Ceza Kanunu'nun 245. Maddesinde düzenlenmiş olup, söz konusu düzenlemede; *"(1) başkasına ait bir banka veya kredi kartını, her ne suretle olursa olsun ele geçirmek veya elinde bulundurmak, kart sahibinin veya kartın kendisine verilmesi gereken kişinin rızası olmaksızın bunu kullanarak veya kullandırtarak kendisine veya başkasına yarar sağlamak, (2) başkalarına ait banka hesaplarıyla ilişkilendirilerek sahte banka veya kredi kartı üretmek, satmak, devretmek, satın almak veya kabul etmek, (3) sahte oluşturulan veya üzerinde sahtecilik yapılan bir banka veya kredi kartını kullanmak suretiyle kendisine veya başkasına yarar sağlamak"* şeklinde ifade edilmektedir (5237 Sayılı TCK, m.245).

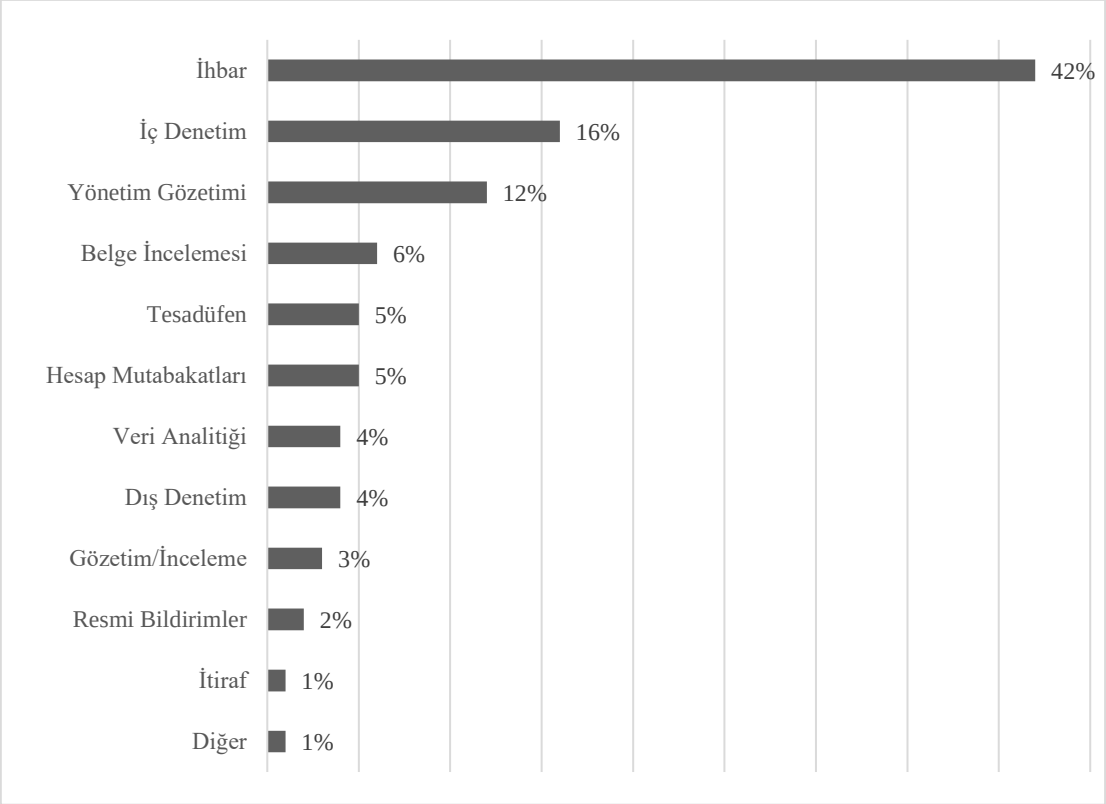
Zimmet suçu, hem Türk Ceza Kanunu'nda hem de Bankacılık Kanunu'nda düzenlenmiş olup, Bankacılık Kanunu'nun 160. Maddesine göre; *"görevi nedeniyle zilyetliği kendisine devredilmiş olan veya koruma ve gözetimiyle yükümlü olduğu para veya para yerine geçen evrak veya senetleri veya diğer malları kendisinin ya da başkasının zimmetine geçirmektir"* (5411 Sayılı BK, m.160).

Özetle, bankacılık suistimali, banka yöneticisinin veya çalışanının, görevi nedeniyle ya da muhafaza etmek veya kullanmak üzere kendisine emanet edilen banka veya müşterilerine ait bilgi, belge, para, kredi ve kredi benzeri ürünleri veya para yerine geçen diğer şeyleri kendisine veya bir başkasına fayda sağlamak ya da başkasını zarara uğratmak gayesiyle görev ve yetkisini kötüye kullanması şeklinde tanımlanabilir. Dolayısıyla bir eylem veya işlemin bankacılık suistimali olarak sınıflandırılabilmesi için, suistimalin banka personeli tarafından bankaya veya üçüncü bir kişiye karşı yapılmış olması, kendisi veya bir başkası lehine ya da karşı tarafın aleyhine olacak şekilde doğrudan veya dolaylı olarak tasarrufta bulunulması ve tüm bu hususların gerçekleşebilmesi için de eylemin bilerek ve isteyerek gerçekleştirilmesi diğer bir deyişle kasıt unsurunu içermesi gerektiği sonucuna ulaşılabilir.

2.4 Suistimallerin Tespitinde Kullanılan Yöntemler

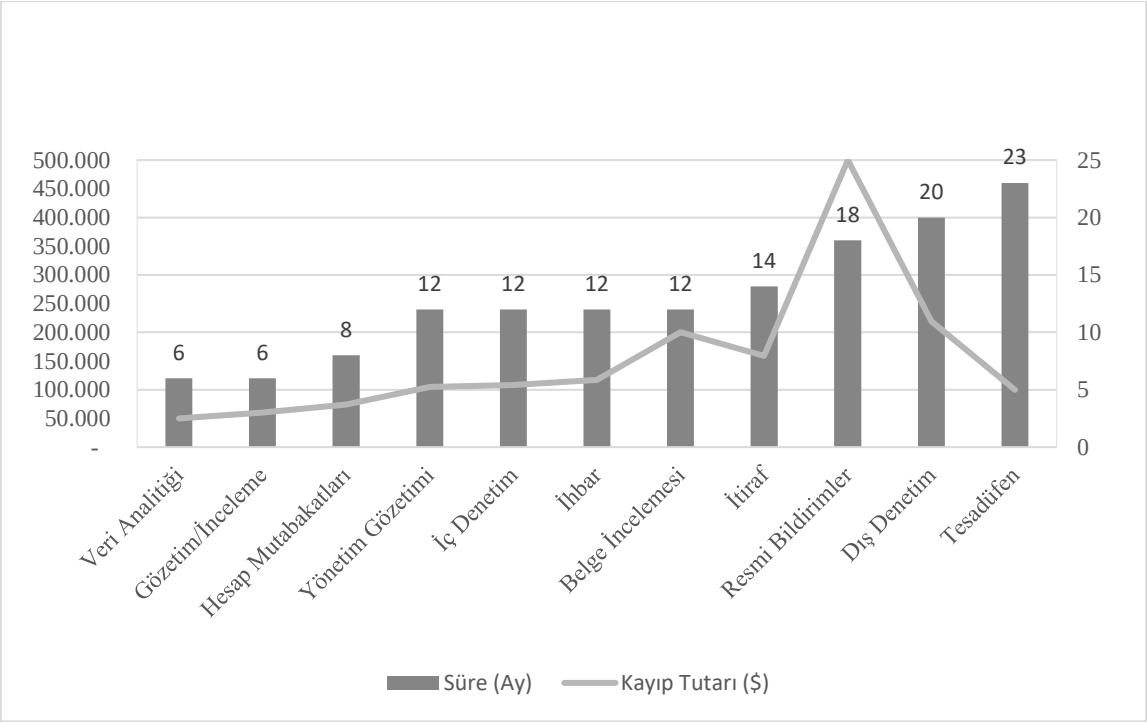
Suistimallerin işletmeye vereceği maddi veya manevi zarar, çok büyük boyutlara ulaşmadan ortaya çıkarılması ve bunun için gerekli aksiyonların alınması önem arz etmektedir. İşte bu noktada kurumların kendine özgü yapısı dikkate alınarak, suistimallerin nasıl tespit edilebileceğine yönelik yöntemlerin geliştirilmesi gerekmektedir. Yapılan araştırmalara göre, suistimal vakalarının başlamasından tespit edilmesine kadar geçen sürenin ortalama 12 ayı bulduğu, hatta bünyesinde etik hattı gibi uygulamalar olmayan organizasyonlarda bu sürelerin ortalama 18 aya kadar uzadığı ve bununla birlikte zararın boyutunun da tespit edilme sürelerine bağlı olarak arttığı ortaya konmaktadır (ACFE, 2022). Bununla birlikte, faillerin de teknolojideki gelişmelere paralel daha karmaşık yapıda suistimaller gerçekleştirdiği dikkate alındığında, suistimallerin geleneksel yöntemler ile tespit edilmesinin güçleştiği söylenebilir. Bu yüzden kurumların, suistimalleri erkenden tespit etmeleri, risk odaklı ve bilgi teknolojileri destekli modern yaklaşımları benimseyerek ve yöntemlerini sürekli geliştirerek faaliyetlerini sürdürmeleri gerekmektedir.

ACFE'nin önceki çalışmalarında olduğu gibi 2022 yılında yapmış olduğu araştırmada da suistimallerin tespitinde en çok kullanılan yöntemin açık ara farkla ihbar olduğu görülmektedir. Araştırmaya konu vakaların tespitine yönelik yapılan ihbarların %55'i kurum içi personel tarafından, %34'ü ise müşteriler, tedarikçiler, rakip firmalar ve şirket sahipleri gibi paydaşlar tarafından gerçekleştirilmiştir. Diğer taraftan, suistimallerin ortaya çıkarılmasında ikinci en sık kullanılan yöntemin ise iç denetim birimleri tarafından yapılan tespitler olduğu görülmektedir (ACFE, 2022).



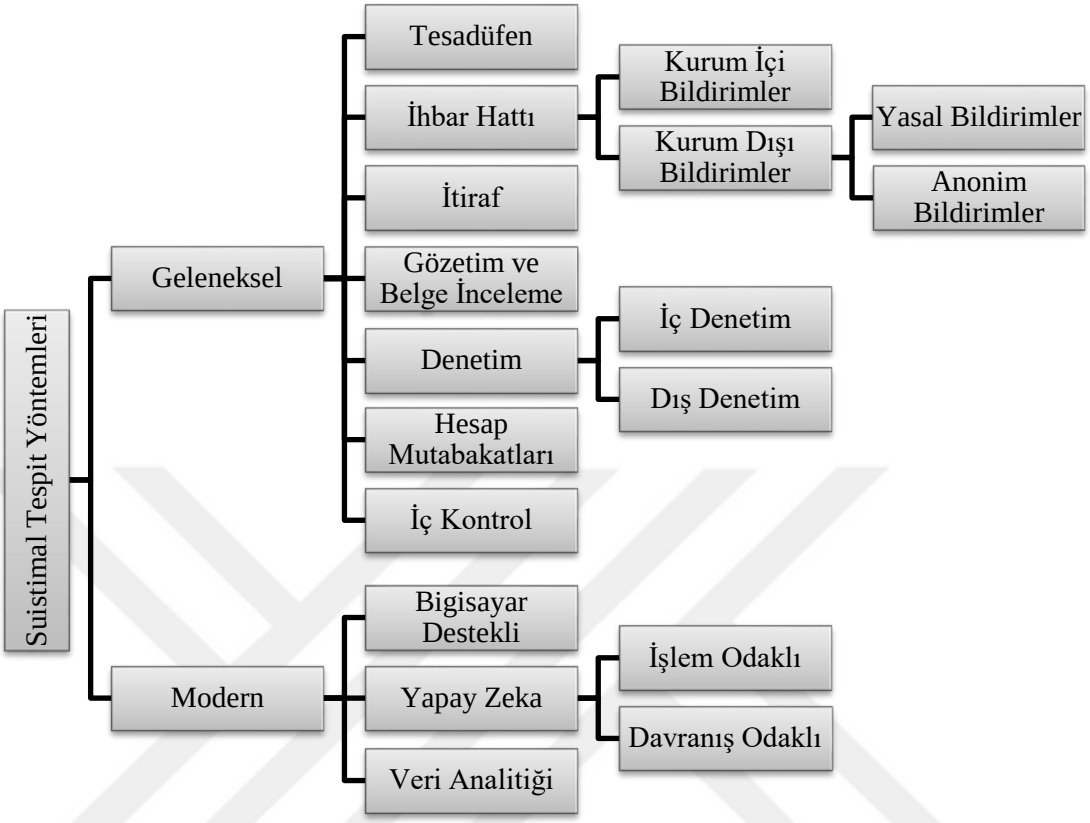
Şekil 2.7: Suistimal Tespit Yöntemleri (ACFE, 2022)

Tespit edilme süresi ile ortalama kayıp tutarı açısından söz konusu yöntemler kıyaslandığında, en etkili yöntemlerin sırasıyla veri analitiği(data monitoring), gözetim/inceleme, hesap mutabakatları, yönetim gözetimi, iç denetim, ihbar ve belge inceleme olduğu, diğer yöntemlerin ise tespit etme süreleri ile ortalama kayıp tutarlarının daha yüksek olduğu görülmektedir. Aşağıdaki grafikte de görüldüğü üzere, veri analitiği ile gözetim/inceleme yöntemleri, diğerlerine nazaran suistimali daha erken sürede ortaya çıkardığı için kurumların maruz kaldığı zararın boyutu da çok yükseklere ulaşamamaktadır.



Şekil 2.8: Suistimal Tespit Yöntemlerinin Etkinliği (ACFE, 2022)

Söz konusu çalışmaya göre ihbar, suistimalleri belirleme hususunda her ne kadar en yaygın kullanılan yöntem olsa da, veri analitiği, gözetim/inceleme ve hesap mutabakatı yöntemleri kadar etkin olmadığı sonucuna ulaşılmaktadır. Bununla birlikte, ihbar hattı olmayan kurumların, olanlara kıyasla %100 daha fazla zarar ile karşılaştığı ve tespit süresinin %50 daha uzun olduğu görülmektedir (ACFE, 2022).



Şekil 2.9: Suistimal Tespit Yöntemleri (Bozkurt, 2009, 173)

Suistimallerin ortaya çıkarılmasında kullanılabilecek yöntemleri genel olarak geleneksel ve modern yöntemler olarak ikiye ayırabiliriz. Modern yöntemler, suistimallerin tespiti için devamlı olarak kullanılırken, geleneksel yöntemler ise genel itibarıyla daha pasif bir yaklaşım sergilemektedir (Bozkurt, 2009, 172). Modern yöntemler daha çok, yalnızca suistimale odaklanan ve sürekli olarak ortaya çıkarılması için mücadele eden kurum içerisinde organize olmuş suistimal, etik veya iç denetim birimleri tarafından kullanılmaktadır.

Suistimal ile mücadelede üç husus bulunmaktadır. Birincisi, suistimal risklerinin yönetilmesi sonucu proaktif bir yaklaşım sergilenerek suistimalin gerçekleşmeden önlenmesi; ikincisi gerçekleşen vakaların tespit edilmesi; üçüncüsü ise ortaya çıkarılan suistimale ilişkin soruşturma faaliyetinin yürütülmesidir (Yılmaz, 2019). Söz konusu süreç, gerek kurum bünyesinde yürütülen faaliyetlere gerek iç kontrol sisteminin çalışma prensibine hakim olması sebebiyle genellikle iç denetim birimleri tarafından yürütülmektedir.

Uluslararası İç Denetim Enstitüsü (IIA) tarafından yayımlanan terimler sözlüğüne göre iç denetim; bir organizasyonun iş ve işlemlerine katma değer sağlamak amacıyla, kurumun iç kontrol, risk yönetimi ve yönetişim süreçlerinin etkinlik ve yeterliliğini değerlendiren güvence, danışmanlık ve suistimal inceleme faaliyetidir.[4] Suistimale ilişkin iç denetçilerden beklenenler ise, Enstitü tarafından yayımlanan Uluslararası İç Denetim Standartlarında belirtilmektedir:[5]

- Madde 1210.A2'ye göre; iç denetçiler, suistimal risklerini ve bu risklerin kurum içerisinde nasıl yönetilebileceğine ilişkin yeterli bilgiye sahip olmak zorundadır.

- Madde 1220.A1'e göre; iç denetçiler, görevlerini icra ettiği sırada suistimal ihtimalini de dikkate alarak mesleğin gereğini yerine getirmek zorundadır.

Ülkemizde başta bankacılık ve finans sektöründe faaliyet gösteren kurumlar olmak üzere, genellikle teftiş kurulu, iç denetim veya suistimal inceleme birimleri tarafından yerine getirilen suistimal tespitine yönelik çalışmalarında, teknoloji destekli sürekli denetim kapsamında veri analitiğine her geçen gün daha fazla önem vermektedir. Denetim faaliyetlerine ve suistimal risklerinin ortaya çıkarılmasına yönelik kullanmış oldukları veri analitiği araçları ise, genellikle veri tabanı üzerinden sorgulamaların ve veri madenciliği yöntemlerinin kullanılabildiği, üretilen senaryoların analizinin yapılabildiği, trend, sapan gözlem ve regresyon gibi istatistiki analizlerin yapılabildiği uygulamalardır.

2.5 Veri Analitiği

Verilerin elde edilmesi, işlenmesi ve incelenmesi süreçlerini kapsayan, aynı zamanda büyük veri olarak tanımlanan yüksek hacimli verilerden taraflar için anlamlı ve değerli bilgilerin keşfedilmesi sürecinin bir parçası olan veri analitiği, gerçek zamanlı içgörüleri esas alarak veri odaklı karar alma mekanizmalarını desteklediği gibi gelişimini de sağlamaktadır (Yılmaz, 2019; Gandomi & Hadier, 2015; Soileau vd., 2015; Kelleher vd., 2015, 36). Veri madenciliği ise veri analitiğinin bir alt kırılımı olup, büyük veri içerisinden çözüme yönelik bilgilerin elde edilebilmesi için kullanılan bir süreçtir (Özmen, 2007; Yılmaz, 2019).

[4] IIA, https://www.theiia.org/globalassets/documents/certifications/the-iia-official-glossary/official-iia-glossary-turkish.pdf, Erişim: 28.09.2022, 19:21

[5] IIA, https://www.theiia.org/globalassets/site/standards/mandatory-guidance/ippf/2017/ippf-standards-2017-turkish.pdf, Erişim: 28.09.2022, 19:21

Veri analitiği, yöntem ve yaklaşım açısından dört grupta incelenmektedir (Geng 2017, 331):

a) **Betimsel Analitik:** Analizin henüz başlarında, veri örüntüsünü daha iyi anlayabilmek için verilerin sınıflandırılması, özetlenmesi ve anlamlı bir hale getirilmesidir. "Ne/Ne oldu" sorularına odaklanmaktadır. Örneğin; şirket için sadık müşteriler kimdir veya geçmiş yıllara nazaran bu yıl gerçekleştirilen suistimallere ilişkin kayıp oranı ne düzeydir?

b) **Tespit Edici Analitik:** Sebep sonuç ilişkisini belirleme noktasında tercih edilen bir yaklaşımdır ve "Neden oldu" sorusu ile ilgilenmektedir. Örneğin; firmanın satışları neden düştü veya önceki yıl vuku bulan suistimallerde neden artış gerçekleşti?

c) **Çıkarımsal Analitik:** Betimsel ve tespit edici analitiğin her ikisinin de kullanılması ile geleceğe yönelik öngörüde bulunulması söz konudur. Geçmiş bilgiler baz alındığında gelecekte "Ne olması muhtemel, trend ne yönde olacak" sorularına yönelik çıkarım yapılması ile ilgilenmektedir. Örneğin; geçmişte karşılaşılan suistimallerin örüntüsü ortaya çıkarılarak, gelecekte olası vakaların belirlenmesi.

d) **Kuralcı Analitik:** Yapay zekâ ve makine öğrenmesi gibi tekniklerin aracılığıyla geleceğe yönelik önerilerin belirlenmesinde kullanılır. Yazılım ve donanım alanındaki teknolojik gelişmelere paralel olarak her geçen gün daha da geliştirilmektedir ve "Ne yapılmalı" sorusuna odaklanmaktadır. Örneğin; yapay zekâ tabanlı sistemin, müşteri hizmetlerini arayan kişinin konuşmasını analiz ederek suistimal olup olmadığını değerlendirmesi ve buna yönelik aksiyon alması.

2.5.1 Makine Öğrenmesi

Yapay zekâ, matematik, istatistik, psikoloji, karmaşıklık teorisi gibi birçok disiplinden beslenen ve her geçen gün önemi artarak devam eden makine öğrenmesi, veriler üzerinden desenler çıkararak tahminlerde bulunan ve sürekli olarak performansını geliştiren bilgisayar modellemesi olarak tanımlanabilir. Birçok farklı alanda uygulanabilir olduğu kanıtlanmış olmakla birlikte, başta sınıflandırma problemlerinde olmak üzere genellikle çıkarımsal veri analitiği uygulamalarında model geliştirmek için kullanılmaktadır. İyi bir öğrenme modeli

oluşturabilmek, eğitim veri setinin kaliteli olmasına ve performans ölçütlerinin iyi belirlenmiş olmasına bağlıdır (Mitchell, 1997, 2-17; Kelleher vd., 2015, 39-40; Mohri vd., 2018, 1).

2.5.1.1 Makine Öğrenmesi Yöntemleri

Makine öğrenmesi yöntemleri, veri setinin türüne, boyutuna ve problemin yapısına göre farklılaşmakta olup (Mohri vd., 2018, 6; Brownlee, 2017, 16), söz konusu yöntemler temelde 3 başlık altında incelenmektedir (Joshi, 2020, 10):

i. Denetimli Öğrenme (Supervised Learning)

ii. Denetimsiz Öğrenme (Unsupervised Learning)

iii. Pekiştirmeli Öğrenme (Reinforcement Learning)

i. Denetimli Öğrenme

Çıkarımsal veri analitiğinde en sık kullanılan yöntem olan denetimli öğrenme, geçmiş verilerden, bilinen örneklerden veya birbiri ile ilişkisi bulunan özelliklerden oluşan veri setini esas alarak, bağımsız değişken ile hedef değişken arasındaki ilişkiyi fonksiyon olarak modeller ve bilinmeyeni öngörmeye çalışır (Kelleher vd., 2015, 39; Mohri vd., 2018, 6). Sınıflandırma ve regresyon olmak üzere iki türü bulunmakta olup (Çelik ve Altunaydin, 2018), hedef değişkenin veri türü kantitatif ise regresyon, kalitatif (kategorik) ise sınıflandırma modelini tercih eder (Osmanoğlu, 2021). Makine öğrenmesi algoritmalarının birçoğu denetimli öğrenme metodolojisini kullanır (Brownlee, 2017, 16).

Çalışmamıza konu uygulamada da denetimli öğrenme yöntemi kullanılarak sınıflandırma algoritmalarından faydalanılmıştır. Bu kapsamda, daha öncesinde suistimal yapan ve yapmayan çalışanlara yönelik bilinen değişkenlerden oluşan eğitim veri setinden bir örüntü çıkarıp, suistimalci olup olmadığı bilinmeyen kişiler hakkında sınıflandırma yapabilen makine öğrenmesi algoritmaları ile model geliştirilmiştir.

ii. Denetimsiz Öğrenme

Denetimsiz öğrenme, bir eğitim seti olmadan veriler arasındaki ilişkileri, veri setinin yapısını ve dağılımını ortaya çıkarmaya çalışan makine öğrenmesi yöntemidir. Bir öğretici veya

bir girdi değişkeni olmadığı gibi model sonucunda doğru veya yanlış olabilecek bir çıktıdan da söz edilemez (Brownlee, 2017, 17; Çelik ve Altunaydin, 2018; Joshi, 2020, 11; Bruce, vd., 2020, 283). Dolayısıyla bu yöntemde elde edilen sonuçların sayısal olarak performansını ölçmek de pek mümkün görünmemektedir (Mohri vd., 2018, 6).

Denetimsiz öğrenme yaklaşımlarından kümeleme, veriler arasında ilişki kurarak sınıflandırılmamış veri setini anlamlı gruplara ayırmak için kullanılırken, boyut indirgeme ise, verilerin daha yönetilebilir hale getirmek için kullanılır (Mohri vd., 2018, 6; Bruce, vd., 2020, 283).

iii. Pekiştirmeli Öğrenme

Denetimli ve denetimsiz öğrenmeden farklı olarak bu yöntemde, eğitim veri seti ile test veri seti iç içe geçmiş durumda olup, sürekli olarak çevre ile etkileşim içindedir (Mohri vd., 2018, 7; Joshi, 2020, 11).

Farklı durumlar karşısında, her seferinde istenilen davranışı sergileyebilmek için en uygun yolu minimum hata ile bulmaya çalışan bu öğrenme metodu, aslında bir ödüllendirme ve cezalandırma sistemi üzerine kuruludur (Çelik ve Altunaydin, 2018; Mohri vd., 2018, 7).

2.5.1.2 Sınıflandırma Algoritmaları

Bu başlık altında yalnızca tez çalışmasında kullanılan Lojistik Regresyon (Logistic Regression), Karar Ağacı (Decision Tree), Rastgele Orman (Random Forest), Naive Bayes ve Destek Vektör Makinesi (Support Vector Machine) algoritmalarına yer verilecektir.

i. Lojistik Regresyon

Sınıflandırma problemlerinde sıkça kullanılan Lojistik Regresyon, adını içinde barındırdığı lojistik fonksiyonundan diğer bir deyişle sigmoid fonksiyonundan almaktadır. Bu yöntemde bağımlı değişken, 0 ile 1 arasında bir olasılık değeri alır. Ancak, bu değer hiçbir zaman 0 veya 1 değeri olmaz (Joshi, 2020, 38; Osmanoğlu, 2021), Olasılık değeri 0 ve 1 değerlerine yakınsadıkça Lojistik Regresyon grafiğinde düzleşme görülür (Başar & Genç, 2020).

Doğrusal Regresyon modelinde bağımlı değişken sürekli bir değer alırken, Lojistik Regresyon modelinde ise "0/1, Doğru/Yanlış, Evet/Hayır, Başarılı/Başarısız" gibi kategorik değerler almaktadır (Osmanoğlu, 2021). Buna karşın bağımsız değişkenlerin veri türünün sürekli/kesikli veya kategorik olup olmaması konusunda herhangi bir kısıtlama bulunmamaktadır. Dolayısıyla, çıktı değişkeninin kategorik olması haricinde, Doğrusal Regresyon modelinden bir farkı yoktur (Joshi, 2020, 38).

Model, ,bağımlı değişkenin kategori sayısına göre ikili (Binary) ve Çoklu (Multinominal) Lojistik Regresyon olarak ikiye ayrılmaktadır (Özdemir, 2021).

ii. Karar Ağacı

Denetimli makine öğrenmesi yöntemlerinden karar ağacı, etiketlenmiş veri setinin sınıflandırılmasında performansı yüksek algoritmalardan biridir. Oldukça esnek bir yapıya sahip olan teknik, büyük veri setinde değişkenler arasındaki bağlantı üzerinden bir dizi kurallar oluşturur ve bu şekilde öğrenmeyi gerçekleştirerek daha küçük kümelere bölerek ilerler (Çelik ve Altunaydin, 2018; Bruce, vd., 2020, 237).

Hastalıkların teşhisi, kullanıcı davranışlarının tespit edilmesi, suistimal faillerinin belirlenmesi ve kredi risk değerlendirmesi gibi birçok farklı alanda kullanılan bu yöntem, hem sınıflandırma hem de regresyon problemlerinde oldukça etkili sonuçlar üretmektedir (Barros, vd., 2015, 7).

Karar ağacı algoritması, soru ve cevaplarına göre yukarıdan aşağıya doğru çıkarımlarda bulunarak ve bu süreçte karar düğümleri, dallar ve bunlara bağlı yapraklar oluşturarak modeli oluşturur (Joshi, 2020, 53). Almış olduğu şekil itibarıyla ağaca benzetildiği için, adını buradan almaktadır.

iii. Rastgele Orman

Makine öğrenmesi içerisinde en sık kullanılanlardan ve performansı en yüksek olanlardan biri olan Rasgele Orman algoritması (Brownlee, 2017, 127), karar ağacı algoritmasının aksine rastgele değerlere göre birden fazla karar ağacı oluşturur. Modelin sonucunu, N tane karar ağacının kendi arasında oylanması sonucunda en yüksek oyu alan karar ağacı belirler (Osmanoğlu, 2021).

Sınıflandırma ve regresyon modellerinin her ikisinde de kullanılabilen bu teknik (Geren, 2020), büyük ve dengesiz veri setlerinde daha yüksek performans sergilemektedir (Osmanoğlu, 2021). Ayrıca, sürekli ağaçların öğrenme yöntemini güncelleyerek daha düşük korelasyonlu ve daha güçlü çıkarımlar elde etmeye çalışır (Brownlee, 2017, 129).

iv. Naive Bayes

Değişkenler arasındaki güçlü bağımsızlık varsayımı ile sonuca etki eden her bir parametrenin olasılık değerini hesaplayarak sınıflandırma yapan, uygulaması kolay, ancak üstün performans gösteren ve bayes teoremini esas alan bir denetimli öğrenme tekniğidir (Friedman vd., 1997; Brownlee, 2017, 83; Çelik ve Altunaydin, 2018).

v. Destek Vektör Makinesi

Makine öğrenmesi algoritmaları içerisinde uygulama açısından en etkili olan Destek Vektör Makinesi (Mohri vd., 2018, 79), hata bazlı öğrenmeye dayalı bir yaklaşım sergileyerek (Kelleher vd., 2015, 408), dağılımı belirsiz olan bir veri seti üzerinde öğrenmeyi gerçekleştirir (Güneren, 2015; Osmanoğlu, 2020).

Denetimli öğrenme yöntemini kullanan bu algoritmanın amacı, doğrusal çizgiler ve hiper düzlemler aracılığıyla sınıflara ait verileri birbirinden optimal şekilde ayırmaya çalışmaktır (Turfan, 2020). Bu ayırma işlemini gerçekleştirirken de, çizeceği doğrunun farklı sınıflara ait verilere olan uzaklığını, maksimum düzeyde tutmaya çalışmaktadır (Osmanoğlu, 2020).

Başlangıçta iki sınıflı hiper düzlem yapısını kullanarak iki sınıflı doğrusal problemlerin çözümü için tasarlanmış olsa da (Joshi, 2020, 65), daha sonrasında lineer olmayan problemlerin çözümünde de kullanılmaya başlanmıştır (Cortes vd., 1995).

2.5.1.3 Performans Değerlendirme Ölçütleri

Bu başlık altında, sınıflandırma algoritmalarının ne kadar başarılı ve ne kadar doğru çıkarımlarda bulunduğuna yönelik performans metriklerine değinilecektir.

Performans değerlendirme ölçütlerine geçmeden önce, sınıflandırmalara yönelik gerçek değerler ile tahmini değerlerin karşılaştırıldığı Karmaşıklık Matrisine (Confusion Matrix) yer verilecektir (Han vd., 2012, 364-369).

Detaylarına Tablo 2.1'de yer verilen matrisin, satır değerleri gerçek sınıflandırmaya, sütun değerleri ise tahmin edilen sınıflandırmaya karşılık gelmektedir.

Tablo 2.1: Karmaşıklık Matrisi

		TAHMİN EDİLEN SINIF	
GERÇEK SINIF		**1**	**0**
	1	**Doğru Pozitif, DP** (True Positive, TP)	**Yanlış Negatif, YN** (False Negative, FN)
	0	**Yanlış Pozitif, YP** (False Positive, FP)	**Doğru Negatif, DN** (True Negative, TN)

- **Doğru Pozitif (DP) :** 1 olarak sınıflandırılmış gözlemlerin, 1 olarak tahminlendiğini gösterir.

- **Yanlış Negatif (YN) :** 1 olarak sınıflandırılmış gözlemlerin, 0 olarak tahminlendiğini gösterir.

- **Doğru Negatif (DN) :** 0 olarak sınıflandırılmış gözlemlerin, 0 olarak tahminlendiğini gösterir.

- **Yanlış Pozitif (YP) :** 0 olarak sınıflandırılmış gözlemlerin, 1 olarak tahminlendiğini gösterir.

Uygulamamıza konu sınıflandırma algoritmalarının başarı ve performans durumları karşılaştırılırken, modellerin Doğruluk (Accuracy), Duyarlılık (Recall/Sensitivity), Kesinlik (Precision) ve F-Skor metrikleri üzerinden değerlendirme yapıldığından, yalnızca söz konusu performans ölçütlerine yer verilmiştir.

Tablo 2.2: Performans Kriterleri (Han vd., 2012, 365)

Performans Kriteri	**Açıklama**	**Formül**
Doğruluk	Gözlemlerin ne kadarının doğru olarak sınıflandırıldığını gösterir	$\frac{DP + DN}{DP + YP + DN + YN}$
Duyarlılık	Gerçek sınıf değeri 1 olanların ne kadarının doğru olarak tahmin edildiğini gösterir	$\frac{DP}{DP + YN}$
Kesinlik	1 olarak tahmin edilenlerin ne kadarının doğru olarak tahmin edildiğini gösterir	$\frac{DP}{DP + YP}$
F-Skor	Duyarlılık ve kesinlik ölçülerinin harmonik ortalamasıdır.	$\frac{2 * Kesinlik * Duyarlılık}{Kesinlik + Duyarlılık}$

2.6 Sentetik Azınlık Örnekleme Tekniği (Synthetic Minority Oversampling Technique, SMOTE)

Veri setinin dengesiz bir şekilde dağılması, hedef değişkenin her bir sınıfındaki veri sayısının birbirinden farklı olması anlamına gelmektedir. Bu durum, makine öğrenmesi algoritmalarının performansını olumsuz yönde etkileyerek, veri sayısı çoğunlukta olan sınıfa daha fazla meyilli olabilmesine sebebiyet verebilmektedir (Chicco & Jurman, 2020, 2). Dolayısıyla, makine öğrenmesi uygulamasına geçilmeden önce, veri kümeleri arasındaki dengesiz dağılımın giderilmesine yönelik yöntemlere başvurulması önem arz etmektedir (Yavaş vd., 2020; Osmanoğlu, 2021).

Sentetik Azınlık Örnekleme Tekniği (SMOTE), azınlık sınıfının veri sayısı sentetik örnekler ile artırılarak, diğer sınıfın veri sayısı ile denkleştirilmesi yöntemidir. Bu işlem, örnekleri doğrudan aynıları ile çoğaltmak yerine, azınlık sınıfındaki verilerin k en yakın komşuları (K-Nearest Neighbors) baz alınarak yeni örnekler oluşturulması şeklinde gerçekleştirilmektedir (Chawla vd., 2002).

2.7 K-Katlı Çapraz Doğrulama (K-Fold Cross Validation)

Makine öğrenmesi modellerinin performansının ölçümünde en sık kullanılan yöntemlerden olan k-katlı çapraz doğrulamada, veri seti, örnek sayısı eşit olacak şekilde k adet alt gruba ayrılır. Her seferinde, söz konusu k adet alt kümenin (k-1) adedi eğitim, 1 adedi de test için kullanılır ve bu işlem k kez tekrarlanır. Dolayısıyla, bu yöntem ile her bir örnek, hem eğitim veri setinde hem de test veri setinde yer alarak çapraz doğrulama gerçekleştirilmiş olur (Görgel, 2011; Turfan, 2020).

2.8 Veri Analitiği Programı KNIME

KNIME (The Konstanz Information Miner), veri ön işlemeden, makine öğrenmesi algoritmalarına kadar birçok veri analitiği araçlarını bünyesinde bulunduran, birbiri ile etkileşimli veri akışı uygulaması imkânı sunan bir platformdur. Program, düğümler (node) aracılığıyla kullanılmakta olup, kullanıcı dostu arayüze sahip olduğu gibi zengin görselleştirme araçlarına da sahiptir. Yanı sıra, eğitim ve araştırma gibi birçok alanda kullanılabilen KNIME, kullanıcılar tarafından geliştirilmesine ve test edilmesine de olanak sağlayan modüler yapıda, açık kaynak kodlu ve ücretsiz kurulabilen bir yazılımdır (Berthold vd., 2008, 319-325).

BÖLÜM 3: UYGULAMA

Bu bölümde, uygulamaya yönelik belirlenen değişkenler analiz edilmiş ve suistimal yapanlar ile yapmayanlara yönelik veriler karşılaştırılarak, suistimalcilere ilişkin profil ortaya konmuştur. Akabinde modele dahil edilecek değişkenlerin seçiminde izlenen adımlara yer verilerek, makine öğrenmesi sınıflandırma algoritmaları ile modeller geliştirilmiş ve elde edilen sonuçlar karşılaştırılmıştır.

3.1 Değişkenler ve Verilerin Analizi

Literatür çalışmalarında genellikle suistimal yapan kişilerin yaşı, cinsiyeti, eğitim düzeyi ile pozisyonuna/unvanına yönelik bilgiler üzerinden profillerine ilişkin çıkarımlarda bulunulduğu görülmektedir. Çalışmamızda ise söz konusu değişkenleri de kapsayacak şekilde demografik (4 adet), mesleki (7 adet) ve finansal (12 adet) bilgilere yönelik olmak üzere toplam 23 adet değişken belirlenmiştir.

Tablo 3.1: Değişkenler

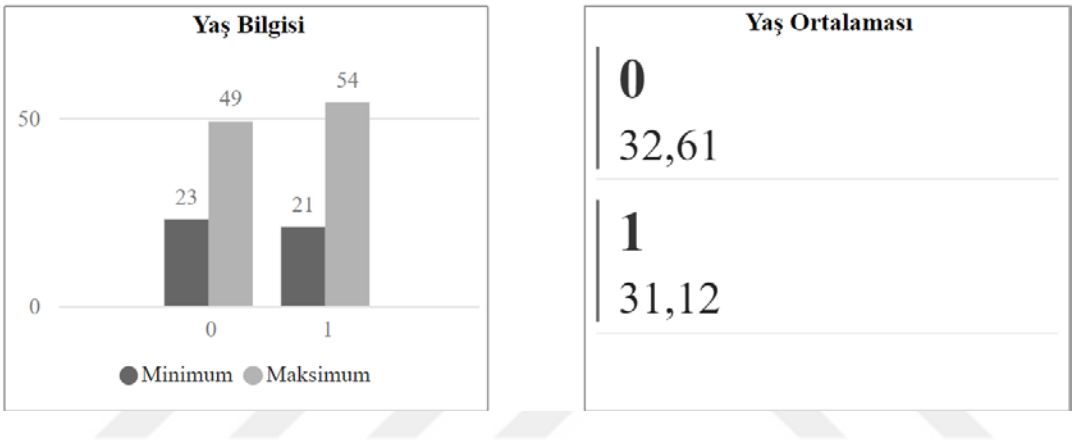
	Değişken	Kategori	Açıklama
1	Yaş	Demografik	Çalışanın Yaş Bilgisi
2	Cinsiyet	Demografik	Çalışanın Cinsiyeti
3	Medeni Durumu	Demografik	Çalışanın Medeni Durumu
4	Eğitim Düzeyi	Demografik	Çalışanın Eğitim Bilgisi (Lise, Önlisans, Lisans, Yüksek Lisans)
5	Pozisyon	Mesleki	Çalışanın Pozisyon Bilgisi (Memur, Yetkili, Yönetmen/Müdür Yardımcısı)
6	Bankada Çalışma Süresi	Mesleki	Çalışanın Bankadaki Toplam Çalışma Süresi (Ay)
7	Birimde Çalışma Süresi	Mesleki	Çalışanın En son Çalıştığı Birimdeki Tecrübe Süresi (Ay)
8	Görevde Çalışma Süresi	Mesleki	Çalışanın En Son Çalıştığı Görev Grubundaki Tecrübe Süresi (Ay)
9	Banka Dışı Deneyim	Mesleki	Çalışanın Banka Haricinde Tecrübesinin Olup Olmadığı Bilgisi

10	Toplam Çalışma Süresi	Mesleki	Çalışanın Banka ve Banka Dışı Toplam Tecrübe Süresi (Ay)
11	İzin Kullanım Oranı	Mesleki	Çalışanın Kullanmış Olduğu Ücretli İzinlerin Tahakkuk Eden Toplam İzin Süresine Oranı (%)
12	KKB Skor	Finansal	Çalışanın KKB AŞ Tarafından Oluşturulan Findeks Notu (1-1900)
13	Aktif Ürün Adedi	Finansal	Çalışanın Sorgulama Tarihinde Aktif Olan Ürün (Kredi, Kredi Kartı, Kredi Mevduat Hesabı) Bilgisi (Adet)
14	Kredi Başvuru Adedi	Finansal	Çalışanın Sorgulama Tarihinde Başvuru Süreci Devam Eden Kredi Bilgisi (Adet)
15	Çalışılan Kurum Sayısı	Finansal	Çalışanın Sahip Olduğu Ürünlere İlişkin Birlikte Çalıştığı Finansal Kurum Sayısı (Adet)
16	Kredi Kartı Doluluk Oranı	Finansal	Çalışanın Tüm Finansal Kurumlar Nezdindeki Kredi Kartlarının Limit-Kullanım Oranı (%)
17	Kredi Kartı Gecikmesi	Finansal	Çalışanın Tüm Finansal Kurumlar Nezdindeki Kredi Kartlarının Toplam Borç Bakiyesi (TL)
18	KMH Kullanım Oranı	Finansal	Çalışanın Tüm Finansal Kurumlar Nezdindeki Kredili Mevduat Hesaplarının Limit-Kullanım Oranı (%)
19	İhtiyaç Kredisi	Finansal	Çalışanın Tüm Finansal Kurumlar Nezdindeki İhtiyaç Kredisi Bakiyesi (TL)
20	İhtiyaç Kredisi Gecikmesi	Finansal	Çalışanın Tüm Finansal Kurumlar Nezdindeki İhtiyaç Kredisi Gecikme Bakiyesi (TL)
21	Konut Kredisi	Finansal	Çalışanın Tüm Finansal Kurumlar Nezdindeki Konut Kredisi Bakiyesi (TL)
22	TL Vadesiz Hesap Bakiyesi	Finansal	Çalışanın Çalıştığı Bankadaki TL Vadesiz Mevduat Hesabı ile Kredili Mevduat Hesabının Soruşturma/Sorgulama Tarihinden Önceki Bir Aylık Ortalama Bakiyesi
23	YP Vadesiz Hesap Bakiyesi	Finansal	Çalışanın Çalıştığı Bankadaki Döviz Vadesiz Mevduat Hesabının Soruşturma/Sorgulama Tarihinden Önceki Bir Aylık Ortalama Bakiyesi
24	Suistimal	-	Çalışanın Suistimal Gerçekleştirip Gerçekleştirmediği Bilgisi (0: Suistimal Gerçekleştirmemiş, 1: Suistimal Gerçekleştirmiş)

3.1.1 Çalışanların Yaşı

ACFE tarafından yapılan araştırmaya göre, suistimal gerçekleştirenlerin %68'si 31 ile 50 yaş aralığında, KPMG Türkiye ile TEİD'in yapmış olduğu çalışmaya göre ise %90'nı 26-45 yaş aralığındadır.

Uygulamamıza konu suistimal gerçekleştirenlerin yaşlarının 21 ile 54 arasında değişiklik gösterdiği ve söz konusu çalışmalar ile benzer sonuçlar ile karşılaşıldığı görülmektedir. Bununla birlikte, suistimal gerçekleştirmeyenlerin yaşları ise 23 ile 49 arasında olup, her iki grubun da yaş ortalaması birbirine yakın seyretmektedir.

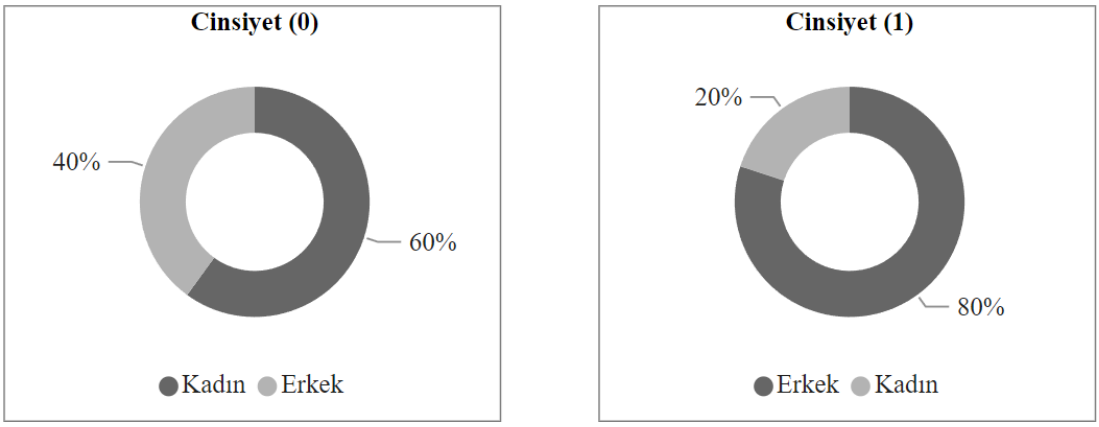


Şekil 3.1: Yaş Bilgileri

3.1.2 Çalışanların Cinsiyeti

ACFE tarafından yapılan araştırmaya göre, suistimal gerçekleştirenlerin %73'ünün cinsiyeti erkek iken, KPMG Türkiye ile TEİD'in yapmış olduğu çalışmaya göre ise %92'si erkektir.

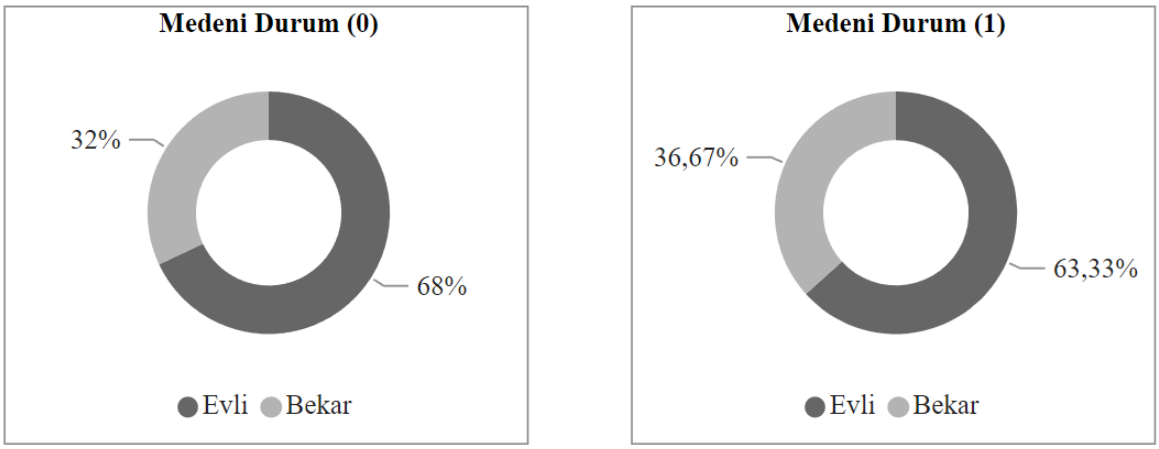
Uygulamamıza konu suistimal gerçekleştirenlerin %80'inin, gerçekleştirmeyenlerin ise %40'ının erkek olduğu görülmektedir.



Şekil 3.2: Cinsiyet Bilgileri

3.1.3 Çalışanların Medeni Durumu

Çalışanların medeni durumlarına bakıldığında, suistimal eylemini işleyenlerin %63,33'ünün, işlemeyenlerin ise %68'inin evli olduğu sonucuna ulaşılmıştır.

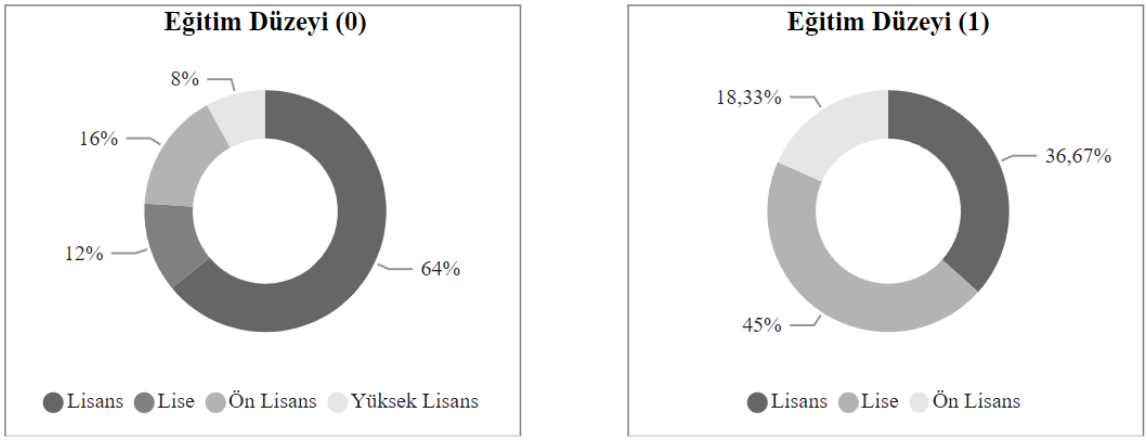


Şekil 3.3: Medeni Durumları

3.1.4 Çalışanların Eğitim Düzeyi

ACFE (2022) tarafından yapılan araştırmaya göre, suistimal gerçekleştirenlerin %20'si lise ve daha düşük bir eğitim düzeyine sahip iken, Durkaya (2008)'nın çalışmasına göre ise, suistimalcilerin %59'u lise mezunudur.

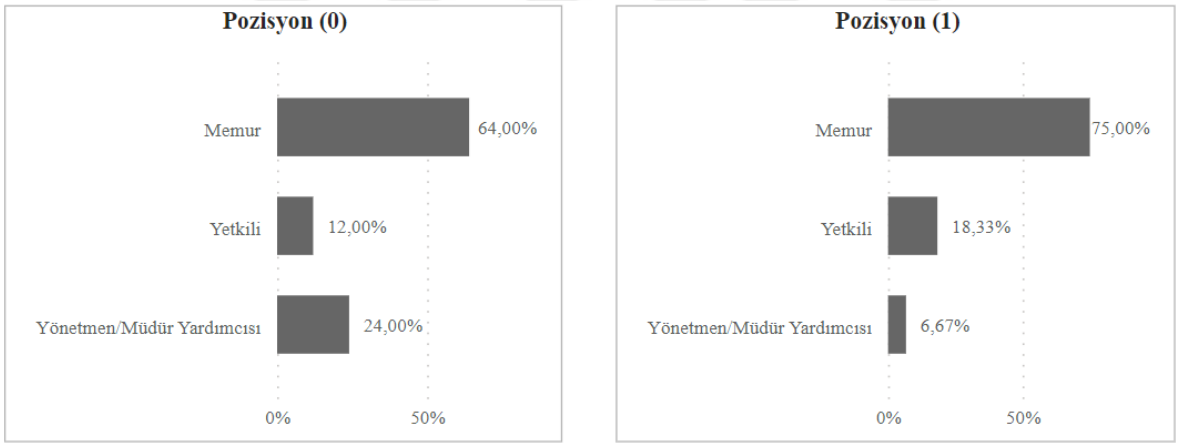
Araştırmamıza konu suistimal gerçekleştiren kişilerin %45'i lise mezunu iken, suistimal gerçekleştirmeyenlerin ise yalnızca %12'si lise mezunudur.



Şekil 3.4: Eğitim Düzeyleri

3.1.5 Çalışanların Pozisyonu

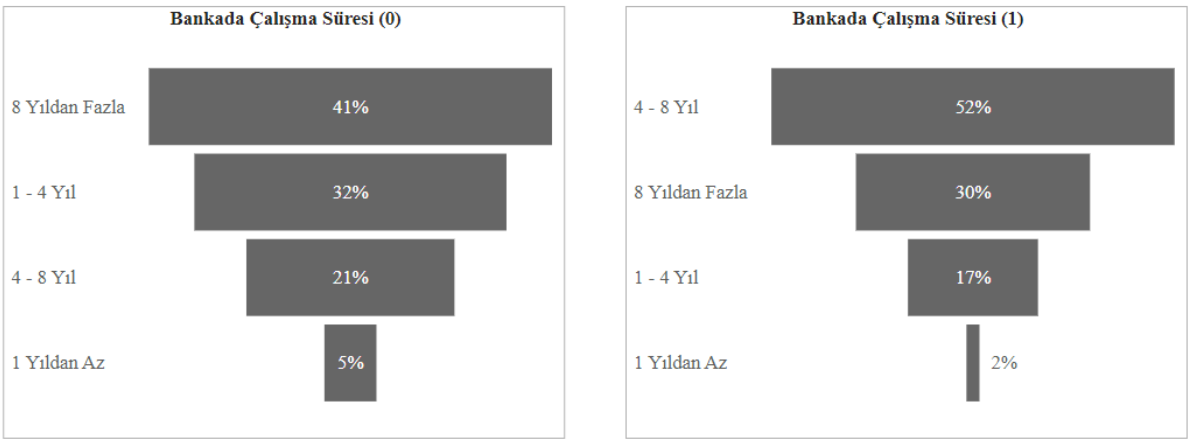
ACFE'ye göre vakaların %37'si, KPMG ve TEİD'e göre ise %49'u yönetici pozisyonun altında çalışan personel tarafından gerçekleştirilmiştir. Çalışmamıza konu suistimalcilerin ise %93'ü yönetici pozisyonu altında çalışmaktadır.



Şekil 3.5: Pozisyon Bilgileri

3.1.6 Çalışanların Bankada Çalışma Süresi

ACFE'ye göre suistimalcilerin %47'sinin 1 ile 5 yıl arasında, %45'inin 6 yıldan fazla hizmet süresi bulunmaktadır. KPMG ve TEİD raporuna göre ise %44'ünün 1 ile 4 yıl arasında ve %54'ünün ise 5 yıl ve üstü bir süredir mağdur olan şirkette çalıştığı görülmektedir. Söz konusu raporlara paralel, çalışmamıza konu suistimalcilerin de çoğunun kıdemli olduğu ve %82'sinin ise 4 yıldan fazla bir süredir bankada çalıştığı görülmektedir.



Şekil 3.6: Bankada Çalışma Süresi

3.1.7 Çalışanların Birimde Çalışma Süresi

Personelin en son çalıştığı birimdeki hizmet süreleri kıyaslandığında; her ne kadar iki grupta da 1 ile 4 yıl arasında çalışanların oranı birbirine yakın olsa da, suistimalcilerin %42'sinin 4 yıldan fazla süredir suistimal yaptığı birimde çalıştığı görülmektedir. Diğer yandan, suistimal eylemine karışmayanların %33'ünün en son çalıştığı birimdeki görev süresinin 1 yıldan daha az olduğu sonucuna ulaşılmıştır. Dolayısıyla, özellikle şubelerde çalışan personelin belirli sürelerle zorunlu olarak rotasyona tabi tutulması gibi bir takım politikaların suistimallerin engellemesinde veya ortaya çıkarılmasında etkili olduğu söylenebilir.

Şekil 3.7: Birimde Çalışma Süresi

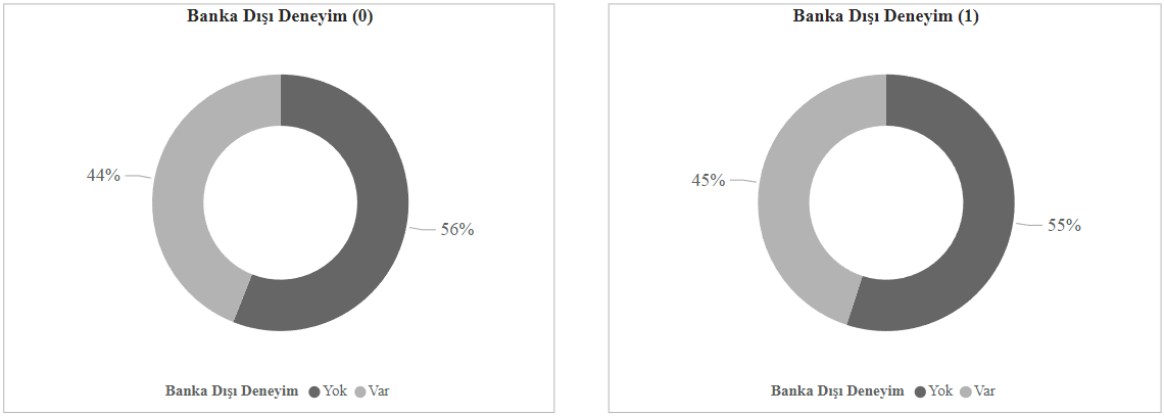
3.1.8 Çalışanların Görevde Çalışma Süresi

Suistimal yapan kişilerin %44'ü, 4 yıldan daha uzun bir süredir suistimal yaptığı esnadaki görevinde bulunurken, suistimal gerçekleştirmeyenlerde bu oran %21'dir. Yanı sıra, suistimalcilerin %5'inin ise bulundukları görev grubundaki tecrübe süresinin1 yıldan az olduğu görülmektedir.

Şekil 3.8: Görevde Çalışma Süresi

3.1.9 Çalışanların Banka Dışı Deneyim Bilgisi

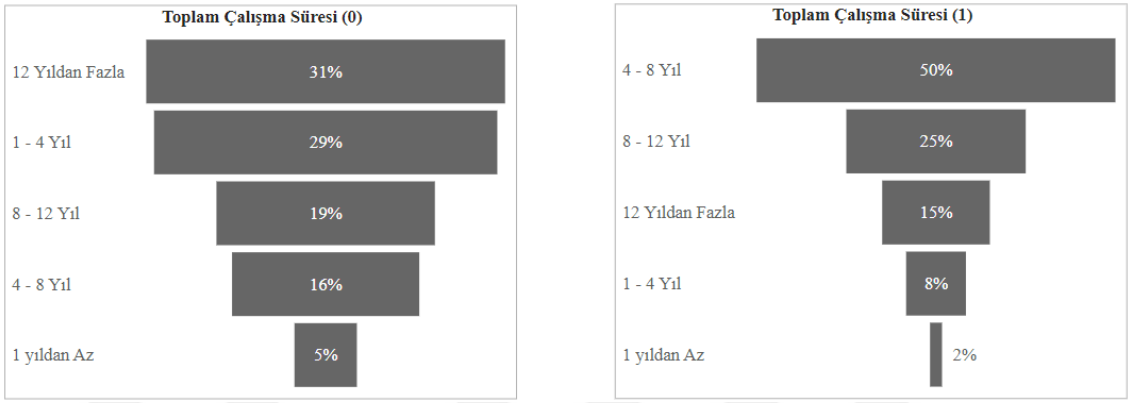
Eylemi gerçekleştirenlerin %45'inin banka öncesinde iş tecrübesinin bulunduğu, geriye kalanın ise ilk iş deneyiminin olduğu sonucuna ulaşılmıştır.



Şekil 3.9: Banka Dışı Deneyim Bilgisi

3.1.10 Çalışanların Toplam Çalışma Süresi

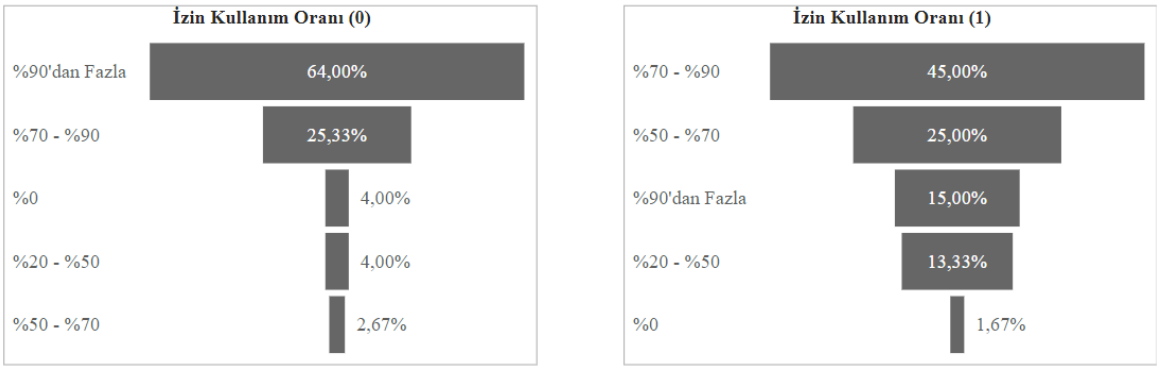
Çalışanların banka içi ve banka dışı toplam çalışma süreleri incelendiği vakit, suistimal yapanların %90'ının toplam tecrübesinin 4 yıldan fazla olduğu görülmektedir.



Şekil 3.10: Toplam Çalışma Süresi

3.1.11 Çalışanların İzin Kullanım Oranı

Çalışanların, hak etmiş oldukları yıllık ücretli izinlerinin ne kadarını kullandıklarına yönelik karşılaştırma yapıldığında, suistimalcilerin ortalama izin kullanım oranı %72 iken, suistimal gerçekleştirmeyenlerin ise %85'tir. Kullanım oranlarının dağılımına bakıldığında ise, suistimal gerçekleştirmeyenlerin %64'ü izinlerinin %90'nından fazlasını kullanırken, suistimalcilerde ise bu oranın %15 olduğu sonucuna ulaşılmaktadır.

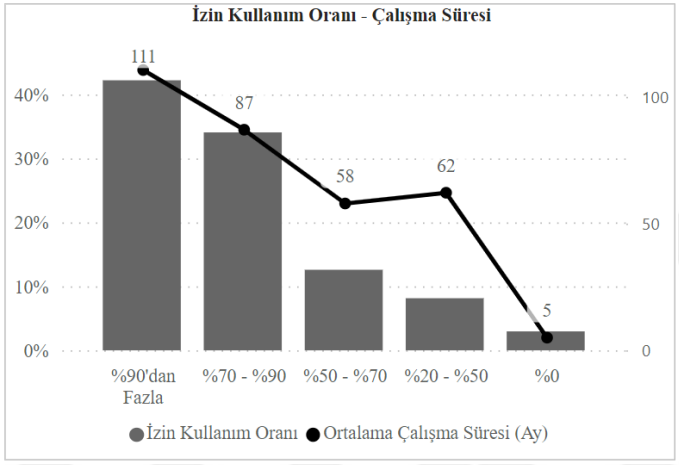


Şekil 3.11: İzin Kullanım Oranı

Suistimal yapan personel, işlemiş olduğu eylemin açığa çıkmaması için genellikle yıllık ücretli izin kullanmamaya gayret eder ve kullanmak zorunda olduğu izinleri de çeşitli bahanelerle ötelemeye çalışır. Çalışmamıza konu suistimalciler, hak etmiş oldukları yıllık

izinlerinin her ne kadar ortalama %72'sini kullanmış olsalar da, hangi dönemlerde izin kullanıp kullanmadıkları, yanı sıra hangi tarihte suistimal eylemini gerçekleştirdiği bilgisi temin edilemediğinden karşılaştırma yapılamamıştır.

Hiç izin kullanmayan çalışanların, bankadaki çalışma sürelerinin bir yılın altında olması ve buna bağlı olarak tahakkuk eden izinlerinin olmaması sebebiyle kullanamadıkları sonucuna ulaşılmıştır.

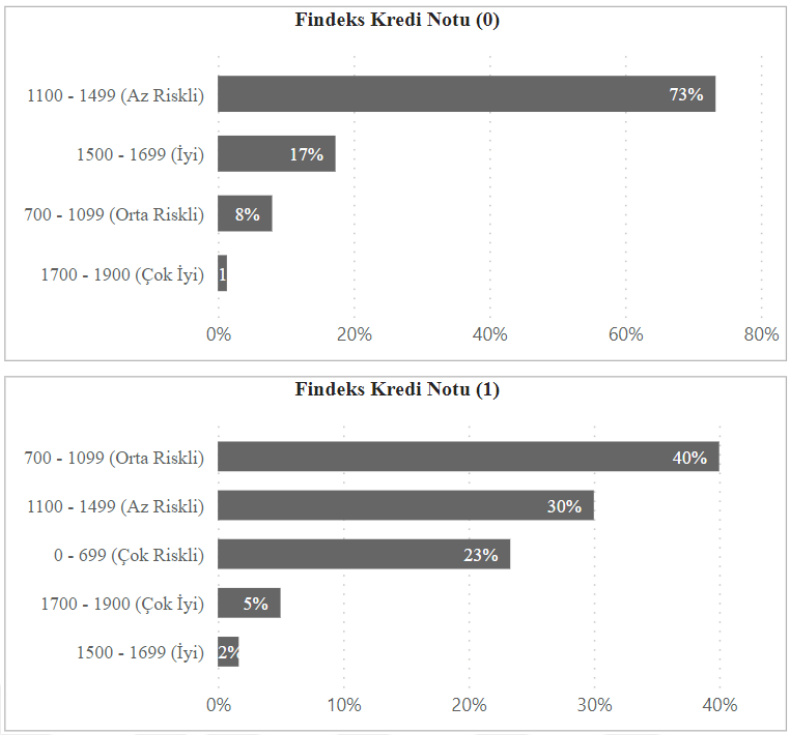


Şekil 3.12: İzin Kullanım Oranı ve Çalışma Süresi

3.1.12 Çalışanların Findeks Kredi Notu

KKB Kredi Kayıt Bürosu AŞ tarafından hesaplanan Findeks kredi notu, *"kişilerin kredi, kredi kartı ve kredili mevduat hesabı borçlarını ödeme alışkanlıklarına, borç durumlarına, yeni kredili ürün ve kredi kullanım yoğunluklarına göre hesaplanan nottur"* şeklinde tanımlanmıştır ve en riskli not 1, en az riskli not 1900 olup, en riskli, orta riskli, az riskli, iyi ve çok iyi şeklinde olmak üzere beş gruba ayrılmıştır.[6] Bankaların risk iştahına bağlık olmakla birlikte, bankalar genel olarak Findeks kredi notu 1000'in altı olan kişilere kredi ve/veya kredili ürünler tahsis etmemektedir. Kurumdan kuruma değişmekle birlikte genel olarak söz konusu bu beş grubun notları 0-699 en riskli, 700-1099 orta riskli, 1100-1499 az riskli, 1500-1699 iyi ve 1700-1900 çok iyi olacak şekilde belirlenmektedir (Yapı Kredi, 2022).

[6] Findeks, https://www.findeks.com/urunler/findeks-kredi-notu, (Erişim Tarihi: 10.10.2022)

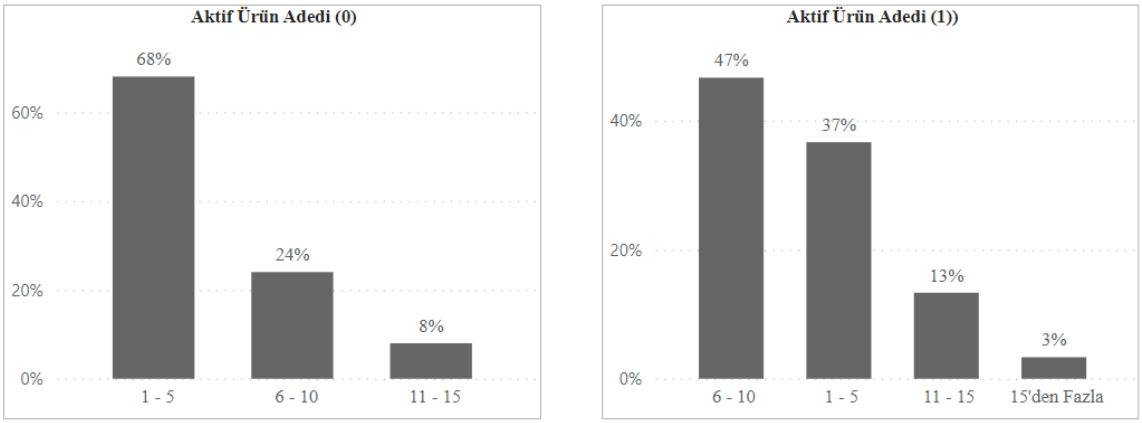


Şekil 3.13: Findeks Kredi Notu

Araştırmamıza konu suistimal yapmayan çalışanların çoğunun Findeks kredi notu az riskli olarak gruplandırılan 1100 ile 1499 arasında bir değer almıştır. Suistimal yapanlar açısından değerlendirildiğinde, %23'ü çok riskli grupta yer alırken, %40'ı ise orta riskli grupta yer almaktadır. Suistimal gerçekleştirmeyenlerin ise çok riskli grupta değerlendirilen notunun olmadığı, yalnızca %8'inin orta riskli gruba dahil olduğu görülmektedir. Ayrıca, suistimacilerin ortalama kredi notu 939 iken, suistimal yapmayanların ise 1375'dir. Böylece, suistimal yapanların, finansal ürünlere ilişkin borç geri ödemelerini düzenli yapmayarak gecikmeye sebebiyet verdikleri veya gelirlerine göre borçluluk durumlarının yüksek olduğu sonucuna ulaşılmaktadır.

3.1.13 Çalışanların Aktif Finansal Ürünleri

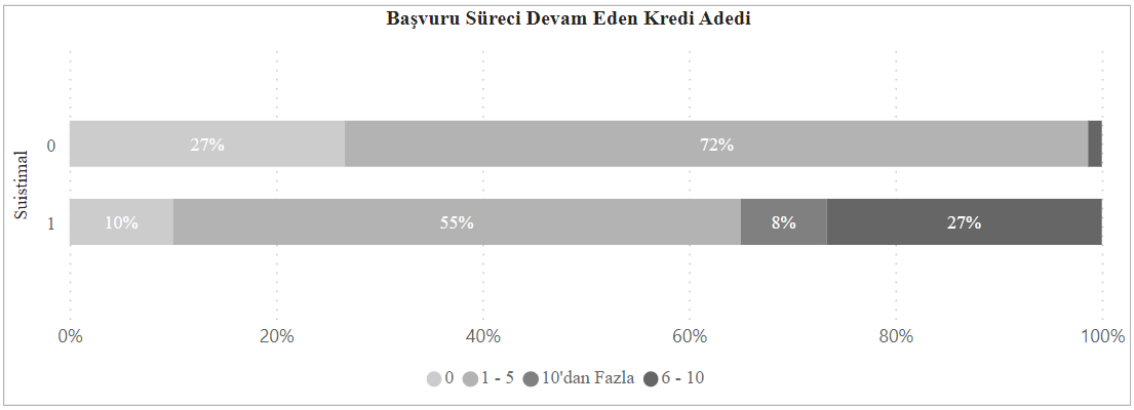
Suistimal yapanların soruşturma tarihinde, suistimal yapmayanların ise verilerin toplanması anında yapılan sorgulamalarda, araştırmaya konu (135) çalışanın tüm finansal kurumlar nezdinde ortalama 6 adet kredi, kredi kartı ve/veya kredili mevduat hesabı gibi aktif finansal ürünün bulunduğu görülmektedir. Ortalama 7 adet aktif ürünü bulunan suistimalcilerin %63'ü, 5 adetten fazla aktif kredi ürününe sahip iken, suistimale karışmayanların %68'i ise 5 adetten daha az aktif kredi ürününe sahiptir.



Şekil 3.14: Aktif Finansal Ürün Adedi

3.1.14 Çalışanların Başvuru Süreci Devam Eden Kredi Bilgisi

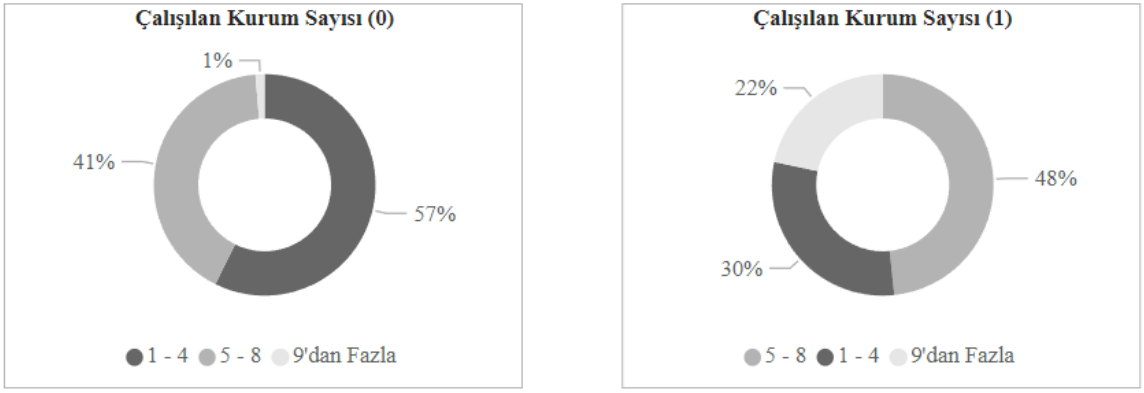
Soruşturma tarihinde yapılan sorgulamalara göre, suistimal gerçekleştiren kişilerin %35'inin, henüz onaylanmamış ancak başvuru süreci devam eden 5 adetten fazla kredisinin bulunduğu, suistimal gerçekleştirmeyenlerde ise bu oranın %1 olduğu görülmektedir. Bu durumun, eylemi gerçekleştirenlerin bir takım baskı veya motivasyon ile suistimal haricinde kaynak arayışında olduklarını işaret ettiği düşünülmektedir. Söz konusu baskı ve motivasyon unsuru, yakalanma durumunda işsiz kalacağı ve sonrasında kaynak bulmakta zorlanacağı düşüncesi olacağı gibi yaptığı suistimalin ihtiyaçlarını karşılamada yetersiz olduğu düşüncesi de olabilir.



Şekil 3.15: Başvuru Süreci Devam Eden Kredi Adedi

3.1.15 Çalışanların Finansal Ürünlerinin Bulunduğu Kurumlar

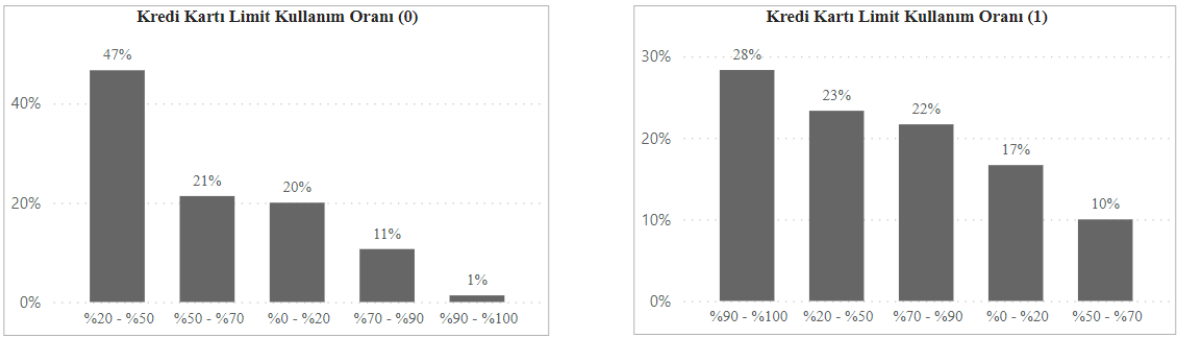
Sahip oldukları finansal ürünlere ilişkin suistimal yapanların, yapmayanlara göre daha fazla finansal kurumla çalışmasının olduğu sonucuna ulaşılmıştır. Ortalama 6 kurumla çalışması olan suistimalcilerin %70'inin, 4'ten fazla finansal kurumla kredi ilişkisi bulunmaktadır.



Şekil 3.16: Çalışılan Kurum Sayısı

3.1.16 Çalışanların Kredi Kartı Limit Kullanım Oranı

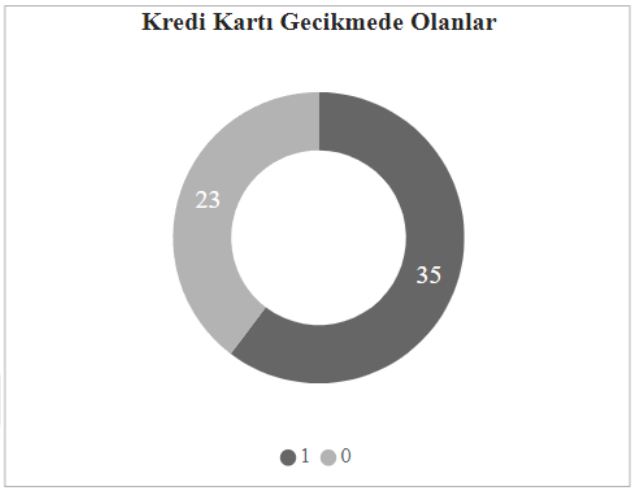
Suistimalcilerin %50'si, tüm bankalar nezdinde bulunan toplam kredi kartı limitlerinin %70'den fazlasını kullanırken, bu oran suistimal yapmayanlarda %12 seviyesindedir. Dolayısıyla genel olarak suistimal yapanların, yapmayanlara kıyasla kredi kartı limit doluluk oranlarının daha yüksek olduğu söylenebilir.



Şekil 3.17: Kredi Kartı Limit Kullanım Oranı

3.1.17 Çalışanların Kredi Kartı Gecikmeleri

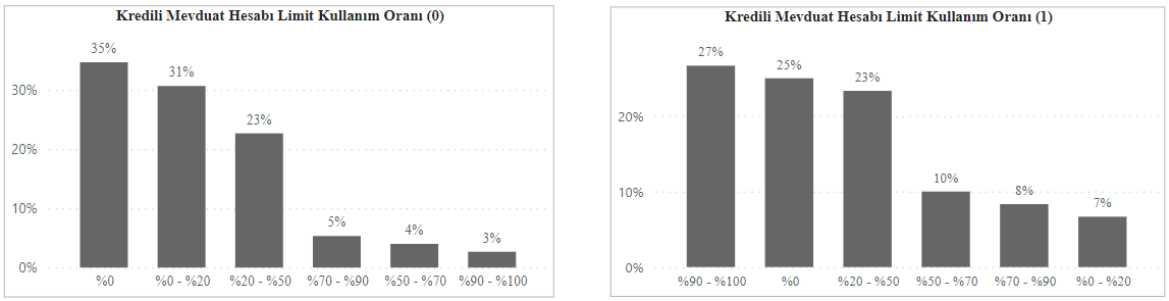
Suistimal yapanların 35'inde (%58), suistimal yapmayanların ise 23'ünde (%31) vadesi geçmiş ancak ödemesi yapılmamış kredi kartı borcu bulunmaktadır.



Şekil 3.18: Kredi Kartı Gecikmede Olanlar

3.1.18 Çalışanların Kredili Mevduat Hesabı Limit Kullanım Oranı

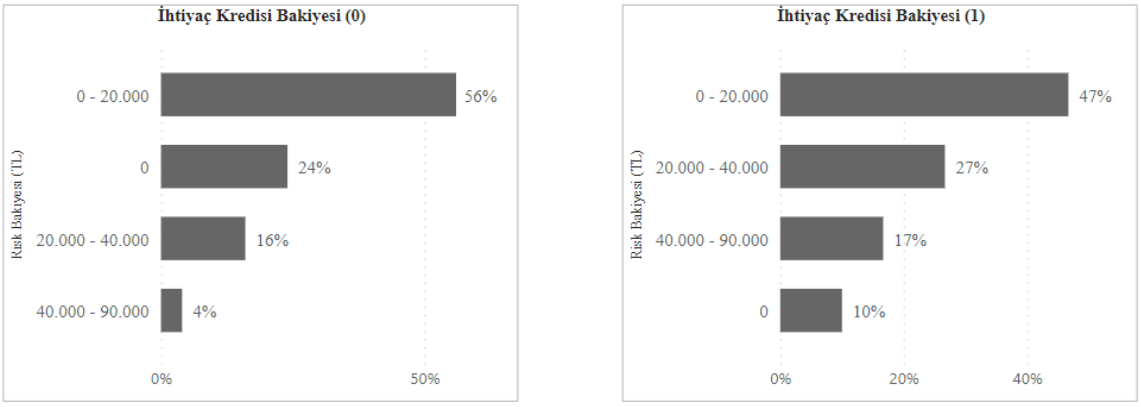
Kredi kartı limitlerinin kullanımında olduğu gibi, benzer şekilde suistimal gerçekleştirenlerin %35'i kredili mevduat hesabı limitlerinin %70'ten fazlasını kullanmaktadır. Yanı sıra, suistimalcilerin %25'inin, suistimale karışmayanların ise %35'inin limitlerinin tamamının sorgulama tarihi itibarıyla kullanılabilir olduğu görülmektedir.



Şekil 3.19: Kredili Mevduat Hesabı Limit Kullanım Oranı

3.1.19 Çalışanların İhtiyaç Kredisi

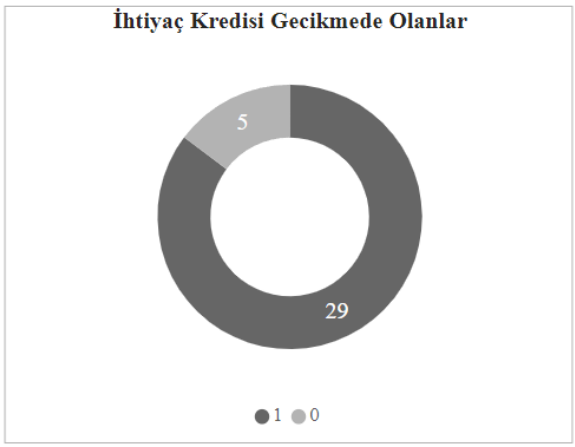
Suistimalcilerin %90'nının, suistimal gerçekleştirmeyenlerin ise %76'sının ihtiyaç kredisi bulunmakta olup, ortalama ihtiyaç kredisi bakiyesi, suistimal yapanlarda 22.300 TL iken, suistimal yapmayanlarda ise 11.024 TL'dir. Ayrıca, suistimalcilerin %44'ünün, suistimale karışmayanların ise %20'sinin ihtiyaç kredisi bakiyesi 20.000 TL'nin üzerindedir.



Şekil 3.20: İhtiyaç Kredisi

3.1.20 Çalışanların İhtiyaç Kredisi Gecikmeleri

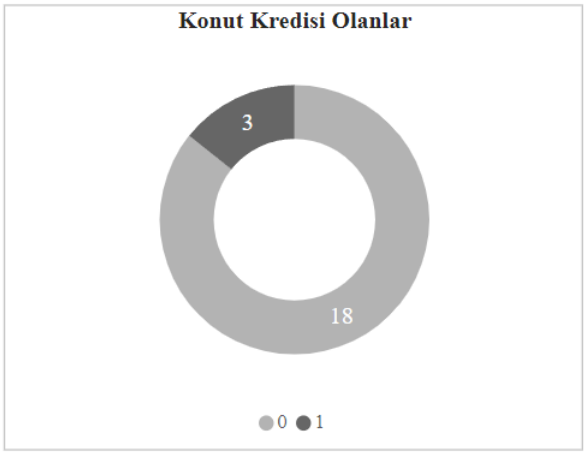
İhtiyaç kredisi riski bulunan suistimalcilerin 29'unun (%48) ortalama 1.055 TL'lik, suistimale karışmayanların ise 5'inin (%7) ortalama 261 TL'lik vadesi geçmiş ancak ödemesi yapılmamış kredi borcu bulunmaktadır.



Şekil 3.21: İhtiyaç Kredisi Gecikmesi

3.1.21 Çalışanların Konut Kredisi

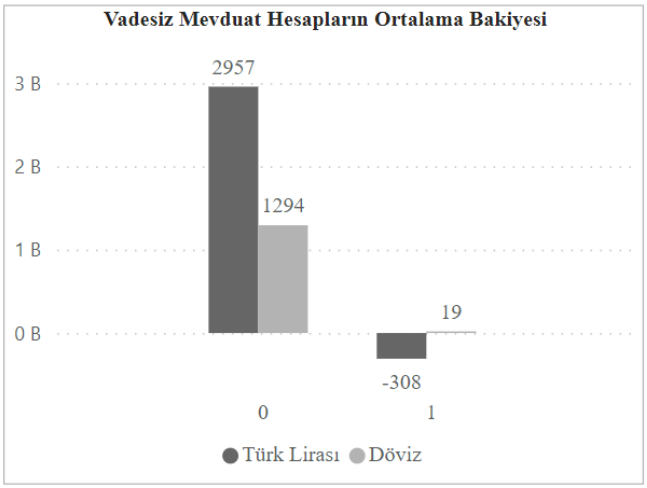
Sorgulama tarihi itibarıyla suistimal yapanların 3'ünün (%5), suistimal yapmayanların ise 18'inin (%24) konut kredisi riski bulunmaktadır.



Şekil 3.22: Konut Kredisi

3.1.22 Çalışanların Vadesiz Mevduat Hesapları

Suistimal gerçekleştirenlerin, soruşturma tarihinden önceki bir aylık dönemde çalıştıkları bankadaki Türk Lirası vadesiz mevduat hesapları ile kredili mevduat hesaplarının ortalama bakiyesi -308 TL, döviz vadesiz mevduat hesaplarının ortalama bakiyesi ise 19 USD/EUR'dur. Buna karşın suistimal gerçekleştirmeyenlerin sorgulama tarihinden önceki bir aylık dönemde çalıştıkları bankadaki vadesiz mevduat hesapları ile kredili mevduat hesaplarının ortalama bakiyesinin 2.957 TL, döviz vadesiz mevduat hesaplarının ortalama bakiyesinin ise 1.294 USD/EUR olduğu görülmektedir. Ayrıca, tüm çalışanların Türk Lirası vadesiz mevduat hesabı bulunurken, suistimal yapanların %27'sinin (ortalama 72 USD/EUR), suistimal yapmayanların ise %60'ının (ortalama 2.156 USD/EUR) döviz vadesiz mevduat hesabı bulunmaktadır. Dolayısıyla, suistimal eylemini gerçekleştiren çalışanın, personeli olduğu bankadaki vadesiz mevduat hesaplarında fazla para tutmadığı, hatta TL hesabında açık hesap limitini de kullanarak hesabın eksi bakiyeye düşmesine sebebiyet verdiği ve bu durumda suistimalcinin nakde ihtiyacının olduğu sonucuna ulaşılabilir.



Şekil 3.23: Vadesiz Hesapların Ortalama Bakiyesi

3.2 Değişkenlerin Seçimi

Suistimal değişkeni bağımlı değişken, diğer 23 değişken ise bağımsız değişken olarak belirlenmiştir. Açıklayıcı gücünün yüksek olması ve model performansını iyileştirmesi yönünden bağımsız değişkenlerin seçimi önem arz etmektedir. Modele uygun değişkenlerin seçilmesi ve değişken sayısının azaltılması, makine öğrenmesinde eğitim ve test süreçlerinin zaman ve maliyet açısından etkinliğini artırabilir, sade, anlaşılabilir ve iyi bir tahminleyici model kurulmasını sağlayabilir, aşırı öğrenmeyi (overfitting) önleyebilir ve modelin sınıflandırma performansı artırılabilir (Ahmad vd., 2019; Turfan, 2020).

Bağımsız değişkenlerin seçilmesinde, adımsal Lojistik Regresyon yöntemi kullanılarak, SPSS Statistics uygulaması aracılığıyla öncelikle her bir bağımsız değişken, bağımlı değişken olan suistimal değişkeni ile modellenmiş ve değişkenlerin %5 önem düzeyinde anlamlılıkları (Variable Significance) ile modele uyumlulukları (Hosmer ve LemesHow Testi) değerlendirilmiştir.

Model sonuçları değerlendirildiğinde, değişken katsayısı ile modele uyumluluğu istatistiksel olarak anlamlı olan 11 değişken elde edilmiştir.

Tablo 3.2: Tek Değişkenli Lojistik Regresyon Analizi

Değişken	Hosmer ve LemesHow Testi (Significance)	Olabilirlik Oranındaki Değişim (Log LikeHood)	Nagelkerke R Kare	Değişkenin Anlamlılık Düzeyi (Significance)
Egitim_Duzeyi	1	26,002	0,235	0,001
KKB_Skor	0,34	53,522	0,438	0
Kredi_Basvuru_Adedi	0,337	45,989	0,387	0
Aktif_Urun_Adedi	0,285	18,302	0,17	0
Calisilan_Kurum_Sayisi	0,647	24,883	0,225	0
Ihtiyac_Kredisi_Bakiyesi	0,114	16,906	0,158	0
Ihtiyac_Kredisi_Gecikme_Tutarı	0,186	36,022	0,314	0,005
Kredi_Karti_Gecikme_Tutari	0,982	26,029	0,235	0
Acik_Hesap_Kullanim_Orani	0,075	20,299	0,187	0
TL_Vadesiz_Hesap_Bakiyesi	0,157	19,162	0,177	0,006
YP_Vadesiz_Hesap_Bakiyesi	0,39	50,803	0,42	0,001

Söz konusu değişkenler bu sefer birlikte Lojistik Regresyon modeline dahil edilerek analiz edilmiş ve sonuçlarına Tablo 3.3'te yer verildiği üzere, söz konusu değişkenler arasından yalnızca "Eğitim Düzeyi, KKB Skoru, Kredi Başvuru Adedi ve Açık Hesap Kullanım Oranı" değişkenlerinin %5 önem düzeyinde istatistiki olarak anlamlı olduğu görülmüştür.

Tablo 3.3: 11 Değişkenli Lojistik Regresyon Modeli

	Eğim/Katsayı (B)	**Standart Hata (S.E.)**	**Serbestlik Derecesi (df)**	**Anlamlılık Düzeyi (Significance)**
Egitim_Duzeyi			**3**	**,012**
Egitim_Duzeyi(1)	24,023	13046,872	1	,999
Egitim_Duzeyi(2)	21,036	13046,872	1	,999
Egitim_Duzeyi(3)	19,115	13046,872	1	,999
KKB_Skor	**-,004**	**,002**	**1**	**,034**
Kredi_Basvuru_Adedi	**,726**	**,256**	**1**	**,005**
Aktif_Urun_Adedi	-,530	,278	1	,056
Calisilan_Kurum_Sayisi	,246	,371	1	,507
Ihtiyac_Kredisi_Bakiyesi	,000	,000	1	,502
Ihtiyac_Kredisi_Gecikme_Tutarı	,003	,003	1	,287
Kredi_Karti_Gecikme_Tutari	,001	,001	1	,514
Acik_Hesap_Kullanim_Orani	**-3,226**	**1,541**	**1**	**,036**
TL_Vadesiz_Hesap_Bakiyesi	,000	,000	1	,965
YP_Vadesiz_Hesap_Bakiyesi	-,005	,002	1	,052
Constant	1,329	3261,719	1	1,000

İstatistiki olarak anlamlı olmayan değişkenler modelden çıkarılmış, yerine her ne kadar tek değişkenli Lojistik Regresyon modelinde istatistiki olarak anlamlı olmasa da suistimal olup olmama konusunda önemli bir değişken olduğu düşünülen "Yaş, Cinsiyet, Birimde Çalışma Süresi, İzin Kullanım Oranı ve Kredi Kartı Doluluk Oranı" değişkenleri modele eklenerek çalışmalara devam edilmiştir. Söz konusu modelin çıktılarında ise bu kez "Açık Hesap Kullanım Oranı ile Kredi Kartı Doluluk Oranı" değişkenleri istatistiki olarak anlamsız çıkmış olup, bu değişkenlerin her biri ayrı ayrı modelden çıkarılıp tekrar modele dahil edilerek tüm değişkenlerin istatistiki olarak anlamlı, açıklayıcı gücü ve olabilirlik oranındaki değişimin yüksek olduğu ve uyum eksikliğinin olmadığı bir model oluşturulmuştur.

Tablo 3.4: 8 Değişkenli Lojistik Regresyon Modeli

	Eğim / Katsayı (B)	Standart Hata (S.E.)	Serbestlik Derecesi (df)	Anlamlılık Düzeyi (Significance)	Odds Oranı Exp(B)
Yas	-,238	,089	1	,007	,788
Cinsiyet(1)	2,920	1,039	1	,005	18,544
Egitim_Duzeyi			3	,010	
Egitim_Duzeyi(1)	24,116	13705,922	1	,999	29745297261,282
Egitim_Duzeyi(2)	21,580	13705,922	1	,999	2356369073,132
Egitim_Duzeyi(3)	19,252	13705,921	1	,999	229735480,599
Birimde_Calisma_Suresi	,060	,017	1	,000	1,062
Izin_Kullanim_Orani	-3,735	1,598	1	,019	,024
KKB_Skor	-,008	,002	1	,001	,992
Kredi_Basvuru_Adedi	,678	,221	1	,002	1,971
Kredi_Karti_Doluluk_Orani	-4,372	2,194	1	,046	,013
Constant	13,618	3426,484	1	,997	821103,895

Olabilirlik oranındaki değişimin 134,51 olduğu ve dolayısıyla bağımsız değişkenlerin, modele dahil edilmesi durumunda tahmin üzerinde 134,51 birimlik katkı sağladığı söylenebilir. Yanı sıra modelin Nagelkerke R^2 değerinin %84,5 olduğu, böylece suistimal değişkeni olan bağımlı değişkenin %84,5'inin söz konusu bağımsız değişkenler tarafından açıklandığı görülmektedir.

Modelin veriye uyumunun iyi olup olmadığını test etmek için uyum iyiliğini gösteren Hosmer ve Lemeshow test sonucunun (Sig.(P) değeri: 1) istatistiksel olarak anlamlı olduğu, dolayısıyla modelde uyum eksikliğinin olmadığı ve modelin verilere uygun olduğu söylenebilir.

3.3 Makine Öğrenmesi Uygulaması

3.3.1 Veri Ön İşleme Süreci

Modelleme aşamasına geçilmeden önce veriler, KNIME veri analitiği programı aracılığıyla aşağıda ayrıntılarına yer verilen düğümler ile ön işleme sürecine tabi tutulmuş ve sonrasında öngörü yapmak üzere 5 farklı sınıflandırma algoritması ile model geliştirilmiştir.

Şekil 3.24: Veri Ön İşleme Süreci

- **Excel Reader:** Microsoft Excel ortamında hazırlanan veri setinin okunması

- **Number To String:** Kategorik değişken olan ancak rakamlar ile ifade edilen Cinsiyet, Medeni Durum, Eğitim Düzeyi, Pozisyon, Banka Dışı Deneyim ve Suistimal değişkenlerinin veri türünün kategorik değişken olarak tanımlanması

- **SMOTE:** En yakın 5 komşusunu baz alarak Sentetik Azınlık Örnekleme Tekniği (Synthetic Minority Oversampling Technique, SMOTE) ile yapay veriler türetilmesi

- **Double To Int:** SMOTE ile türetilen ve ondalık sayı tipinde olan Yaş, KKB Skoru, Kredi Başvuru Adedi, Aktif Ürün Adedi ve Çalışılan Kurum Sayısı gibi değişkenlere ilişkin verilerin tam sayıya yuvarlanması

- **Shuffle:** Verilerin rastgele olarak karıştırılması

- **Column Filter:** Adımsal Lojistik Regresyon yöntemi ile seçilen 8 bağımsız değişken ile suistimal değişkeninin modele eklenmesi

- **Normalizer:** Sayı biçiminde olan ham verilerin Z skor normalleştirme ile standartlaştırılması

- **X-Partitioner:** Her verinin eğitim ve test verisi olarak kullanılması için k-katlı çapraz doğrulama (k-fold cross validation) işlemine tabi tutulması (K=10 ve tabakalı örnekleme (Stratified Sampling) tercih edilmiştir.)

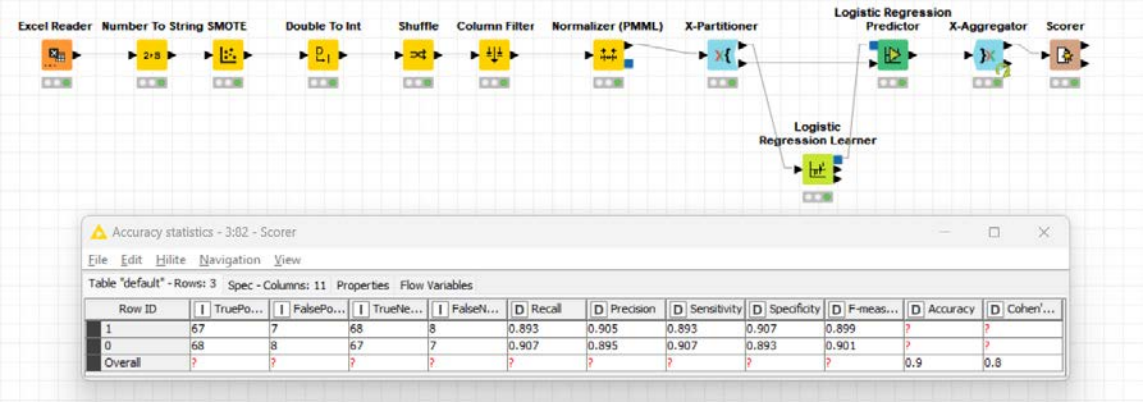
3.3.2 Makine Öğrenmesi Model Geliştirme Süreci

Değişkenlerin belirlenmesinin ve veri ön işleme sürecinin akabinde, KNIME veri analitiği programı aracılığıyla çalışanların suistimal gerçekleştirip gerçekleştirmediğine yönelik öngörü yapmak amacıyla makine öğrenmesi sınıflandırma algoritmalarından Lojistik Regresyon (Logistic Regression), Rastgele Orman (Random Forest), Karar Ağacı (Decision Tree), Naive Bayes ve Destek Vektör Makinesi (Support Vector Machine) kullanılarak modeller geliştirilmiştir.

Her bir modellemeye ilişkin düğümlerin sonuna, tüm satırlar için öngörülen sınıf ile gerçek sınıfı karşılaştırmak amacıyla X-Aggregator düğümü, hata matrisi (confusion matrix) ve doğruluk (accuracy), duyarlılık (recall), kesinlik (precision), seçicilik (specifity) ve F ölçümü gibi istatistiksel bilgileri elde edebilmek için ise Score düğümü eklenmiştir.

3.3.2.1 Lojistik Regresyon

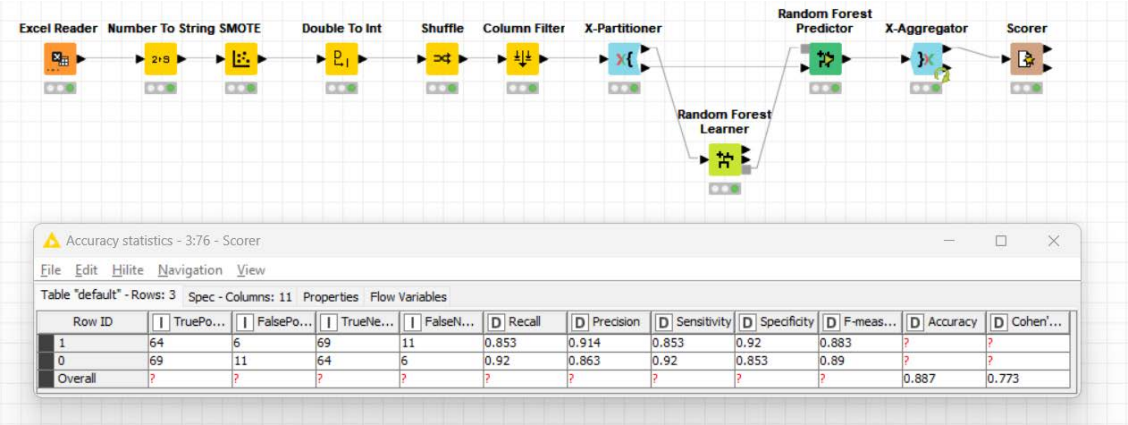
Lojistik Regresyon modeli, Logistic Regression Learner ve Logistic Regression Predictor düğümleri ile oluşturulmuştur. Söz konusu modelin doğruluk oranı %90, F-skor değeri ise %89,90'dır.



Şekil 3.25: Lojistik Regresyon

3.3.2.2 Rastgele Orman

Rastgele Orman modeli, Random Forest Learner ve Random Forest Predictor düğümleri ile oluşturulmuştur. Söz konusu modelin doğruluk oranı %88,70, F-skor değeri ise %88,30'dur.

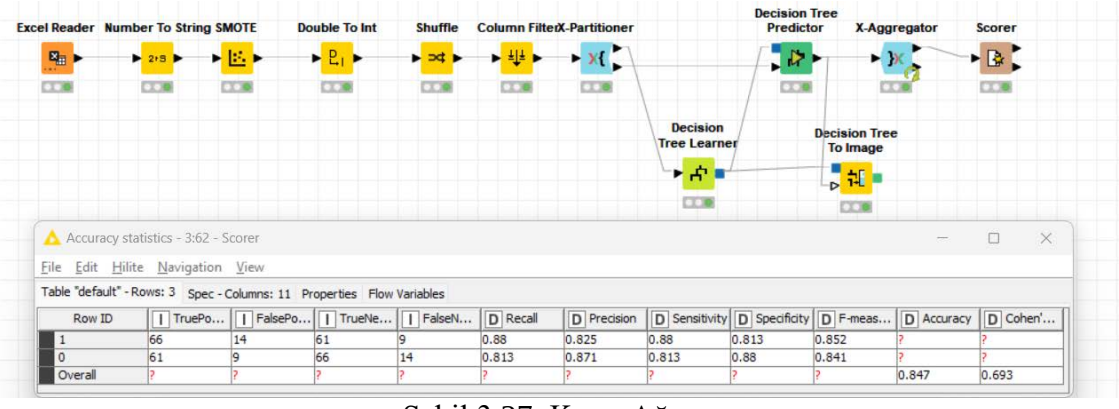


Row ID	TruePo...	FalsePo...	TrueNe...	FalseN...	Recall	Precision	Sensitivity	Specificity	F-meas...	Accuracy	Cohen'...
1	64	6	69	11	0.853	0.914	0.853	0.92	0.883	?	?
0	69	11	64	6	0.92	0.863	0.92	0.853	0.89	?	?
Overall	?	?	?	?	?	?	?	?	?	0.887	0.773

Şekil 3.26: Rastgele Orman

3.3.2.3 Karar Ağacı

Karar Ağacı modeli, Decision Tree Learner ve Decision Tree Predictor düğümleri ile oluşturulmuştur. Söz konusu modelin doğruluk oranı %84,70, F-skor değeri ise %85,20'dir.

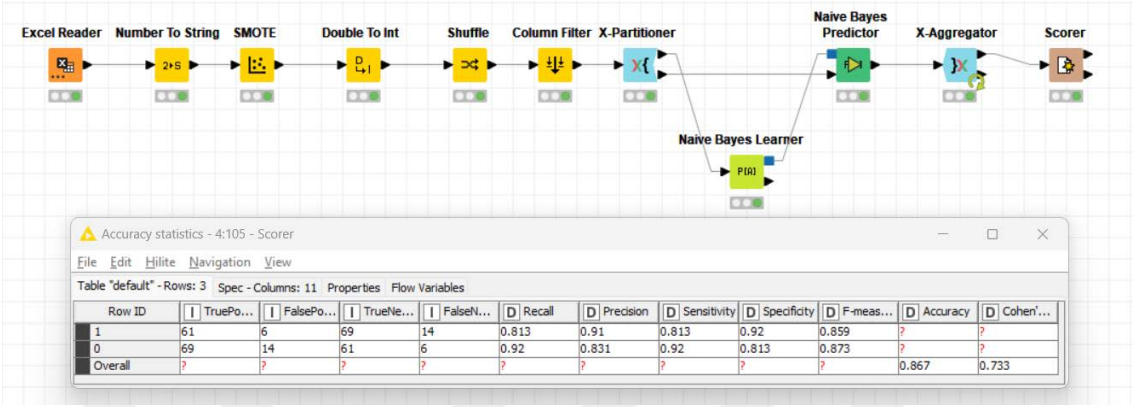


Row ID	TruePo...	FalsePo...	TrueNe...	FalseN...	Recall	Precision	Sensitivity	Specificity	F-meas...	Accuracy	Cohen'...
1	66	14	61	9	0.88	0.825	0.88	0.813	0.852	?	?
0	61	9	66	14	0.813	0.871	0.813	0.88	0.841	?	?
Overall	?	?	?	?	?	?	?	?	?	0.847	0.693

Şekil 3.27: Karar Ağacı

3.3.2.4 Naive Bayes

Naive Bayes modeli, Naive Bayes Learner ve Naive Bayes Predictor düğümleri ile oluşturulmuştur. Söz konusu modelin doğruluk oranı %86,70, F-skor değeri ise %85,90'dır.

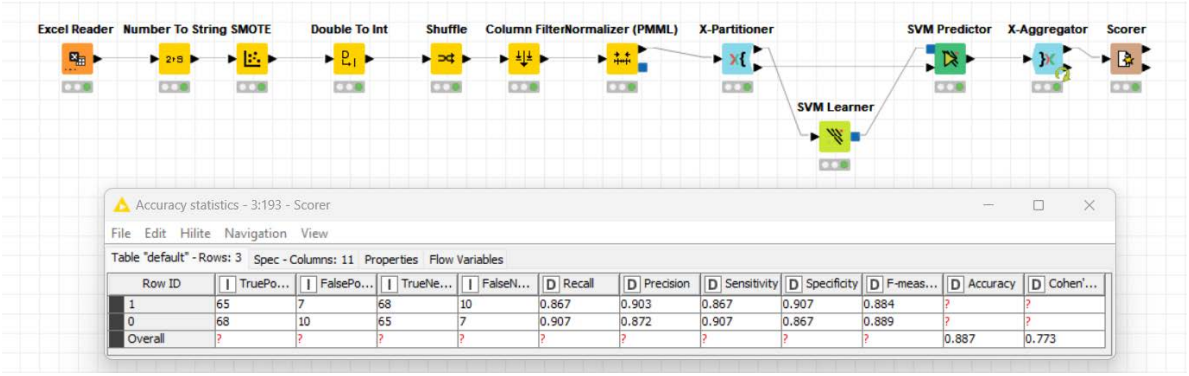


Row ID	TruePo...	FalsePo...	TrueNe...	FalseN...	Recall	Precision	Sensitivity	Specificity	F-meas...	Accuracy	Cohen'...
1	61	6	69	14	0.813	0.91	0.813	0.92	0.859	?	?
0	69	14	61	6	0.92	0.831	0.92	0.813	0.873	?	?
Overall	?	?	?	?	?	?	?	?	?	0.867	0.733

Şekil 3.28: Naive Bayes

3.3.2.5 Destek Vektör Makinesi

Destek Vektör Makinesi modeli, SVM Learner ve SVM Predictor düğümleri ile oluşturulmuştur. Söz konusu modelin doğruluk oranı %88,70, F-skor değeri ise %88,40'tır.



Row ID	TruePo...	FalsePo...	TrueNe...	FalseN...	Recall	Precision	Sensitivity	Specificity	F-meas...	Accuracy	Cohen'...
1	65	7	68	10	0.867	0.903	0.867	0.907	0.884	?	?
0	68	10	65	7	0.907	0.872	0.907	0.867	0.889	?	?
Overall	?	?	?	?	?	?	?	?	?	0.887	0.773

Şekil 3.29: Destek Vektör Makinesi

3.3.3 Model Performans Sonuçlarının Değerlendirilmesi

Suistimal olup olmadığına yönelik 8 bağımsız değişken ile öngörüde bulunan 5 farklı sınıflandırma algoritması üzerinden geliştirilen modellerin performans değerlerine, karşılaştırmalı olarak Tablo 3.5'te yer verilmiştir.

Doğruluk ile duyarlılık ve kesinlik ölçülerinin harmonik ortalaması olan F-skor performans ölçütü 0 ile 1 arasında bir değer almakta olup, 1'e yaklaştıkça, modelin sınıflandırma başarısı da buna parelel artmaktadır (Chicco & Jurman, 2020, 4-5).

Performans ölçütlerinden doğruluk oranı ile F-Skor değerine göre, personelin suistimal yapıp yapmadığının belirlenmesinde %90 doğruluk oranı ve %89,90 F-Skor değeri ile en iyi sınıflandırma algoritmasının Lojistik Regresyon olduğu, Karar Ağacı algoritmasının ise %84,70 doğruluk oranı ve %85,20 F-Skor değeri ile diğerlerine kıyasla daha düşük performans sergilediği görülmüş olsa da, çalışmada kullanılan 5 modelin de genel olarak sınıflandırma performanslarının yüksek olduğu, böylece suistimallerin belirlenmesinde makine öğrenmesi algoritmalarının etkili bir yöntem olduğu sonucuna ulaşılmıştır.

Tablo 3.5: Modellerin Performans Sonuçları

Model	Doğruluk	Duyarlılık	Kesinlik	F-Skor
Lojistik Regresyon	90,00%	89,30%	90,50%	89,90%
Destek Vektör Makinesi	88,70%	86,70%	90,30%	88,40%
Rastgele Orman	88,70%	85,30%	91,40%	88,30%
Naive Bayes	86,70%	81,30%	91,00%	85,90%
Karar Ağacı	84,70%	88,00%	82,50%	85,20%

BÖLÜM 4: SONUÇ VE ÖNERİLER

Yapılan ulusal ve uluslararası çalışmalar, bankacılık ve finans sektöründe gerçekleştirilen suistimal oranının ve kayıp tutarının diğer sektörlere kıyasla daha yüksek olduğunu göstermektedir (ACFE, 2018; ACFE, 2020; ACFE, 2022; Yılmaz, 2019) ve bu suistimallerin yıkıcı etkisi, sektörde yer alan kurumların zaman zaman itibar kaybı ile karşı karşıya gelmesine neden olduğu gibi, bazen de vermiş olduğu maddi zarar neticesinde faaliyetlerini sonlandırmasına neden olabilmektedir (Benston, 2004). Dolayısıyla, finansal kurumlar da suistimal konusuna her geçen gün daha fazla önem vermekte ve bu durumun üstesinden gelebilmek amacıyla, suistimallerin önlenmesi ve büyük boyutlara ulaşamadan ortaya çıkarılabilmesine yönelik gerek insan kaynağına gerekse teknolojiye yatırım yapmaktan çekinmemektedirler.

Makine öğrenmesi olarak bilinen bir yöntemin yeni bir alana uygulanarak suistimallerin tespitinde kullanılması, bunun birlikte, suistimalci profiline yönelik literatürde yapılan çalışmalarda kullanılan yaş, cinsiyet, eğitim düzeyi, pozisyon ve çalışma süresi gibi değişkenlere ilave farklı demografik, mesleki ve finansal değişkenlerin de eklenmesi ile toplam 23 değişken üzerinden suistimalcilerin özelliklerinin ortaya konması çalışmamızın önemini göstermekte olup, suistimalci profiline ilişkin detaylara Tablo 4.1'de yer verilmiştir.

Tablo 4.1: Suistimalci Profili

	Değişken	Kategori	Açıklama
1	Yaş	Demografik	21 ile 54 yaşları arasında
2	Cinsiyet	Demografik	%80'i erkek
3	Medeni Durumu	Demografik	%63'ü evli
4	Eğitim Düzeyi	Demografik	%45'i lise, %18'i önlisans mezunu
5	Pozisyon	Mesleki	%93'ü yönetici pozisyonu altında çalışmakta
6	Bankada Çalışma Süresi	Mesleki	%82'si 4 yıldan fazla bir süredir bankada çalışmakta
7	Birimde Çalışma Süresi	Mesleki	%42'si 4 yıldan fazla bir süredir suistimal yaptığı birimde çalışmakta
8	Görevde Çalışma Süresi	Mesleki	%44'ü 4 yıldan fazla bir süredir suistimal yaptığı esnadaki görevinde çalışmakta
9	Banka Dışı Deneyim	Mesleki	%45'inin banka öncesi deneyimi bulunmakta

10	Toplam Çalışma Süresi	Mesleki	%90'ının toplam iş tecrübesi 4 yıldan fazla
11	İzin Kullanım Oranı	Mesleki	Ortalama izin kullanım oranları %72 (Yalnızca %15'i izinlerinin %90'ınından fazlasını kullanmakta)
12	KKB Skor	Finansal	%23'ünün skoru 0-699 (çok riskli) arasında, %40'ının ise 700-1099 (orta riskli) arasında (Ortalama Findeks kredi notu 939)
13	Aktif Ürün Adedi	Finansal	%63'ünün 5 adetten fazla kredi, kredi kartı ve/veya kredili mevduat hesabı gibi finansal ürünü bulunmakta
14	Kredi Başvuru Adedi	Finansal	%35'inin soruşturma anında 5 adetten fazla onaylanmamış kredi başvurusu bulunmakta
15	Çalışılan Kurum Sayısı	Finansal	%70'inin 4'ten fazla finansal kurumla kredi ilişkisi bulunmakta
16	Kredi Kartı Kullanım Oranı	Finansal	%50'si toplam limitlerinin %70'ten fazlasını kullanmakta
17	Kredi Kartı Gecikmesi	Finansal	%58'inin kredi kartı gecikmesi bulunmakta
18	KMH Kullanım Oranı	Finansal	%35'i KMH limitlerinin %70'ten fazlasını kullanmakta
19	İhtiyaç Kredisi	Finansal	%90'nının ihtiyaç kredisi bulunmakta
20	İhtiyaç Kredisi Gecikmesi	Finansal	%48'inin ihtiyaç kredisi gecikmesi bulunmakta
21	Konut Kredisi	Finansal	%5'inin konut kredisi bulunmakta
22	TL Vadesiz Hesap Bakiyesi	Finansal	Ortalama bakiyesi -308 TL
23	YP Vadesiz Hesap Bakiyesi	Finansal	Ortalama bakiyesi ise 19 USD/EUR

Suistimallerin ortaya çıkarılmasına yönelik hâlihazırda uygulanmakta olan veri analitiği yöntemlerine ilave olarak, bu tez çalışmasında, çalışanların demografik, mesleki ve finansal bilgilerine yönelik belirlenen değişkenler üzerinden suistimal yapıp yapmadığı hakkında öngörüde bulunmak amacıyla 5 farklı makine öğrenmesi algoritması kullanılmıştır. Çalışmanın sonucunda, çalışanın suistimal yapıp yapmadığının öngörüsünde, makine öğrenmesi sınıflandırma algoritmalarının kullanımının, başta suistimal denetçileri olmak üzere bu alanda çalışan kişiler için önemli bir yöntem olduğu görülmüştür.

Çalışmamızda, 135 çalışana yönelik belirlenen 23 adet değişken, SPSS Statistics uygulaması aracılığıyla Lojistik Regresyon modeli üzerinden değerlendirilmiş ve çalışanın suistimal yapıp yapmadığı konusunda istatistiksel olarak anlamlı bulunan, açıklayıcı gücü yüksek, uyum eksikliği olmayan ve modele dahil edilmesi durumunda tahmin üzerinde olumlu katkı sağladığı görülen 8 adet değişken (Yaş, Cinsiyet, Eğitim Düzeyi, Birimde Çalışma Süresi, İzin Kullanım Oranı, KKB Skoru, Kredi Başvuru Adedi ve Kredi Kartı Doluluk Oranı) ile model geliştirilmiştir.

Makine öğrenmesi uygulaması ve veri ön işleme sürecinde KNIME veri analitiği programı kullanılmış olup, sınıflar arasındaki dengesiz veri dağılımının üstesinden gelmek amacıyla SMOTE tekniği ile sentetik örnekler oluşturularak diğer grubun veri sayısı ile denkleştirilmiş ve ön işleme süreci sonrası veriler, k-katlı çapraz doğrulama ve tabakalı örnekleme yöntemi ile eğitim ve test verisi olarak ayrılmıştır. Akabinde, çalışanların suistimal yapıp yapmadığına yönelik öngürü yapabilmek için makine öğrenmesi sınıflandırma algoritmalarından Lojistik Regresyon (Logistic Regression), Rastgele Orman (Random Forest), Karar Ağacı (Decision Tree), Naive Bayes ve Destek Vektör Makinesi (Support Vector Machine) yöntemleri kullanılmış ve algoritmaların sonuçları karşılaştırılmıştır. Söz konusu performans ölçütlerinden doğruluk oranı ile duyarlılık ve kesinlik ölçülerinin harmonik ortalaması olan F-Skor değerine göre, %90 doğruluk oranı ve %89,90 F-Skor değeri ile en iyi sınıflandırma algoritmasının Lojistik Regresyon olduğu, Karar Ağacı algoritmasının ise %84,70 doğruluk oranı ve %85,20 F-Skor değeri ile diğerlerine kıyasla daha düşük performans sergilediği görülmüş olsa da, çalışmada kullanılan 5 modelin de genel olarak sınıflandırma başarılarının yüksek olduğu, böylece suistimallerin belirlenmesinde makine öğrenmesi algoritmalarının etkili bir yöntem olduğu sonucuna ulaşılmıştır.

Makine öğrenmesi uygulamalarında, eğitim veri setindeki örneklem sayısının, bağımsız değişken sayısının ve veri kümelerinin dengeli olarak dağılmasının çok önemli olduğu düşünülmekle birlikte, değişken sayısı ile bunlara ilişkin veri sayısı artırıldıkça ve gruplar arası veri dağılımı dengeli oldukça algoritmaların sınıflandırma başarısının da buna paralel artacağı beklenmektedir. Ayrıca, bu tez çalışmasında kullanılan makine öğrenmesi algoritmaları haricinde Yapay Sinir Ağları gibi farklı algoritmalar üzerinden de modeller kurularak sınıflandırma performanslarının karşılaştırılmasının faydalı olabileceği düşünülmektedir.

Çalışmamızda model için, 23 bağımsız değişken içerisinden istatistiksel olarak anlamlı olan 8 değişken belirlenmiş olsa da, diğer 15 değişken farklı kurumlar için suistimal konusunda daha etkili öngörüler elde edilmesini sağlayabilir. Yanı sıra temin edilemediğinden her ne kadar çalışmamıza konu edilemese de, aşağıda belirtilen bağımsız değişkelerin suistimallerin ortaya çıkarılmasında etkili değişkenler olabileceği öngörülmektedir.

- Çalışanın performans notu
- Çalışanın kurum bilgisayarını mesai saatleri dışında kullanma durumu
- Çalışanın işe giriş ve işten çıkış saatleri
- Çalışanın finansal ürünlerine ilişkin geri ödeme gecikme gün sayısı
- Çalışanın yasal veya idari takibe konu kredisinin olup olmama durumu
- Çalışanın vadeli mevduat hesabının (TL/YP) sorgulama tarihi itibarıyla aylık ortalama bakiyesi

Yapay zekâ ve makine öğrenmesi çalışmalarının, özellikle bankacılık ve finans sektöründe olmak üzere suistimallerin ortaya çıkarılmasında her geçen gün daha fazla yer edineceği, dolayısıyla bu tez çalışmasının da gerek literatüre gerek bankacılık ve finans sektöründe faaliyet gösteren kurumların mevcut uygulamalarına katkıda bulunacağı düşünülmektedir.

KAYNAKÇA

5237 Sayılı Türk Ceza Kanunu. *T.C. Resmi Gazete* (25611, 12 Ekim 2004).

5411 Sayılı Bankacılık Kanunu. *T.C. Resmi Gazete* (25983, 1 Kasım 2005).

Association of Certified Fraud Examiners (ACFE). (2018). *Occupational Fraud 2018: A Report To The Nations*. https://s3-us-west-2.amazonaws.com/acfepublic/2018-report-to-the-nations.pdf, Erişim Tarihi: 18.10.2022

Association of Certified Fraud Examiners (ACFE). (2020). *Occupational Fraud 2020: A Report To The Nations.* https://acfepublic.s3-us-west-2.amazonaws.com/2020-Report-to-the-Nations.pdf, Erişim Tarihi: 18.10.2022

Association of Certified Fraud Examiners (ACFE). (2022). *Occupational Fraud 2022: A Report To The Nations.* https://acfepublic.s3.us-west-2. amazonaws.com/2022+Report+to+the+Nations.pdf, Erişim Tarihi: 18.10.2022

Ahmad, S. R., Bakar, A. A., & Yaakub, M. R. (2019). A review of feature selection techniques in sentiment analysis. *Intelligent Data Analysis*, 23(1), 159–189.

Aksoy, K. (2017). *Detection of Debit Card Fraud Through Random Forest* (Yüksek Lisans Tezi). Doğuş University, Instıtute of Science and Technology, İstanbul.

Albrecht, S. (2014). Iconic Fraud Triangle Endures. *Fraud Magazine*, July/August 2014. https://www.fraud-magazine.com/article.aspx?id=4294983342, Erişim Tarihi: 02.09.2022

Ataman, B., & Aydın, R. (2017). Hile Denetimi ve Denetçilerin Hile Tespitine Yönelik Bir Araştırma. *Marmara Business Review*, 2(1), 1-23.

Azam, M. R. (2018). Theory Application: Why People Commit Fraud. *International Journal of Management, Accounting and Economics*, 5(1), 54-65.

Bağımsız Denetim Standardı 240. Finansal Tabloların Bağımsız Denetiminde Bağımsız Denetçinin Hileye İlişkin Sorumlulukları. *T.C. Resmi Gazete* (30272, 16 Aralık 2017).

Barros, R. C., de Carvalho, A.C.P.L.F., & Freitas, A. A. (2015). *Decision-Tree Induction. In: Automatic Design of Decision-Tree Induction Algorithms*. SpringerBriefs in Computer Science.

Benston, G. J. (2004). What's Special about Banks?. *The Financial Review*, 39, 13-33.

Berthold, M. R., Cebron, N., Dill, F., Gabriel, T. R., Kötter, T., Meinl, T., Ohl, P., Sieb, C., Thiel, K., & Wiswedel, B. (2008). *KNIME: The Konstanz Information Miner. In Data Analysis, Machine Learning and Application* (ed. Preisach, C. vd.). Springer. 319–326.

Bozkurt, N. (2009). *İşletmelerin Kara Deliği: Hile, Çalışan Hileleri*. İstanbul: Alfa Basım Yayın Dağıtım, 3.Basım.

Bozlak, E. (2019). *İç Denetim Süreçlerinde Hile Denetimi* (Yüksek Lisans Tezi). Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Brownlee, J. (2017). *Master Machine Learning Algorithms: Discover How They Work and Implement Them from Scratch*. Jason Brownlee, v1.12 Edition.

Bruce, P., Bruce, A., & Gedeck, P. (2020). *Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python*. Sebastopol: O'Reilly Media, Second Edition.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357.

Chicco, D., & Jurman, G. (2020). The Advantages of The Matthews Correlation Coefficient (MCC) Over F1 Score and Accuracy in Binary Classification Evaluation. *BMC Genomics*, 21(6).

Coenen, T. L. (2008). *Essentials of Corporate Fraud*. New Jersey: John Wiley & Sons.

Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine learning*, 20(3), 273-297

Çalış, A., Kayapınar, S., & Çetinyokuş, T. (2014). Veri Madenciliğinde Karar Ağacı Algoritmaları ile Bilgisayar ve İnternet Güvenliği Üzerine Bir Uygulama. *Endüstri Mühendisliği Dergisi*, 25(3-4), 2-19.

Çalıyurt, K. & Tezgel, D. T. (2015). Sigorta Şirketlerinde Finansal Tablo Suistimalleri. *Muhasebe ve Denetime Bakış Dergisi*, 46, 33-51

Çamlıfidan, E. (2019). *Sigorta Sektöründe Kasko Sigortası için Makine Öğrenmesi Kullanılarak Sahte Hasarların Tahmini* (Yüksek Lisans Tezi). Bahçeşehir Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Çatıkkaş, Ö., & Çalış, E. (2010). Hile Denetiminde Proaktif Yaklaşımlar, *Muhasebe ve Finansman Dergisi*, 45, 146-156.

Çelik, Ö., & Altunaydın, S. S. (2018). A Research on Machine Learning Methods and Its Applications. *Journal of Educational Technology & Online Learning*, 1(3), 25-40.

Demir, V., Özdarak, E., & Sebilcioğlu, F. (2022). *İşletme Yolsuzlukları*. https://archive.ismmmo.org.tr/YAYINLAR/e_kitap/15042022_isletme_yolsuzluklari.pdf, Erişim Tarihi: 25.05.2022

Deniz Başar, Ö., & Güneren Genç, E. (2020). Ülkelerin Güvenli Olmalarının Tahmininde Lojistik Regresyon, Yapay Sinir Ağları ve Moora Yöntemlerinin Karşılaştırılması. *Journal of Life Economics*, 7(2), 123–134.

Doğan, S., Kayakıran, D. (2017). İşletmelerde Hile Denetiminin Önemi. *Maliye Finans Yazıları*, 108, 167-188.

Dorminey, J., Scott Fleming, A., Kranacher, M. J., & Riley, R. A. (2012). The Evolution of Fraud Theory. *Issues in Accounting Education*, 27(2), 555–579.

Dönmez, A., & Bağışlar, H. (2017). Bankacılık Sektöründe Yaşanan Hile Olaylarına Yönelik Bir Araştırma. *International Journal of Academic Value Studies*, 3(16), 170-180.

Durkaya, İ. (2008). *Türk Bankacılık Sisteminde Suiistimallerin Engellenmesinde İç Denetimin Rolü* (Yüksek Lisans Tezi). İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Elhakan, N. (2017). *Hile Soruşturması ve Denetimin Yürütülmesi* (Yüksek Lisans Tezi). İstanbul Arel Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Erol, S. (2016). *Hile Denetiminde Proaktif Yaklaşımlar* (Yüksek Lisans Tezi). İstanbul Ticaret Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Emir, Murat. (2008). Hile Denetimi. *Mali Çözüm Dergisi*, 86, 109-121.

Friedman, N., Geiger, D., & Goldszmidt, M. (1997). Bayesian Network Classifiers. *Machine Learning*, 29, 131-163.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144.

Geng, H. (2017). *Internet of Things and Data Analytics Handbook*. New Jersey: John Wiley & Sons.

Geren, Y. (2020). *Makine Öğrenmesi ile Sigorta Sektöründe Sahte Hasar Tespiti* (Yüksek Lisans Tezi). İstanbul Aydın Üniversitesi, Bilgisayar Mühendisliği Ana Bilim Dalı, İstanbul.

Golden, T., Skalak, S. & Clayton, M. (2006). *A Guide to Forensic Accounting Investigation*. New Jersey: John Wiley & Sons.

Görgel, P. (2011). *Görüntü İşleme Teknikleri Kullanılarak, 2-Boyutlu Mamografik Verilerde Kanserli Bölge Tanısı* (Doktora Tezi). İstanbul Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Güneren, H. (2015). *Destek Vektör Makineleri Kullanarak Gömülü Sistem Üzerinde Yüz Tanıma Uygulaması*. (Yüksek Lisans Tezi). Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Güven, Ö. (2021). *Makine Öğrenmesi Teknikleriyle Mobil Ödemede Sahtekarlık Tespiti* (Yüksek Lisans Tezi). Dokuz Eylül Üniversitesi, Sosyal Bilimler Enstitüsü, İzmir.

Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques*. Waltham: Morgan Kaufmann Publishers.

Hazım, L. R. (2018). *Four Classification Methods Naive Bayesian, Support Vector Machine, K-Nearest Neighbors and Random Forest Are Tested For Credit Card Fraud Detection* (Yüksek Lisans Tezi). Altınbaş Üniversitesi, Fen Bilimleri Enstitüsü, İstnabul.

Joshi, A. V. (2020). *Machine Learning and Artificial Intelligence*. Cham: Springer.

Kaban, İ., Gük, M., & Eryılmaz, Y. (2018). Hile Denetimi Araçları Olarak Hile Riski Değerlemesi ve Kırmızı Bayraklar: Bankacılıkta Örnek Bir Uygulama. *Bankacılar Dergisi*, (106), 59-68.

Kaban, İ. (2018). Bankalarda Hile Denetimi: *Merkezden Denetim ve Personel Algısına İlişkin Nitel Bir Araştırma* (Doktora Tezi). Gaziosmanpaşa Üniversitesi, Sosyal Bilimler Enstitüsü, Tokat.

Kandemir, C., & Kandemir, Ş. (2012). Muhasebe Hilelerinin Önlenmesi ve Ortaya Çıkarılmasında Kullanılan Çağdaş Araç ve Yöntemler. *Mali Çözüm*, 22(114), 37-70.

Kaya, H. P., & Uzay, Ş. (2018). Hileli Finansal Raporlama ve Bağımsız Denetçinin Sorumluluğu. *Muhasebe Bilim Dünyası Dergisi*, 20 (Özel Sayı), 721- ös740

Kelleher, J. D., Namee, B. M., & D, Aoife. (2015). *Fundamentals of Machine Learning For Predictive Data Analytics*. London: The MIT Press.

KPMG Türkiye, Etik ve İtibar Derneği. (2020). *Bir Suistimalcinin Profili*. https://assets.kpmg/content/dam/kpmg/tr/pdf/2020/12/bir-suistimalcinin-profili-2020.pdf, Erişim Tarihi: 18.10.2022

Kramer, B. (2015). Trust, but verify: fraud in small businesses, *Journal of Small Business and Enterprise Development*, 22(1), 4-20.

Mitchell, T. M. (1997). *Machine Learning*. New York: McGraw-Hill.

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). *Foundations of machine learning*. London: The MIT Press, Second Edition.

Mui, G., & Mailley, J. (2015). A tale of two triangles: comparing the Fraud Triangle with criminology's Crime Triangle. *Accounting Research Journal*, 28(1), 45-58.

Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The Application of Data Mining Techniques in Financial Fraud Detection: A Classification Framework and An Academic Review of Literature. *Decision Support Systems*, 50(3), 559–569.
Osmanoğlu, U. Ö. (2021). *Makine Öğrenmesi Sınıflama Algoritmalarıyla Kalp Yetersizliği Mortalitesinin Tahminlenmesi* (Doktora Tezi). Eskişehir Osmangazi Üniversitesi Sağlık Bilimleri Enstitüsü, Eskişehir.

Özdemir, R. (2021). *Makine Öğrenmesindeki Sınıflandırma Yöntemlerinin Karşılaştırılması Ve E-Ticaret Üzerinde Bir Uygulama* (Yüksek Lisans Tezi). T.C. İstanbul Ticaret Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Özmen, Ş. (2001). İş Hayatı Veri Madenciliği ile İstatistik Uygulamalarını Yeniden Keşfediyor. *V. Ulusal Ekonometri ve İstatistik Sempozyumu*, *Çukurova Üniversitesi* Adana, Türkiye, Eylül 19-22.

Roy, A. M. (2001). *A Manager's Dilemma: Corruption-Related Decisionmaking* (Yüksek Lisans Tezi). University of Canterbury, New Zealand. Erişim Adresi: https://ir.canterbury.ac.nz/bitstream/handle/10092/16759/Roy_thesis_2001.pdf?sequence=1&isAllowed=y

Sánchez, M., Torres, J., Zambrano, P., & Flores, P. (2018). FraudFind: Financial Fraud Detection by Analyzing Human Behavior. *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, 281-286.

Sandhu, N. (2016). Behavioural Red Flags of Fraud A Qualitative Assessment, *Journal of Human Values*, 22(3), 221-237.

Sandhu, N. (2019). Behavioural Red Flags of Fraud: An Ex Post Assessment of Types and Frequencies. *Global Business Review*, 21(2), 507–525.

Soileau, J., Soileau, L., & Sumners, G. (2015). The Evolution of Analytics and Internal Audit, *EDPACS*, 51(2), 10-17.

Tatlıtürk, U. Y. (2021). *İç Denetimde Veri Analitiğinin Kullanımı ve Bir Uygulama* (Yüksek Lisans Tezi). Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Terzi, S. (2012). Hile ve Usulsüzlüklerin Tespitinde Veri Madenciliğinin Kullanımı, *Muhasebe ve Finansman Dergisi*, 54, 51-64.

Timofeyev, Y. (2015). Analysis of predictors of organizational losses due to occupational corruption. *International Business Review*, 24(4), 630–641.

Turfan, D. (2020). *Gen Açıklama Verilerinin Sınıflandırılmasında Yeni Bir Özellik Seçimi Yöntemi* (Doktora Tezi). Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

Vona, L. W. (2008). *Fraud Risk Assessment: Building a Fraud Audit Program*. New Jersey: John Wiley & Sons.

Yapı Kredi Bankası. *Kredi Notu Nedir? Kredi Notu Hakkında Bilmeniz Gerekenler*. https://www.yapikredi.com.tr/blog/surdurulebilirlik/finansal- okuryazarlik/ detay/kredi-notu-nedir-kredi-notu-hakkinda-bilmeniz-gerekenler, Erişim Tarihi: 10.10.2022

Yavaş, M., Güran, A., & Uysal, M. (2020). Covid-19 Veri Kümesinin SMOTE Tabanlı Örnekleme Yöntemi Uygulanarak Sınıflandırılması. *Avrupa Bilim ve Teknoloji Dergisi*, Özel Sayı, 258-264.

Yıldırım, S., & Turgut, Ö. (2016). Hilelerinin Demografik Analizi ve Bir Uygulama. *World of Accounting Science*, 18(1), 215-236.

Yıldız, E. & Baskan, T. D. (2014). Muhasebe Hilelerinin Önlenmesinde Kullanılan Araçlar: BİST Şirketleri Üzerine Bir Araştırma. *Muhasebe ve Finansman Dergisi*, (62), 1-18.

Yılmaz, G. (2019). *Hilenin Önlenmesi ve Tespitinde Büyük Veri Analitiğinin Kullanımı ve Türkiye'deki Çalışan Hileleri Üzerine Bir Araştırma* (Doktora Tezi). Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.

Yılmaz, G., Ataman, B., & Ayboğa, H. (2018). Hile Önlenme ve Tespit Etme Yöntemlerinin Etkinliğinin Değerlendirilmesi: Türkiye Araştırması. *Finans Ekonomi ve Sosyal Araştırmalar Dergisi*, 3(3), 646-664.

West, J., & Bhattacharya, M. (2016). Intelligent Financial Fraud Detection: A Comprehensive Review. *Computers & Security*, 57, 47–66.