

Bağımsız Denetimde Makine Öğrenmesi: Teorik Bir Yaklaşım

Machine Learning in External Auditing: A Theoretical Approach

Fatih Faydalı^a , Bilal Solak^b 

^a *Department of Business Administration, Kyrgyz-Turkish Manas University, Bishkek, Kyrgyzstan, fatihfaydali@gmail.com (Corresponding Author)*

^b *Department of Business Administration, Fırat University, Elazığ, Türkiye*

Özet

Bu çalışmada, yapay zekâ ve makine öğrenmesi tekniklerinin muhasebe denetimi alanındaki uygulamaları incelenmektedir. Kapsamlı bir literatür taraması ile muhasebe denetimi alanındaki yapay zekâ ve makine öğrenmesi çalışmaları altı başlık altında organize edilmiş ve sunulmuştur. Denetçi seçimi, denetim görüşü, işletmenin sürekliliği, hile denetimi ve risk yönetimi başlıkları altındaki çalışmalarda kullanılan metotlar, veri setlerinin yapısı ve veri setlerinde yer alan öznitelikler hakkında detaylı bilgi verilmektedir. Muhasebe denetiminde metin madenciliği ve doğal dil işleme başlığı altında yer alan çalışmalar sonuçları ile birlikte tartışılmaktadır. Çalışma şu şekilde organize edilmiştir, birinci bölümde yapay zekâ, makine öğrenmesi, öğrenme stratejileri, derin öğrenme ve metin madenciliği hakkında bilgi verilmektedir. İkinci bölümde yedi başlık altında yapay zekanın ve makine öğrenmesinin denetim alanında nasıl uygulamalarda kullanıldığı mevcut literatür ve uygulamadaki yansımalarının ortaya koymak maksadı ile dört büyük denetim firmasının yayınladığı raporlardan yola çıkılarak tartışılmaktadır.

Anahtar Kelimeler: Muhasebe Denetimi, Yapay Zekâ, Makine Öğrenmesi, Metin Madenciliği, Doğal Dil İşleme

Abstract

In this study, the applications of artificial intelligence and machine learning techniques in the field of auditing are examined. With a comprehensive literature review, artificial intelligence and machine learning studies in the field of auditing are organized and presented under six topics. Detailed information is given about the methods used in the studies, the structure of the data sets, and the features in the data sets under the topics of auditor selection, audit opinion, going concern, fraud detection, and risk management. Studies on the topic of text mining and natural language processing in auditing are discussed with their results. The study is organized as follows, in the first part, information about artificial intelligence, machine learning, learning strategies, deep learning, and text mining is given. In the second part, under seven topics, how artificial intelligence and machine learning are used in the field of auditing is discussed based on the current literature and the reports published by the big four audit firms in order to reveal reflections in practice.

Keywords: Auditing, Artificial Intelligence, Machine Learning, Text Mining, Natural Language Processing

Acknowledgments

Bu çalışma, 25-27 Kasım 2021 tarihlerinde MÖDAV tarafından düzenlenen 18. Uluslararası Muhasebe Konferansında sunulan bildirinin, yeniden düzenlenmiş ve genişletilmiş halidir.

For Citation: Faydalı, F., Solak B. (2023). Bağımsız Denetimde Makine Öğrenmesi: Teorik Bir Yaklaşım. *Journal of Academic Value Studies*, 9(1), 95-114. <http://dx.doi.org/10.29228/javs.68546>

Received: 24.02.2023 Accepted: 29.03.2023

This article was checked by *intihal.net*



1. Giriş

Hızlı bir şekilde dönüşüm geçiren, küreselleşen ve sürekli rekabetin arttığı pazarlarda, işletmeler rekabetçi kalmak amacı ile farklı ve yenilikçi yönetim yaklaşımları ile yeni bilişim sistemlerini benimsemekte ve uygulamaktadır. İşletmeler, veri odaklı teknolojilerin geliştirilmesi ve adaptasyonu konusunda oldukça açık bir tutum sergilemektedir. Özellikle yapay zekâ ve ilgili teknolojiler geleceğin muhasebesinde çok önemli bir yer tutmaktadır (Elliott, 1992).

Muhasebe “Bir örgütün kaynaklarının oluşumunu, bu kaynakların kullanılma biçimini, örgütün işlemleri sonucunda bu kaynaklarında meydana gelen artış ve azalışları ve örgütün finansal açıdan durumunu açıklayan bilgileri üreten ve bunları ilgili kişi ve kuruluşlara ileten bilgi sistemidir (Sevilengül, 2016, s. 3). Muhasebe, işletmeler içerisinde bilgi yönetimi sürecinin temel unsurudur. Muhasebe işletmelerdeki raporlama, maliyetlendirme, tahmin ve değerlendirme ile denetleme gibi pek çok işlevi yerine getiren kapsamlı bir bilgi sistemidir. Muhasebe bilgi sisteminin, tanımda da ifade edildiği üzere, en doğal fonksiyonu raporlamadır. Sistemin ürettiği raporlarda yer alan bilgiler yapısal veya yapısal olmayan formatta olabilmektedir.

Muhasebe denetimi ise “Bir ekonomik birim veya döneme ait bilgilerin önceden belirlenmiş ölçütlere olan uygunluk derecesini araştırmak ve bu konuda rapor düzenlemek amacıyla bağımsız bir uzman tarafından yapılan kanıt toplama ve değerlendirme sürecidir.” (Bozkurt, 2018, s. 29). Denetim süreci muhasebe bilgi sistemi tarafından üretilen yapısal ve yapısal olmayan formattaki verileri girdi olarak kabul etmektedir. Ayrıca denetim, yapısı nedeni ile belirsizliği ve riski arttıran analiz ve karar süreçlerine de sahiptir.

Mevcut denetim ortamı tartışma götürmez bir şekilde büyük belirsizliğe sahiptir. Mali açıdan büyük kayıplara neden olan denetim skandallarından sonra tüm Dünya’da yapılan düzenlemeler ile denetim sürecinin girdileri giderek yapısal formattan uzaklaşmışlardır. Bu durum nedeni ile muhasebe ve denetim alanı, veri analitiği ve yapay zeka teknolojilerindeki gelişmelerden etkilenmekte ve bir dönüşüm geçirmektedir (Lombardi vd., 2014; Wagner, 2016).

Muhasebe denetimi sürecinin, son yıllarda teknoloji tabanlı bir dönüşüm geçirdiği söylenebilir. Akıllı sistemler, denetim sürecinin etkinliğini arttırmaktadır (Granlund, 2011). Bu nedenle yapay zekanın ve dolayısı ile onun bir alt dalı olan makine öğrenmesinin, muhasebe ve denetim uygulamaları bağlamında daha fazla incelenmesi gerektiği söylenebilir (Debreceňy, 2011).

Denetim alanının yakın gelecekte çok farklı bir yapıya sahip olması beklenmektedir (Ramlukan, 2015). Blok zinciri teknolojisi, veri analitiği, yapay zekâ, makine öğrenmesi, nesnelerin interneti gibi kavramlar günümüzde örgüt çevresini etkileyen ve dönüştüren unsurlar olarak nitelendirilmektedir. İşletmelerin günlük operasyonları pek çok farklı teknolojik gelişmeden etkilenmektedir. Denetçiler yeni teknolojik gelişmeleri, denetim sürecine dahil etme konusunda teşvik edilmekte ve böylece denetim sürecinin etkinliğinin artması beklenmektedir (Dai & Vasarhelyi, 2016; Manyika vd., 2017). Ayrıca denetim sürecinin kendini tekrarlayan bir yapıya sahip olması sebebi ile yapay zekâ teknolojilerinin sürece entegrasyonu, denetim sürecinin otomatikleşmesine büyük katkı sağlayacaktır ve denetim sürecinde kendini tekrarlayan unsurların yönetimini kolaylaştıracaktır. Bu yolla denetçi, denetim sürecinde değer yaratma adına daha fazla imkân elde edebilecek ve denetçinin süreç içerisinde değer yaratma potansiyelinin artması denetimin verimliliğini de artıracaktır. Yakın gelecekte, denetim firmaları yeni denetim metotları ile geleneksel denetim uygulamalarını terk edecektir (Luo vd., 2018; Raphael, 2015)

Denetim sürecinde otomasyonun kullanımının artması, süreçte insan kaynaklı hataların azalmasına ve performansın artmasına katkı sağlayacaktır. Veri analitiğinin denetçiler adına faydalı ve anlamlı hale getirilmesi için otomatikleştirilmiş süreçlerin denetim sürecine adapte edilmesi gerekmektedir (T. Sun & Vasarhelyi, 2017). Denetim sürecinde yapay zekâ ve veri analitiğinin kullanımı, denetim kalitesinin artmasına olanak tanıyacaktır. Denetim sürecinde, yapay zeka ve veri analitiğinin kullanılması denetlenen işletmenin süreçlerinin, faaliyetlerinin ve muhasebe bilgi sisteminin risk yönetim bağlamında daha iyi analiz edilmesine olanak tanıyacaktır (Manita vd., 2020).

Veri analitiği uygulamaları, işletmelere gelecek eğilimleri tahmin etmek ve daha etkin karar almak adına büyük yarar sağlamaktadır. Veri analitiği ve makine öğrenmesi yöneticilere bir problem hakkında proaktif karar alabilmek adına işletmenin ürettiği veri yığınları içerisinde mantıklı ve nedensel örüntüler sağlayabilmektedir (Yigitbasioglu & Velcu, 2012).

Veri analitiği ve makine öğrenmesi, denetimin sürecinde denetçilerin elde ettiği kanıtların kalitesinin ve genel denetim etkinliğinin artmasına büyük ölçüde katkı sağlamaktadır (C. E. Brown vd., 2007; Chan & Vasarhelyi, 2011). Veri analitiği sürekli denetim uygulamalarını da kolaylaştırabilmekte ve ayrıca yönetsel kontrol sistemlerinin işleyişini köklü bir şekilde değiştirebilmektedir (C. E. Brown vd., 2007; Granlund, 2011; Sutton vd., 2011). Makine öğrenmesi, işletmelerin performans ölçüleri arasındaki istatistiksel ilişkileri kolaylıkla tanımlamasına, bir olay ile ilgili nicel tahminlerde bulunma

ve yöneticilerin nitel yargılarını tamamlayarak kontrol sağlamasına olanak tanımaktadır. Makine öğrenmesi, işletmelerin geleneksel teknikler yardımı ile tespit edemeyeceği ya da tespiti için uzun zaman ayırması gereken örüntüleri hızlı bir şekilde ortaya koyabilmektedir.

Denetim alanında dijital bir dönüşüm kaçınılmazdır. Bu nedenle veri analitiği, makine öğrenmesi, metin madenciliği ve doğal dil işleme gibi denetim sürecinin verimliliğini artıracak olan yeni ve güncel yaklaşımların farklı denetim görevlerini icrası için kullanıldığı modellerin geliştirilmesi ve uygulamada da yaygınlaştırılması gerekmektedir. Bu çalışmada, muhasebe denetiminde yapay zekâ teknikleri uygulamaları ile ilgili kapsamlı bir literatür çalışması sunmak amaçlanmaktadır. Çalışma şu şekilde organize edilmiştir. Çalışmanın ikinci bölümünde makine öğrenmesi kavramı hakkında bilgi verilmektedir. Üçüncü bölümünde, yapısal formattaki finansal verileri ve/veya yapısal olmayan metinleri veri olarak kabul eden çalışmalar kapsamlı bir literatür taraması ile aktarılmaktadır. Dördüncü bölümde ise literatür taraması sonuçları tartışılmaktadır.

2. Yapay Zekâ, Makine Öğrenmesi ve Metin Madenciliği Kavramları

Yapay zekâ, temel olarak insan beyninin çalışma şeklini taklit eden yazılım veya donanımları tanımlamak için kullanılan bir ifadedir (Tung vd., 2004). Yapay zekâ uygulamalarının temel amacı, düşünmekten çok temsil ve doğru şeyleri yapmaktan çok insan beyni ile yarışmaktır. Yapay zeka dengeli kararlar veren ve çevresini tanıyarak amaca ulaşmak için aksiyon alan programlardır (Issa vd., 2016). Veri Analitiği ve makine öğrenmesi kavramları geçtiğimiz yirmi yıl içerisinde giderek popülerlik kazanmış olan kavramlardır. Bireyler ve kurumlar tarafından üretilen veri miktarı artıkça elde edilen büyük boyutlu veri yığınlarının analizi için daha gelişmiş yöntemlere duyulan ihtiyaç da artmaktadır. Bu nedenle veri analitiği ve makine öğrenmesi metodlarından işletmenin her fonksiyonunda yararlanılmaya başlanılmıştır (Smola & Vishwanathan, 2008).

2.1. Makine Öğrenmesi

Makine öğrenmesi yapay zekanın bir alt türüdür. Makine öğrenmesi temel olarak geçmiş verilerden öğrenmek sureti ile geleceği tahmin eden modeller şeklinde tanımlanabilir. Makine öğrenmesi metodları öznelilikler arasındaki ilişkilerden öğrenme sürecini gerçekleştirir. Bir başka ifade ile algoritmalar, koşullu bir yönlendirme olmaksızın çok boyutlu bir veri uzayındaki anlamlı örüntülerden öğrenmektedir (Abbott, 2014). Makine öğrenmesi veri setindeki öznelilikler arasındaki ilişkiyi en iyi tanımlayacak olan muhtemel fonksiyonlar arasından, en doğru olanı seçme sürecidir (Kelleher, 2019, s. 11)

2.1.1. Öğrenme Stratejileri

Makine öğrenmesinde temel olarak gözetimli öğrenme ve gözetimsiz öğrenme olmak üzere iki öğrenme stratejisi bulunmaktadır.

2.1.1.1. Gözetimli Öğrenme

Gözetimli öğrenme makine öğrenmesi uygulamaları içerisinde yaygın olarak kullanılmaktadır. Gözetimli öğrenme modellerinde algoritma, modelin girdi ve çıktısı ile ilgili bilgiye sahiptir. Model eğitim sürecinde, eğitim veri setinden öğrenme yolu ile her yeni girdi için çıktı tahmin eden fonksiyon oluşturmaktadır (Hastie vd., 2009). Gözetimli öğrenme algoritmaları, girdi değişkenleri ile girdi vektörleri arasındaki ilişkiyi tanımlamakta ve geleceğe yönelik tahmin yürütmek maksadı ile tanımlanan matematiksel ilişkiyi yeni değerlere uygulamaktadır. Öğrenme süreci sonucunda gözlemlenen ve tahmin edilen çıktılar hedef değişkenler ya da hedef vektörler olarak tanımlanmaktadır.

Gözetimli öğrenme modelleri sınıflandırma ve regresyon problemleri olmak üzere iki başlık altında toplanmaktadır (Bishop, 2006, s. 3). Sınıflandırma Problemleri, girdi veri setini belirli sayıda sınıfa ayırmayı amaçlamaktadır. Regresyon problemleri ise girdiyeye dayalı olarak, değişkene devamlı bir değer atamayı amaçlamaktadır.

2.1.1.2. Gözetimsiz Öğrenme

Gözetimsiz öğrenme algoritmaları, karmaşık ve etiketlenmemiş olan veri setinden yararlı örüntüler elde etme amaçlanmaktadır. Verinin etiketlenmemiş olması, analiz aşamasında verinin ele alınış biçimi için genel bir doğrunun bulunmaması anlamına gelmektedir. Kümeleme, etiketlenmemiş veriler ile yürütülmektedir. Kümeleme modelleri boyut azaltma için çok başarılı bir yaklaşımdır. Model benzerlik veya uzaklıklar üzerine inşa edilmekte olup, belirlenecek olan kümeler içerisindeki benzerliğin maksimizasyonu ya da bir başka ifade ile uzaklıkların minimizasyonunu amaçlanmaktadır

(Abbott, 2014) Kümeler kendi içlerinde homojen bir yapı sergilerken, evren içerisindeki diğer kümeler ile heterojen yapı göstermektedir. Kümeleme modellerinde bir diğer yaklaşım değişkenler arasındaki varyans-kovaryans yapısının açıklanmasına dayanmaktadır, “ p ” değişkenlerinin birkaç farklı doğrusal kombinasyonu yolu ile farklılığın çoğu, daha az küme ile açıklanabilmektedir (Johnson & Wichern, 2007).

Gözetimsiz öğrenme problemleri kümeleme ve boyutsal küçültme olmak üzere iki başlık altında toplanmaktadır. Kümeleme, etiketlenmiş veri içeren, veri setinin benzer veri merkezleri etrafında kümelere ayırmayı amaçlamaktadır. Boyutsal küçülme ise genellikle veri ön işleme aşamasının ya da diğer makine öğrenmesi metotları dâhilinde veriyi dönüştürme aşamasının bir parçasıdır.

2.2. Derin Öğrenme

Derin öğrenme belirli yapay zekâ metotlarını ifade etmektedir ancak kesin bir tanımı bulunmamaktadır. Derin öğrenme terimi, araştırmacılar tarafından farklı biçimlerde çalışılmıştır. Derin öğrenme, genellikle çok sayıda katmandan oluşan yapay sinir ağları modelleri ile ilgilidir (Schmidhuber, 2015).

Derin öğrenme birden fazla katmana sahip olan modellerin veriden öğrenmesini sağlamaktadır. Derin öğrenme yöntemleri, konuşma tanımlama, görsel tanımlama ve obje tespiti gibi farklı alanlarda kullanılmaktadır. Derin öğrenme, her bir geçmiş katmandaki temsili kullanarak makinenin içsel parametrelerini ne şekilde değiştireceğini belirlemek amacı ile geri yayılım algoritması kullanmakta ve büyük veri setleri içerisindeki karmaşık yapıları keşfetmektedir (LeCun vd., 2015)

Derin öğrenme, makine öğrenmesinden oldukça farklı bir yapıya sahiptir. Makine öğrenmesi tahmin ortaya koymaktayken, derin öğrenme tahmin ortaya koyarken aynı zamanda yaratıcı bir davranış sergilemektedir. Derin öğrenme, yapısal olmayan veri yapılarını yapısal forma sokma maksadı ile kullanılmaktadır.

2.2. Metin Madenciliği

Metin madenciliği, yapısal formatta olmayan metinleri kullanarak bilgi keşfetme sürecidir (Mecca vd., 2007; Witten, 2003) bu noktada makine öğrenmesi metotlarından farklılaşmaktadır. Makine öğrenmesi metotları yapısal veriler ile çalışmaktayken, metin madenciliği yapısal olmayan verileri konu almaktadır.

Metin madenciliği ile analiz edilen metinlerdeki görülmeyen gizli ilişkiler, hipotezler ve trendler tespit edilmeye çalışılır. Metin madenciliğinden, metinleri kategorize etmek, etiketlemek, kümelemek, özetlemek veya metinler içerisindeki kelimelerin frekanslarını belirlemek gibi amaçlar için yararlanılmaktadır (Campesato, 2021; Çelik, 2020). Metin madenciliği, metin ön işleme, metin dönüştürme, öznelik seçimi, bilgi keşfi ve analiz aşamalarından oluşmaktadır.

3. Muhasebe Denetiminde Yapay Zekâ, Makine Öğrenmesi, Metin Madenciliği ve Doğal Dil İşleme

Denetim sürecinde, denetçiler denetim verimliliğini arttırmak için yeni araçlara ihtiyaç duymaktadır. Denetim doğası gereği veri analitiği ve makine öğrenmesi metotlarının kullanımı için araştırmacı ve uygulamacılara motivasyon sağlamaktadır. Denetim süreci risk, tekrarlayan yapısal ve yapısal olmayan kararlar, eksik bilgi ve belirsizliğin değerlendirilmesi süreçlerini kapsamaktadır (Baldwin vd., 2006). Denetim sözleşmesinin yapılmasını takiben denetim süreci risk değerlendirme ve denetimin kapsamını ve hedeflerini kapsayan bir denetim planının oluşturulması ile devam eder. Bunu takiben denetçi, denetim kanıtı toplar ve analiz eder ve ayrıca yönetim tarafından sağlanan bilgilerin güvenilirliğinin yanı sıra iç kontrol yapısına ilişkin görüş oluşturur. Denetim süreci denetçinin denetlenen işletme ile ilgili görüş oluşturması ile tamamlanır. Yapay zeka ve makine öğrenmesi metotlarının denetim süreci içerisinde farklı amaçlar ile kullanımını örnekleyen pek çok çalışma mevcuttur (Amani & Fadlalla, 2017). Denetim sürecinin her aşamasında kullanılan analitik prosedürlerin etkinliğini arttırmak maksadı ile de makine öğrenmesi metotlarından yararlanılmaktadır (Jans vd., 2014; Koskivaara, 2004).

3.1. Denetçi Seçimi Problemi

Denetim süreci her ne kadar gün geçtikçe dijitalleşse de denetçinin süreç içerisindeki önemi tartışılmaz bir gerçektir. İnsan odaklı bir süreç olan denetimde denetçinin seçimi de çok önemli bir problemdir. Bu nedenle literatürde denetçi seçimi problemini konu alan modeller geliştirilmiştir.

Tablo 1. Denetçi Seçimi Problemi ile İlgili Literatürdeki Çalışmalar

	Çalışmanın Başlığı	Kullanılan Metotlar	Veri Setinin Yapısı ve Öznitelikler
1	Applying Data Mining Methodologies for Auditor Selection (Kirkos, Spathis, & Manolopoulos, 2007a)	Karar Ağaçlar – C4.5 Çok Katmanlı Algılayıcılar K-En Yakın Komşu Algoritması	Çalışmada kullanılan veri seti 2003-2005 yılları arasında denetim firması değiştirmiş olan işletmelerden oluşmaktadır. Veri setinde toplam 338 gözlem yılı yer almaktadır. Çalışmada nicel nitelikteki mali öznitelikler ile birlikte nitel özniteliklerden faydalanılmıştır.
2	Support vector machines, Decision Trees and Neural Networks for auditor selection (Kirkos vd., 2008)	Karar Ağaçlar – C4.5 Çok Katmanlı Algılayıcılar *DMV-Doğrusal DVM- Quadratic DVM- Radial Tabanlı	Çalışmada kullanılan veri seti 2003-2005 yılları arasında denetim firması değiştirmiş olan işletmelerden oluşmaktadır. Veri setinde toplam 338 gözlem yılı yer almaktadır. Çalışmada ANOVA analizi ile anlamlılığı test edilmiş olan on sekiz öznitelik ile birlikte dört kukla öznitelik kullanılmıştır.
3	Audit-firm group appointment: an artificial intelligence approach (Kirkos vd., 2009)	Karar Ağaçlar – C4.5 Çok Katmanlı Algılayıcılar K-En Yakın Komşu Algoritması Lojistik Regresyon	Çalışmada kullanılan veri seti 2003-2005 yılları arasında denetim firması değiştirmiş olan işletmelerden oluşmaktadır. Veri setinde toplam 338 gözlem yılı yer almaktadır. Çalışmada ANOVA analizi ile anlamlılığı test edilmiş olan on sekiz öznitelikten faydalanılmıştır.
4	Auditor Choice Prediction Model Using Corporate Governance and Ownership Attributes: Machine Learning Approach (Rahman vd., 2021)	Naive Bayes Sınıflandırıcısı Karar Ağaçları Rastgele Orman Algoritması Derin Öğrenme Algoritması	Çalışmada kullanılan veri seti 2000 – 2007 yılları arasında kapsamaktadır. Veri setinde toplam 2262 gözlem yılına yer verilmiştir. Çalışmada üç ana kategoride, işletme karakteristiği, kurumsal yönetim, ortaklık yapısı, toplam on üç öznitelik yer almaktadır.

*Destek Vektör Makineleri

3.2. Denetim Görüşü

Denetim sürecinin sonucunda, denetçi tüm denetim sürecini özetleyen, denetlenen işletme ile ilgili kararını denetim raporunda finansal bilgi kullanıcıları ile paylaşmaktadır. Veri analitiği ve makine öğrenmesi metotları literatürde denetim görüşü tahmin etmek maksadı kullanılmıştır.

Tablo 2. Denetim Görüşü ile İlgili Literatürdeki Çalışmalar

	Çalışmanın Başlığı	Kullanılan Metotlar	Veri Setinin Yapısı ve Öznitelikler
1	Explaining qualifications in audit reports using a support vector machine methodology (Douplos vd., 2005)	DMV-Doğrusal DVM- Quadratic DVM- Radial Tabanlı	Çalışmada 1998-2003 yılları arasında kapsayan bir veri seti yer almaktadır. Birleşik Krallık faaliyet gösteren 1754 işletmeye ait toplam 6048 gözlem yılı kullanılmıştır. Bunları 859 adedi şartlı görüştür. Beş mali oran ile birlikte iki adet kredi riskini temsil eden yapay kukla değişken yer almaktadır.
2	Identifying Qualified Auditors' Opinions: A Data Mining Approach (Kirkos, Spathis, Nanopoulos, vd., 2007)	Karar Ağaçları Yapay Sinir Ağları Bayes Ağları	Çalışmada kullanılan veri seti 1995-2004 yıllarını kapsamaktadır. Veri setinde 450 adet işletmeye ait gözlem yer almaktadır. Çalışmada ANOVA testi sonucunda anlamlı olduğu tespit edilen on altı öznitelik yer almaktadır.
3	Modeling the auditors' opinions by using association rules (Kirkos, 2009)	Birliktelik Kuralları Analizi	Çalışmada kullanılan veri seti FAME veri tabanından elde edilmiştir. Veri seti toplam 450 adet işletmenin verilerini içermektedir. Çalışmada 24 adet mali oran öznitelik olarak yer almaktadır.
4	Finding Rules for Audit Opinions Prediction Through Data Mining Methods (Saif vd., 2012)	Karar Ağaçları DVM – Doğrusal DVM – Polinomsal DVM – Radial tabanlı DVM - Sigmoid	Çalışmada kullanılan veri seti 2001-2007 yılları arasında kapsamaktadır. Veri seti toplam 1018 gözlemden oluşmaktadır. Çalışmada toplam 29 adet mali oran ve Altman Z-skor öznitelik olarak yer almaktadır.
5	An Expert System with Neural Network and Decision Tree for	Karar Ağaçları	Çalışmada kullanılan veri seti 2001-2007 yılları arasında kapsamaktadır. Veri seti toplam 1018 gözlemden oluşmaktadır.

	Predicting Audit Opinions (Saif vd., 2013)	Yapay Sinir Ağları	Çalışmada toplam 29 adet mali oran ve Altman Z-skor öznelik olarak yer almaktadır.
6	Predicting Qualified Audit Opinions Using Financial Ratios: Evidence from the Istanbul Stock Exchange (Yaşar vd., 2015)	Karar Ağaçları – C5.0 Lojistik Regresyon Diskriminant Analizi	Çalışmada kullanılan veri seti 2010 – 2013 yılları arasını kapsamaktadır. Veri setinde yer alan gözlem yılı Borsa İstanbul A.Ş' de işlem gören işletmelere aittir. Çalışmada toplam 12 adet mali oran öznelik olarak yer almıştır.
7	Predicting the type of auditor opinion: Statistics, machine learning, or a combination of the two? (Stanišić vd., 2019)	Karar Ağaçları – C5.0 Rastgele Orman Algoritması Regularize Rastgele Orman Algoritması Stokastik Gradient Boosting Extreme Gradient Boosting K-En Yakın Komşu Algoritması Çok Katmanlı Algılayıcılar DVM-Radial Tabanlı Doğrusal Diskriminant Analizi Lojistik Regresyon Probit Regresyon Karma Etkili Lojistik Regresyon	Çalışmada 13561 adet denetim görüşünün yer aldığı bir veri seti yer almaktadır. Sırası ile 71 olumsuz, 644 görüş bildirmekten kaçınma, 3706 şartlı görüş ve 9140 olumlu görüş yer almaktadır. Literatüre dayalı olarak seçilen ve veriye dayalı olarak üretilen toplam 1546 uygun öznelik tahminci belirlenmiştir.
8	Bağımsız Denetim Görüşlerinin Makine Öğrenmesi Algoritmaları ile Tahmin Edilmesi (Büyüktanır & Toraman, 2020)	Lojistik Regresyon Naive Bayes Sınıflandırıcısı Rastgele Gradyan Azaltma K-En Yakın Komşu Algoritması Karar Ağaçları Rastgele Orman Algoritması AdaBoost Çok Katmanlı Algılayıcılar Karesel Diskriminant Analizi Extreme Gradient Boosting AdaBoosted Karar Ağacı	Çalışmada kullanılan veri seti 2009 – 2019 yıllarını kapsamaktadır. Çalışmada toplam 7128 gözlem yılına ait denetim görüşü kullanılmıştır. Denetim görüşlerinden 6259 adedi olumlu, 777 adedi şartlı ve 8 adedi olumsuz görüştür. 83 gözlem yılı için denetçi görüş bildirmekten kaçınmıştır. Çalışmada veri ön işleme sonrasında 52 adet özneliğe yer verilmiştir.
9	Predicting Audit Opinion in Consolidated Financial Statements with Artificial Neural Networks (Sánchez-Serrano vd., 2020)	Çok Katmanlı Algılayıcılar Radyal Tabanlı Fonksiyon Ağları Derin öğrenme Ağları	Çalışmada kullanılan veri setinde toplam 298 gözleme yer verilmiştir. Veri seti 87 şartlı görüş ve 211 olumlu görüşten oluşmaktadır. Veri setinde literatüre dayanarak seçilen 35 öznelik yer almaktadır.

3.3. İşletmenin Sürekliliği

Denetim sürecinde, denetlenen işletmenin sürekliliği de denetçinin değerlendirmesi gereken konulardan birisidir. Bu kapsamda denetçinin işletmenin sürekliliği ile görüşünü desteklemek amacı ile farklı yöntemler ile modeller geliştirilmiştir.

Tablo 3. İşletmenin Sürekliliği ile İlgili Literatürdeki Çalışmalar

	Çalışmanın Başlığı	Kullanılan Metotlar	Veri Setinin Yapısı ve Öznitelikler
1	Going concern prediction using data mining techniques (Chye Koh & Kee Low, 2004)	Lojistik Regresyon Karar Ağaçları Yapay Sinir Ağları	Çalışmada kullanılan veri seti Koh (1991) tarafından yürütülen çalışmadan alınmıştır. Veri seti toplam 330 gözlemden oluşmaktadır. Gözlemlerden 165 adedi süreklilik görüşü almıştır. Veri setinde bulunan her bir gözlem, mali verilerden elde edilen 6 öznitelik ile temsil edilmektedir.
2	Analyzing the financial distress of Chinese public companies using probabilistic neural networks and multivariate discriminate analysis (Wu vd., 2008)	Çoklu Diskriminant Analizi Olasılıksal Yapay Sinir Ağları	Çalışmada kullanılan veri setinde toplam 48 adet işletmenin verilerinden yararlanılmıştır. Bunlardan 32 adedi eğitim verisini oluştururken, 16 adedi test verisini oluşturmaktadır. Mali sıkıntılı olarak tanımlanan işletmeler Çin’ de ST(Special Treatment) etiketi almış olan işletmelerdir. Veri seti dengeli forma sahiptir. Çalışmada 7 adet mali oran öznitelik olarak yer almıştır.
3	Predicting going concern opinion with data mining Predicting going concern opinion with data mining (Martens vd., 2008)	AntMiner+ Karar Ağaçları – C4.5 Lojistik Regresyon DVM – Radyal Tabanlı Fonksiyon Çoğunluk Oylamasına dayalı Topluluk Öğrenmesi	Çalışmada kullanılan veri 271’i süreklilik görüşü olmak üzere toplam 10047 gözlem yılı içermektedir. Çalışmada işletmenin mali durumunu yansıtan 14 adet öznitelik yer almaktadır.
4	Financial failure prediction using efficiency as a predictor (Xu & Wang, 2009)	Lojistik Regresyon Çoklu Diskriminant Analizi Destek Vektör Makineleri	Çalışmada kullanılan veri seti 1999 – 2005 yılları arasında kapsamakta ve 60 adedi ST (Special Treatment) etiketi almış olmak üzere toplam 120 adet gözlemden oluşmaktadır. Çalışmada istikrar, karlılık, büyüme, faaliyet ve nakit akışı kategorileri altında toplam 20 mali oran muhtemel öznitelikler olarak belirlenmiştir. Bunlar içerisinde anlamlı olan 8 adedi seçilmiştir.
5	Business failure prediction using decision trees (Gepp vd., 2010)	Tekrarlamalı Ayırma Analizi Çoklu Diskriminant Analizi Karar Ağaçları – C5.0 Karar Ağaçları- CART	Çalışmada kullanılan veri seti karşılaştırmaya olanak tanınması maksadı ile Frydman vd., (1985) tarafından kullanılmış olan veri setinin aynısıdır.
6	Prediction of hotel bankruptcy using support vector machine, artificial neural network, logistic regression, and multivariate discriminant analysis (Kim, 2011)	Çoklu Diskriminant Analizi Lojistik Regresyon Yapay Sinir Ağları DVM-Doğrusal DVM-Polinomal DVM-Sigmoid DVM-Radyal Tabanlı	Çalışmada kullanılan veri seti, 33 adedi iflas etmiş olan 66 işletmeye ait gözlemi içermektedir. Veri seti 1995-2002 yılları arasında kapsamaktadır. Veri seti anlamlılığı test edilmiş olan, likidite, stabilite, karlılık ve faaliyet kategorilerinin altında bulunan mali oranlardan oluşmaktadır.
7		DVM-Doğrusal	

	Predicting business failure using support vector machines with straightforward wrapper: A re-sampling study (Li & Sun, 2011)	DVM-Polinomal DVM-Gaussian DVM-Radyal Tabanlı	Çalışmada veri ön işleme aşaması sonrasında 153 adet gözlem bulunan bir veri seti kullanılmıştır. Veri ön işleme öncesi karlılık, faaliyet, yükümlülük yapı, büyüme ve hisse başına varlık ve getiri kategorileri altında organize edilmiş olan 30 adet mali oran seçilmiştir. Çalışmada birden fazla öznitelik seçim yöntemi kullanılmış ve performansları karşılaştırılmıştır.
8	A Multicriterion Decision Support Methodology For Audit Opinions: The Case Of Audit Reports Of Distressed Firms In Turkey (Adiloglu & Vuran, 2011)	Lojistik Regresyon	Çalışmada kullanılan veri setinde Borsa İstanbul AŞ.' De işlem gören ve 1998-2006 yılları arasında mali sıkıntı yaşamış 33 imalar işletmesine yer verilmiştir. Çalışmada gözlemleri temsil eden 10 adet öznitelik bulunmaktadır.
9	Comparative analysis of data mining methods for bankruptcy prediction (Olson vd., 2012)	Yapay Sinir Ağları Karar Ağaçları-ID3 Karar Ağaçları-C4.5 Karar Ağaçları-CART Karar Ağaçları-BFD Karar Ağaçları-AD Lojistik Regresyon Destek Vektör Makineleri	Çalışmada kullanılan veri seti toplam 1321 adet gözlemden oluşmaktadır. Gözlemlerden 697 adedi iflas etmiş olan işletmelerdir. Çalışmada kullanılan veri seti toplam 17 öznitelige yer verilmiştir. Kullanılan öznitelikler içerisinde "fiyat" dışındakiler işletmelerin mali tablolarından elde edilen verilerdir.
10	Partial Least Square Discriminant Analysis for bankruptcy prediction (Serrano-Cinca & Gutiérrez-Nieto, 2013)	Doğrusal Diskriminant Analizi Lojistik Regresyon Stepwise Lojistik Regresyon Çok Katmanlı Algılayıcılar K-En Yakın Komşu Algoritması Bayes Ağları Destek Vektör Makineleri Karar Ağaçları - C4.5(Boosting) Rastgele Orman Kısmi En Küçük Kareler Diskriminant Analizi	Çalışmada kullanılan veri setinde toplam 8293 gözlem bulunmaktadır. Bunların toplam 320 adedi 2009 krizi sonrasında 2011 yılına kadar iflas etmiş bankalardır. Kalan gözlemler aynı dönemde faaliyetlerine devam etmiş olan bankalardır. Çalışmada 17 adet mali oran öznitelik olarak kullanılmıştır.
11	Predicting restaurant financial distress using decision tree and AdaBoosted decision tree models (Kim & Upneja, 2014)	Karar Ağaçları Adaptive Boosting	Çalışmada kullanılan veri setinde toplam 163 işletmeye ait 826 gözlem yılı yer almaktadır. Gözlemlerin içerisinde 21 işletmeye ait 42 gözlem yılı mali sıkıntılı olarak etiketlenmiştir. Çalışmada 25 adet mali oran öznitelik olarak kullanılmıştır.
12	Going-concern prediction using hybrid random forests and rough set approach (Yeh vd., 2014)	Kaba Küme Teorisi Karar Ağaçları Rastgele Orman Algoritması Yapay Sinir Ağları Destek Vektör Makineleri	Çalışmada kullanılan veri seti 220 süreklilik görüşü ve 165 olumlu denetim görüşü içermektedir. Süreklilik görüşüne sahip gözlemler 55 farklı işletmeye aittir. Çalışmada kullanılan veri setinde toplam 27 adet öznitelik bulunmaktadır. Bunların 18 adedi mali oran, 2 adedi finansal olmayan oran ve 7 adedi ise entelektüel sermaye ile ilgili oranlardır.
13	Incorporating sequential information in bankruptcy prediction with predictors based on Markov for	Lojistik Regresyon Doğrusal Diskriminant Analizi Destek Vektör Makineleri	Çalışmada Belfirst veri tabanından faydalanılmıştır. Çalışmada toplam karlılık, likidite ve ödeme gücü kategorilerinin altında toplam 18 adet mali oran ile işletmelerin iş gücü profilini ortaya koyan 7 adet oran öznitelik olarak kullanılmıştır.

	discrimination (Volkov vd., 2017)	Rastgele Orman Algoritması	
		Extreme Gradient Boosting	
14	An investigation of bankruptcy prediction in imbalanced datasets (Veganzones & Séverin, 2018)	Doğrusal Diskriminant Analizi	Çalışmada toplam 1500 işletmeden oluşan dengeli bir veri seti kullanılmıştır. Çalışmada farklı sektörler için ve veri setinin tamamı için öznelik seçimi yapılmıştır. Kullanılan öznelikler çalıştırılan filtreleme yöntemlerinin en az ikisi tarafından seçilmiştir.
		Lojistik Regresyon	
		Yapay Sinir Ağları	
		Destek Vektör Makineleri	
		Rastgele Orman Algoritması	
15	Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting (J. Sun vd., 2020)	Destek Vektör Makineleri	Çalışmada toplam 2628 işletmeden oluşan bir veri seti kullanılmıştır. Veri setindeki 438 işletme mali sıkıntılı olarak etiketlenmiştir. Dengesiz veri setinin ortaya çıkaracağı sorunun üstesinden gelmek maksadı ile SMOTE metodundan yararlanılmıştır. Çalışmada ilk olarak karlılık, faaliyet kapasitesi, ödeme gücü, yapısal oranlar, büyüme yeteneği, risk seviyesi, hisse başına oranlar ve ortak karlılığı kategorileri altında organize edilmiş olan 55 mali oran belirlenmiştir. Öznelik seçimi sonrasında 19 mali oran belirlenmiş ve öznelik olarak kullanılmıştır.
16	Combining corporate governance indicators with stacking ensembles for financial distress prediction (Liang vd., 2020)	Destek Vektör Makineleri	Çalışmada kullanılan veri setinde toplam 286 işletme yer almaktadır. İşletmelerin 143 adedi iflas etmiştir. Çalışmada iki mali oranlar ve kurumsal yönetim göstergeleri olmak üzere iki kategoride toplam 61 öznelik kullanılmıştır. Eğitilen ilk modelde öznelik seçimi yapılmamıştır. İkinci modelde aşamalı diskriminant analizi ile seçilen öznelikler kullanılmıştır.

3.4. Risk Kontrolü ve Değerleme

Yapay zeka ve makine öğrenmesi metodlarından iç denetçilerin faydalanabileceği modeller geliştirmek ve iç denetim sürecinin etkinliğini ölçmek maksadı ile de faydalanılmaktadır (Tang vd., 2017; Werner & Gehrke, 2015, 2019). Ayrıca risk kontrolü ve değerlendirme amacı ile de pek çok model geliştirilmiştir.

Tablo 4. Risk Kontrolü ve Değerleme ile İlgili Literatürdeki Çalışmalar

	Çalışmanın Başlığı	Kullanılan Metotlar	Veri Setinin Yapısı ve Öznelikler
1	Earnings management prediction: A pilot study of combining neural networks and decision trees (Tsai & Chiou, 2009)	Yapay Sinir Ağları	Kazanç yönetimi uygulamalarının tespitini amaçlayan bu çalışmada 2002-2005 yılları arasında Tayvan Borsasında işlem gören işletmelerin yer aldığı bir veri seti kullanılmıştır. Veri setinde yer alan işletmelerin tamamı elektronik sektörde faaliyet göstermektedir. Çalışmada yer alan 11 öznelik literatür taraması ile seçilmiştir.
		Karar Ağaçları	
2	Towards a Machine Learning Approach for Earnings Manipulation Detection (Dbouk & Zaarour, 2017)	Naive Bayes Sınıflandırıcısı	Çalışmada Lübnan' da sıvı yakıt sektöründe faaliyet göstermekte olan 53 işletmenin verilerinden faydalanılmıştır. Çalışmada kazanç manipülasyonunun tespiti maksadı ile 19 mali oran öznelik olarak kullanılmıştır.
3	Detecting financial restatements using data mining techniques (Dutta vd., 2017)	Yapay Sinir Ağları	Çalışmada kullanılan iki veri seti 2001-2014 yılları arasını aittir. Bu veri seti gözlemlerin filtrelemesi neticesinde elde edilmiştir. Veri seti toplam, veri ön işleme sonrasında toplam 3513 gözlem yığını içermektedir. Veri setin nihai olarak 109 hile beyanı içermektedir. Dengesiz veri seti sorununun üstesinden gelmek maksadı ile SMOTE metodu kullanılmıştır. Çalışmada literatür temel alınarak ilk olarak 116 farklı öznelik belirlenmiştir. Bunların içerisinde 14 adedi boyut indirgeme sürecinde silinmiştir.
		Karar Ağaçları	
		Bayes Ağları	
		Naive Bayes Sınıflandırıcısı	
		Destek Vektör Makineleri	

4	Research on Application of Data Mining Technology in Risk Assessment Process of Audit (Yan vd., 2019)	DBSCAN Kümeleme (density-based spatial clustering of applications with noise)	Çalışma teorik bir model önermektedir.
---	---	---	--

3.5. Hile Denetimi

Hileli finansal raporlamanın tespiti, yapılan hileli işlemin gelecekteki olumsuz etkilerinin önlenmesi açısından çok önemlidir. Hileli finansal raporlama tespiti, hileli verileri gerçek verilerden ayırt etmeyi ve böylece hileli davranış ve/veya hileli faaliyeti ifşa ederek, karar vericilerin hilenin olumsuz etkisine karşı uygun stratejileri geliştirmesini amaçlamaktadır. Veri analitiği ve makine öğrenmesi metotları bu maksat ile çok yaygın bir şekilde kullanılmaktadır.

Tablo 5. Hile Denetimi ile İlgili Literatürdeki Çalışmalar

	Çalışmanın Başlığı	Kullanılan Metotlar	Veri Setinin Yapısı ve Öznitelikler
1	Forecasting Fraudulent Financial Statements using Data Mining (Kotsiantis vd., 2006)	Karar Ağaçları – C4.5 Yapay Sinir Ağları Bayes Ağları K-En Yakın Komşu Algoritması Destek Vektör Makineleri	Çalışmada kullanılan veri seti Atina Borsasında işlem gören 164 Yunan imalat işletmesine aittir. Veri setinde bulunan işletmelerin 41 adedi hakkında hile kanıtları sunulmuştur. Çalışmada kullanılan karlılık, kaldıraç, likidite, verimlilik ve nakit akışı kategorileri altında toplam 23 adet öznitelik kullanılmıştır.
2	Data Mining techniques for the detection of fraudulent financial statements (Kirkos, Spathis, & Manolopoulos, 2007b)	Karar Ağaçları Yapay Sinir Ağları Bayes Ağları	Çalışmada kullanılan veri seti 76 Yunan imalat firmasına ait gözlemleri içermektedir. Veri setindeki işletmelerin 38 adedi hakkında hile kanıtları sunulmuştur. İlk olarak 27 mali oran öznitelik olarak belirlenmiştir. ANOVA testi sonucunda anlamlılığı tespit edilen 10 mali oran son olarak öznitelik olarak veri setinde yer almıştır.
3	False Financial Statements: Characteristics of China's Listed Companies and CART Detecting Approach (Bai vd., 2008)	Karar Ağaçları – CART Lojistik Regresyon	Çalışmada kullanılan veri seti toplam 148 işletmeye ait gözlemden oluşmaktadır. Gözlemlerin 24 adedi hatalı finansal raporlardır. Hileli raporlama yapan işletmelerin hayali işlemler yaratmak sureti ile gelirlerini arttırması sebebi ile ilk olarak işletmelerin varlık ve gelirleri ile ilgili mali oranlar öznitelik olarak belirlenmiştir. Bu işlemlere paralel olarak işletmelerin başka kalemler üzerinde yapabileceği tahrifleri temsil etmesi için 17 mali oran da öznitelik olarak seçilmiştir.
4	Classification techniques for the identification of falsified financial statements: a comparative analysis (Gaganis, 2009)	Yapay Sinir Ağları Çoklu Diskriminant Analizi Lojistik Regresyon UTADIS MHDIS DVM- Doğrusal DVM-Polinomal	Çalışmada kullanılan veri setinde yer alan işletmeler Atina Borsasında işlem görmektedir. Veri seti 2001-2004 yıllarını içine almaktadır. İlk olarak 199 hileli raporlama yapan işletme tespit edilmiştir. İkinci adımda seçilen ilk işletmeler aynı sektörden ve aynı büyüklüğe sahip hileli raporlama yapmamış olan işletmeler ile eşleştirilmek sureti ile veri seti oluşturulmuştur. Literatüre dayalı olarak seçilen 28 mali oran, karlılık, performans, ödeme kabiliyeti gibi boyutlarda temsil kabiliyeti yeteneğine sahiptir.

		DVM-Radyal Tabanlı Fonksiyon	
		K-En Yakın Komşu Algoritması	
		Olasılıksal Yapay Sinir Ağları	
5	Detecting Management Fraud in Public Companies (Cecchini vd., 2010a)	Destek Vektör Makineleri	Çalışmada yer alan veri seti 122 işletmeye ait 205 hileli raporlama gözlem yılını içermektedir. Bu gözlemler ile birlikte veri setinde toplam 6427 hile içermeye gözlem yer almaktadır. Çalışmada literatürde kullanılmış olan 40 öznelikten faydalanılmıştır.
6	MetaFraud: A Meta-Learning Framework for Detecting Financial Fraud (Abbasi vd., 2012)	Lojistik Regresyon	Çalışmada kullanılan eğitim veri seti toplam 3862 gözlem içermektedir. Bunları 406 adedi hile barındırmaktadır. Test veri seti ise 5144 gözlem içermektedir. Test veri seti içerisindeki 409 gözlem hile barındırmaktadır. Çalışmada toplam 12 adet mali oran öznelik olarak kullanılmıştır. Seçilen mali oranlar literatür destekli olarak belirlenmiştir.
		Yapay Sinir Ağları	
		DVM-Doğrusal	
		DVM-Polinomal	
		DVM-Radyal Tabanlı Fonksiyon	
		Navie Bayes Sınıflandırıcısı	
		Bayes Ağları	
		Karar Ağaçları-J48	
		Navie Bayes Ağaçları	
		ADTree	
		Rastgele Orman Algoritması	
		REPTree	
		En Yakın Komşu Algoritması	
JRip			
7	The Application of Data Mining Techniques in Manipulated Financial Statement Classification: The Case of Turkey (Ozdoglu vd., 2016)	Karar Ağaçları	Çalışmada kullanılan veri seti 2009 – 2013 taranarak oluşturulmuştur. Veri setinde yer alan işletmeler ilgili periyotta Borsa İstanbul AŞ. 'de işlem görmektedir. Veri seti toplam olarak 224 gözlem içermektedir. Bunları 110 adedi riskli olarak etiketlenmiştir. Çalışmada literatüre dayalı olarak seçilen 13 adet mali oran öznelik olarak kullanılmıştır.
		Lojistik Regresyon	
		Yapay Sinir Ağları	
8	Detection of fraudulent financial statements using the hybrid data mining approach (Chen, 2016)	Karar Ağaçları-CART	Çalışmada kullanılan veri seti toplam olarak 176 işletmeye ait gözlemi içermektedir. Veri setindeki 44 gözlem hile içermektedir. Çalışmada 23 adet mali oran ile 7 adet mali olmayan oran öznelik olarak yer almaktadır. Çalışmada öznelik seçim yöntemi olarak CHAIN ve CART karar ağacı algoritmaları kullanılmıştır.
		Karar Ağaçları-CHAID	
		Bayes Ağları	
		Destek Vektör Makineleri	
		Yapay Sinir Ağları	
		Adaboost	
		Rastgele Orman Algoritması	
		Destek Vektör Makineleri	
		Doğrusal Probit Model	
		Decision Stump Algoritması	
		Karar Ağaçları-J48	
		Bayes Ağları	
		Navie Bayes Sınıflayıcıları	
Yapay Sinir Ağları			

3.6. Muhasebe Denetiminde Metin Madenciliği ve Doğal Dil İşleme

Muhasebe bilgi sisteminin içerisinde üretilen verinin büyük bir çoğunluğu yapısal formata sahip değildir. (N. C. Brown vd., 2020). Muhasebe ve denetim alanında, veri analitiği ve makine öğrenmesi metotlarının kullanıldığı çalışmalarda çoğunlukla, yapısal formattaki finansal nitelikteki verilere yer verilmiştir. Ancak son yıllarda yapılan çalışmalar, modellerde sadece finansal nitelikteki verilerin kullanılmasının modelin başarısını sınırladığını belirtmektedir. Finansal nitelikteki özniteliklerin işletme ile ilgili aktardığı bilgi sınırlıdır. Finansal nitelikteki veri kümeleri tek başına işletmenin iç ve dış çevresi ile ilgili tüm faktörler hakkında yeterli bilgiyi aktarabilecek niteliğe sahip değildir (Jo & Shin, 2016). Bu nedenle finansal nitelikte olmayan veri yapılarının modellerde kullanılabilmesi belirtilmektedir (Cecchini vd., 2010b; Hájek & Olej, 2015; Loughran & Mcdonald, 2011). Yapısal formatta olmayan metinlerin analizi nitel tahmine dayalı bilgileri elde etmenin etkili bir yoludur. Metin analizi, nitel analizin bir alt kümesi olarak ifade edilmektedir (Loughran & Mcdonald, 2016).

Son yıllarda, metin analizinin popülerliği kurumsal raporların içeriğinin analizinde giderek artmaktadır. Veri analitiği ve metin madenciliği metotları muhasebe ve denetim alanında on yıldan fazla süredir kullanılmaktadır (Baldwin vd., 2006; Fisher, 2018; Loughran & Mcdonald, 2016). Metin analizi metotlarının kullanımı muhasebe ve denetim alanında oldukça yeni olsa da yapılan çalışmalar, bu metotlardan faydalanılan yayınların sayısının giderek arttığını belirtmektedir (Fisher vd., 2016). İşletmelerin yayınladığı farklı kurumsal raporlardan elde edilecek olan anahtar kelimelerden veya söz öbeklerinden model başarısını arttırmak maksadı ile yararlanılabilir (Qiu vd., 2013). Bu bağlamda, denetim alanında literatürde farklı çalışmalar bulunmaktadır.

Tablo 6. Metin Madenciliği ve Doğal Dil İşleme Metotlarının Kullanıldığı Çalışmalar

	Çalışmanın Adı	Çalışmanın İçeriği
1	The language of quarterly reports as an indicator of change in the company's financial status (Magnusson vd., 2005)	Çalışmada bilişim sektöründeki ekonomik büyümenin ciddi bir düşüş yaşadığı 2000 ile 2001 yılları arasında odaklanılmaktadır. Bu bağlamda çalışmada, bilişim sektöründeki gerilemeden ciddi şekilde etkilenen telekomünikasyon sektöründe faaliyet gösteren işletmelere ait mali raporlar ile analizler yürütülmüştür. Çalışmanın temel amacı bilişim sektöründe yaşanan gelişmelerin işletmelerin üçer aylık mali raporlarının diline nasıl yansıdığı ve metinlerde belirtilen değişikliklerin mali rakamlardaki değişikliklerden önce gelip gelmediğini incelemektir.
2	Can Linguistic Predictors Detect Fraudulent Financial Filings? (Goel vd., 2010)	Çalışmada hile tespiti maksadı ile işletmelerin yıllık mali raporlarının metinsel içeriğinin dilsel analizini içeren bir metodoloji sunulmaktadır. Metinlerin içerdiği dilsel ip uçları ya da dil özellikleri, çalışmada araştırmacıların metinle etkileşiminin ilerisine geçmiş ve raporların içeriğinin ötesini incelemelerine de olanak tanımıştır. Çalışma sonucunda yıllık mali raporların niteliksel anlatı içeriğinin, mali rakamlar tarafından doğru bir şekilde tespiti mümkün olmayan hileleri tespit etmek için faydalı bilgiler içerdiği ortaya koyulmuştur. Çalışma sonucunda hileli finansal raporların yazım tarzlarında farklılıklar tespit edilmiştir.
3	Making words work: Using financial text as a predictor of financial events (Cecchini vd., 2010b)	Çalışmada mali nitelikteki metinleri otomatik olarak analiz etmek maksadı ile bir metodoloji geliştirilmiştir. İşletmeye ciddi zararları olan finansal olayların tespiti maksadı ile geliştirilen bu metodoloji hile ve mali sıkıntı üzerinde test edilmiştir. Araştırmacılar belirli sözcüklerin belirli sonuçlar ile olumlu ya da olumsuz ilişki içerisinde olduğu varsayımına sahiptir. Bu bağlamda mali metinlerden bir olay ile ilgili anahtar terimler sözlüğü tanımlamayı amaçlayan araştırmacılar tarafından kullanılabilir bir metodoloji geliştirilmiştir.
4	Identification of fraudulent financial statements using linguistic credibility analysis (Humpherys vd., 2011)	Çalışmanın temel amacı hile içeren olaylar ile hile içermeyen olaylar arasındaki, kullanılan dil hakkındaki anlayışın derinleştirilmesidir. Bu bağlamda doğal dil işlemeden kurumsal raporları düzenleyen şahısların niyetini ortaya çıkarmak maksadı ile faydalanılmıştır. Çalışmanın sonuçları, yararlanılan modellerin mali raporlarda hilenin tespiti maksadı ile kullanılabilmesini ortaya koymuştur.
5	When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks (Loughran & Mcdonald, 2011)	Muhasebe ve finans literatüründe mali raporların sahip olduğu dil tonunun analizi maksadı ile metin analizinin kullanılması giderek yaygınlaşmaktadır. Günümüze kadar literatürde elde edilmiş olan sonuçlar, olumsuz kelimelerin sınıflandırılmasının metinlerin sahip olduğu tonun ölçülmesinde etkili olabileceğini göstermektedir. Çalışmanın metin analizi bağlamında literatüre çeşitli katkıları bulunmaktadır. Bunların içerisinde en önemlisi, Harvard listesinde bulunan negatif kelime sayımlarının yaklaşık olarak dörtte üçünün mali anlamda genellikle negatif olmayan kelimelere atfedilebileceği ortaya koyulmaktadır. Çalışmada genellikle mali anlamda olumsuz etkileri bulunan 2337 kelimelik bir liste oluşturulmuştur.
6	Extracting Key Phrases as Predictors of Corporate Bankruptcy: Empirical Analysis of Annual Reports	İşletmelerin yıllık mali raporları üzerinden iflas tahmini yapmak oldukça zordur. Bu sorunun üstesinden gelmek maksadı ile çalışmada Japonya' da kamuya sunulan raporlardan bilgi çıkarımına odaklanılmıştır. Çalışmada 90 adedi iflas etmiş toplam 180 adet işletmenin yer aldığı bir veri seti kullanılmıştır. İflas etme durumunu ayırtıran anahtar ifadelerin ayrıştırılması maksadı ile kamu ile paylaşılmış yıllık raporlardan faydalanılmıştır. Metin

	by Text Mining (Shirata vd., 2011)	madenciligi temeli üzerine inşa edilmiş olan çalışma sonucunda, faaliyet raporlarında "dividends" veya "retained earnings" ifadelerinin belirli bazı özel ifadeler ile birlikte yer alması durumunun iflas durumunun tespitinde etkili olabileceği tespit edilmiştir. İflas etmemiş olan işletmeler için "capital investment", "new business" ve "research and development" ifadelerinin "dividend" ve "retained earnings" ifadelerinin aynı cümlelerde bulunduğu tespit edilmiştir.
7	Beyond the Numbers: Mining The Annual Reports For Hidden Cues Indicative of Financial Statement Fraud (Goel & Gangolly, 2012)	Çalışmada hile tespiti maksadı ile işletmeler tarafından yayınlanmış olan yıllık raporlardaki metin içeriği incelenmektedir. Bu yolla hileli raporlama yapan ve yapmayan işletmelerin arasındaki yazım ve sunum stilineki farklılıkların tespiti amaçlanmaktadır. Yıllık raporlarda yapılan hilenin sayısal mali bilgiler üzerinde gizlenebileceği varsayımında hareketle çalışmada oluşturulan hipotezler ki-kare yöntemi ile test edilmiştir. Çalışmada sonuçları hileli raporlamada özellikle altı ipucu kategorisine ışık tutmaktadır. Bunlar karmaşık cümle yapısı kullanımı, okunabilirlik indeksi ile ölçüldüğünde okuma ve anlama güçlüğü, pozitif ton kullanımı, pasif ses kullanımı, belirsizlik belirtecilerinin kullanımı ve zarflarını kullanımıdır. Raporlar üzerinde yapılan analizlerde hileli raporlar ile hileli olmayan raporlar arasında büyük harf ve noktalama işareti kullanımı arasında bir farklılık tespit edilmemiştir.
8	Forecasting Corporate Financial Performance Using Sentiment in Annual Reports For Stakeholders' Decision-Making (Hajek vd., 2014)	Çalışmada yıllık mali raporların üslubunun* mali performansın tahminindeki rolü incelenmesi amaçlanmaktadır. Çalışmada ABD' de faaliyet gösteren işletmelere ait veriler kullanılmaktadır. Çalışmada destek vektör makineleri en iyi sınıflama sonucu vermiştir.
9	Word Categorization of Corporate Annual Reports for Bankruptcy Prediction by Machine Learning Methods (Hájek & Olej, 2015)	Çalışmada yıllık mali raporların çeşitli kelime kategorilerinin tespit edilmesi ve bunların iflas tahmini üzerindeki etkisinin incelenmesi amaçlanmaktadır. Çalışmada sonuçları, iflas eden işletmelerin raporlarında kullandıkları dilin daha güçlü azim,, başarı, aşinalık, ilgi, dışlama ve inkar ile karakterize edildiği tespit edilmiştir. Ayrıca iflas etmiş olan işletmeleri daha fazla olumlu, kesinlik taşımayan, olumsuz dil tercih ettiği ve kip kullandığı da tespit edilmiştir. Bir başka ifade ile iflas eden işletmeler paydaşlarının endişelerini gidermek maksadı ile hem olumlu hem de olumsuz üslup özelliklerini daha fazla kullandığı tespit edilmiştir. Ayrıca çalışma sonuçları iflas eden işletmelerin raporların kendinden emin ve hedef odaklı görünmeye çalıştıklarını ortaya koymaktadır.
10	Bankruptcy Prediction Modeling Using Qualitative Information Based on Big Data Analytics (Jo & Shin, 2016)	Çalışmada mali oranlar ile birlikte ekonomi haberlerinden elde edilen nitel bilginin birlikte iflas tahmini maksadı ile kullanılması amaçlanmaktadır. Çalışmada ekonomi haberlerinden duygu çıkarımı için alana özgü bir duygu sözlüğü kullanılmıştır. Çalışma sonucunda büyük veri analitiğine dayalı elde edilen nitel verilerin, muhasebe verilerine dayalı iflas tahmin modellerinden daha başarılı sonuçlar ortaya koyduğu tespit edilmiştir.
11	Do Sentiments Matter in Fraud Detection? Estimating Semantic Orientation of Annual Reports: Do Sentiments Matter in Fraud Detection: Estimating Semantic Orientation of Annual Reports (Goel & Uzun, 2016)	Çalışmada yıllık mali raporların metin içeriğinin analizi maksadı ile yeni bir yaklaşım sunulmaktadır. Doğal dil işleme teknikleri kullanılarak raporlardaki "Management Discussion & Analysis" bölümleri üslup açısından değerlendirilmiştir. Çalışma sonucunda hile riskinin daha yüksek olduğu işletmelerde negatif ve pozitif bir üslubun daha belirgin kullanıldığı, nesnel ifadeleri kıyasla öznel ifadelerin daha fazla kullanıldığı ve duygu yoğunluğu sergileyen ifadelerin daha fazla kullanıldığı tespit edilmiştir.
12	Can Language Predict Bankruptcy? The Explanatory Power of Tone in 10-K Filings (Lopatta vd., 2017)	Bu çalışmada işletmelerin yıllık raporlarında kullanılan dilin iflas riskini yansıtıp yansıtmadığı incelenmektedir. Bu amaçla çalışmada 1994-2015 yılları arasında iflas etmiş olan, ABD faaliyet göstermiş 424 işletmeye 10-k raporları analiz edilmektedir. Çalışmada bir logit model kullanılmaktadır. Logit model sonuçları iflas riski taşıyan işletmelerin 10-k raporlarında daha fazla olumsuz kelime kullandığını göstermektedir. Bu ilişkinin, işletmenin iflas başvurusundan üç yıl öncesine kadar sürdüğü tespit edilmiştir. Çalışma sonuçları genel vekalet teorisine dayanan ve işletme yönetiminin paydaşlardan işletmenin gerçek durumunu gizlemek yönünde hareket edeceği varsayımına sahip olan çalışmalardan farklılaşmaktadır.
13	A new random subspace method incorporating sentiment and textual information for financial distress prediction (Wang vd., 2018)	Çalışmada mali sıkıntı tahmini maksadı ile duygu analizi ve metinsel bilgileri kullanan Random Subspace metodunun dahil edildiği yeni bir yaklaşım önerilmektedir. Çalışma Şenzen ve Şangay borsalarında işlem gören işletmelerin yer aldığı ve 129 adedi mali sıkıntılı ("Special Treatment" etiketi almış) olarak etiketlenmiş toplam 1726 işletmenin gözlemlerini içine alan bir veri seti kullanılmıştır. Dengesiz bir veri setinin kullanıldığı çalışmada farklı temel öğreniciler, topluluk öğrenmesi metotları karşılaştırılmış olup aynı zamanda under-sampling, over-sampling ve SMOTE metotları da dengesiz veri setinin ortaya çıkarabileceği

		sorunların üstesinden gelmek için kullanılmıştır. Çalışmada mali nitelikli özneliklere, metinlerden ve duygu analizinden elde edilen bilgiler de entegre edilme amaçlanmaktadır.
14	Detecting manager's fraud risk using text analysis: evidence from Iran (Rahrovi Dastjerdi vd., 2019)	Çalışmada işletmelerdeki hile riskini tespit etmek maksadı ile metin madenciliğini temel alan bir yaklaşım sunulmaktadır. Bu amaçla, işletmelerin raporları metin analizine tabi tutulmuş, CVX metodu ve LASSO regresyonu kullanılarak raporlardaki hile riskini açıklamada güçlü olan kelimeler tespit edilmiştir. Çalışma sonucunda metin temelli olarak hile riskini açıklamada kullanılacak bir model sunulmaktadır.

*" Sentiment" kelimesi üslup olarak çevrilmiştir.

3.7. Yapay Zekâ ve Makine Öğrenmesi Tekniklerinin Uygulamaya Yansımaları

Denetim firmalarının, sürece yapay zekayı nasıl entegre ettiğini inceleyen pek çok araştırma bulunmaktadır. Ayrıca denetim firmaları tarafından yayınlanan raporlarda da veri analitiği ile ilgili yapılan yatırımlar hakkında bilgi mevcuttur. Dört büyük denetim firmasında birisi olan KPMG, "IBM Watson" yapay zekasını, zeki denetim araçları geliştirmek maksadı ile kullanmaktadır (Melendez, 2016). KPMG, yapısal olmayan veri yapıları, sosyal, dijital ve basılı medyadan elde edilen, veri yığınlarının işlenmesi amacı ile kullanılacak bilişsel sistemlerin geliştirilmesi amacı yatırım yapmaktadır (KPMG, 2018). KPMG tarafından yayınlanan "Audit 2025: The Future is Now" raporuna göre, Birleşik Devletlerdeki büyük işletmelerin risk yönetiminin etkinliğinin arttıracak unsurları talep edilmektedir (KPMG, 2017).

PriceWaterhouseCooper "PwC", Halo isimli bir analitik platform geliştirmiştir. Geliştirilen "Halo" isimli uygulama büyük boyutlu verileri okuyarak, denetçiye geçerli ve güvenilir veri analizi ve risk değerlendirme yapma imkanı sağlamaktadır (PwC, 2016). Ayrıca uygulama muhasebe işlemlerini tarayarak yüksek riskli işlemlerin ya da şüpheli işlemlerin tespiti hususunda da denetçiye destek sağlamaktadır (PwC, 2016). PwC, muhtemel robotik otomasyon uygulamalarını araştırmakta ve yapay zekâ temelli anomali tespiti ve yapay zekâ temelli tahminler üzerine geliştirme çalışmaları yürütmektedir. PwC ayrıca denetim kalitesini arttırmak maksadı ile blockchain teknolojisini de kullanmaktadır (PwC, 2019).

Deloitte, yapay zeka geliştirmeleri için Argus' u ve veri analitiği için ise OtiX' i geliştirmiştir (Deloitte, 2018). Ayrıca Deloitte de sürekli denetim uygulamalarında etkinlik artışı sağlamak amacı ile blockchain teknolojisi de kullanmaktadır.

Ernst & Young "EY", Microsoft Azure yardımı ile kendi yapay zeka platformunu geliştirmiş ve süreçte çeşitli iyileştirmeler gerçekleştirmek maksadı ile yeni teknolojilere yatırım yapmayı sürdürmektedir (Makrygiannis, 2019). EY, yapısal olmayan veri yapılarının analizi maksadı ile derin öğrenme temelli uygulamalara yatırım yapmaktadır. Yapısal olmayan verinin analizi, denetim sürecindeki elle yapılan işleri azaltacaktır. Dahası, EY işlemleri gerçek zamanlı ve otomatik olarak yürütmek maksadı ile blockchain teknolojisini incelemektedir. Böylece kayıtlardaki çarpıklıklar günlük olarak tespit edilebilecek ve denetimin etkinliği arttırılacaktır.

4. Sonuç

Denetim sürecinde, denetçinin mesleki yargısını kullanması hem denetim teorisi açısından hem de yasal olarak bir zorunluluktur. Ancak denetim sürecinde teknoloji kullanımının mesleki yargı kullanımı ile çeliştiği algısı nedeni ile süreçte karar destek modellerinin kullanımına bir direnç bulunduğu söylenebilir (Meredith vd., 2020). Bu çalışmada kapsamlı bir literatür taraması vasıtası ile yapay zekâ ve makine öğrenmesi metodlarının muhasebe denetiminde hangi amaçlar ile kullanıldığına ışık tutmak amaçlanmaktadır. Bu bağlamda literatür taraması sonuçları altı başlık altında organize edilmiş ve sunulmuştur. Ayrıca yapay zekâ ve makine öğrenmesi metodlarının uygulamadaki yansımalarının tespiti maksadı ile dört büyük denetim firması tarafından yayınlanan yıllık raporlar incelenmiş, bu alana yapılan yatırımlar ile metodların hangi amaçlar ile kullanıldığı raporlanmıştır.

Denetçiler, denetim süreci içerisinde de süreci sonuçlandırırken de sezgisel olarak karar vermekte ve geçmiş deneyimlerinde yola çıkarak sonuca varmaktadır. Ayrıca mesleki tecrübe denetim sürecinde bir yasal zorunluluktur. Sezgisel olarak karar veren denetçilerin, onlara analiz imkânı sağlamayan karar destek modelleri ile desteklenmesi, modelden beklenen başarının düşmesine neden olabilir. Denetim sürecinde, denetçiden görülenin ötesinin görmesi beklenmektedir. Bu nedenle modellerin mutlaka denetçinin mesleki yargısının ve mesleki şüpheciliğinin önüne geçmemesi ya da denetçinin yeni uygulama hakkında mesleki yargısının ve mesleki şüpheciliğinin önüne geçtiği yönünde bir düşüncesinin olmaması gerekmektedir.

Denetçiler süreçte kullanılan veri ve veri kaynaklarının kalitesi konusunda şüpheli davranmaktadır. Denetim sürecine yeni bir araç adaptasyonunda, entegre edilecek olan bu yeni aracın, denetçinin iş kalitesine bir tehdit oluşturmaması gerekmektedir. Denetim görüşü ile sonuçlanan denetim süreci temel olarak sezgisel bir süreçtir. Denetçi, denetim

görüşüne giden süreçte mesleki tecrübesine dayanarak ilerlemektedir. Denetçinin kararını destekleyecek olan yeni aracın da denetçinin karar mekanizmasına uyum göstermesi gerekmektedir. Bir diğer taraftan denetçinin, geleneksel olarak süreç içerisindeki operasyonel ve taktiksel düzeydeki işlemlerin neticesinde stratejik bir karar aldığı ve bu kararı denetim raporunda ilgili taraflara iletmiş olduğu söylenebilir. Bu açıdan denetçinin kararını destekleyecek olan veri analitiği aracının stratejik düzeyde olması denetçinin aracı benimsemesi noktasında olumlu etki yaratabileceği söylenebilir.

Kaynakça

- Abbasi, Albrecht, Vance, & Hansen. (2012). MetaFraud: A Meta-Learning Framework for Detecting Financial Fraud. *MIS Quarterly*, 36(4), 1293. <https://doi.org/10.2307/41703508>
- Abbott, D. (2014). *Applied predictive analytics: Principles and techniques for the professional data analyst*. Wiley.
- Adiloglu, B., & Vuran, B. (2011). A Multicriterion Decision Support Methodology For Audit Opinions: The Case Of Audit Reports Of Distressed Firms In Turkey. *International Business & Economics Research Journal (IBER)*, 10(12), 37. <https://doi.org/10.19030/iber.v10i12.6648>
- Amani, F. A., & Fadlalla, A. M. (2017). Data mining applications in accounting: A review of the literature and organizing framework. *International Journal of Accounting Information Systems*, 24, 32-58. <https://doi.org/10.1016/j.accinf.2016.12.004>
- Bai, B., Yen, J., & Yang, X. (2008). False Financial Statements: Characteristics of China's Listed Companies and CART Detecting Approach. *International Journal of Information Technology & Decision Making*, 07(02), 339-359. <https://doi.org/10.1142/S0219622008002958>
- Baldwin, A. A., Brown, C. E., & Trinkle, B. S. (2006). Opportunities for artificial intelligence development in the accounting domain: The case for auditing. *Intelligent Systems in Accounting, Finance and Management*, 14(3), 77-86. <https://doi.org/10.1002/isaf.277>
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Bozkurt, N. (2018). *Muhasebe Denetimi*. ALFA Basım Yayım Dağıtım San. ve Tic. Ltd Şti.
- Brown, C. E., Wong, J. A., & Baldwin, A. A. (2007). A Review and Analysis of the Existing Research Streams in Continuous Auditing. *Journal of Emerging Technologies in Accounting*, 4(1), 1-28. <https://doi.org/10.2308/jeta.2007.4.1.1>
- Brown, N. C., Crowley, R. M., & Elliott, W. B. (2020). What Are You Saying? Using topic to Detect Financial Misreporting. *Journal of Accounting Research*, 58(1), 237-291. <https://doi.org/10.1111/1475-679X.12294>
- Büyüktanır, T., & Toraman, T. (2020). *Bağımsız Denetim Görüşlerinin Makine Öğrenmesi Algoritmaları ile Tahmin Edilmesi*. 1-4. <https://doi.org/10.1109/SIU49456.2020.9302441>
- Campeato, O. (2021). *NLP fundamentals for developers*. Mercury Learning and Information.
- Cecchini, M., Aytug, H., Koehler, G. J., & Pathak, P. (2010a). Detecting Management Fraud in Public Companies. *Management Science*, 56(7), 1146-1160. <https://doi.org/10.1287/mnsc.1100.1174>
- Cecchini, M., Aytug, H., Koehler, G. J., & Pathak, P. (2010b). Making words work: Using financial text as a predictor of financial events. *Decision Support Systems*, 50(1), 164-175. <https://doi.org/10.1016/j.dss.2010.07.012>
- Chan, D. Y., & Vasarhelyi, M. A. (2011). Innovation and practice of continuous auditing. *International Journal of Accounting Information Systems*, 12(2), 152-160. <https://doi.org/10.1016/j.accinf.2011.01.001>
- Chen, S. (2016). Detection of fraudulent financial statements using the hybrid data mining approach. *SpringerPlus*, 5(1), 89. <https://doi.org/10.1186/s40064-016-1707-6>
- Chye Koh, H., & Kee Low, C. (2004). Going concern prediction using data mining techniques. *Managerial Auditing Journal*, 19(3), 462-476. <https://doi.org/10.1108/02686900410524436>
- Çelik, S. (2020). Metin Madenciliği ile Shakespeare Külliyyatının İncelenmesi. *MANAS Sosyal Araştırmalar Dergisi*, 1343-1357. <https://doi.org/10.33206/mjss.561919>
- Dai, J., & Vasarhelyi, M. A. (2016). Imagineering Audit 4.0. *Journal of Emerging Technologies in Accounting*, 13(1), 1-15. <https://doi.org/10.2308/jeta-10494>

- Dbouk, B., & Zaarour, I. (2017). Towards a Machine Learning Approach for Earnings Manipulation Detection. *Asian Journal of Business and Accounting*, 10(2), 215-251.
- Debreceeny, R. S. (2011). Betwixt and Between? Bringing Information Systems and Accounting Systems Research Together. *Journal of Information Systems*, 25(2), 1-9. <https://doi.org/10.2308/isys-10140>
- Deloitte. (2018). *16 Artificial Intelligence projects from Deloitte Practical cases of applied AI*. Deloitte. <https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/innovatie/deloitte-nl-innovatie-artificial-intelligence-16-practical-cases.pdf>
- Doumpos, M., Gaganis, C., & Pasiouras, F. (2005). Explaining qualifications in audit reports using a support vector machine methodology. *Intelligent Systems in Accounting, Finance and Management*, 13(4), 197-215. <https://doi.org/10.1002/isaf.268>
- Dutta, I., Dutta, S., & Raahemi, B. (2017). Detecting financial restatements using data mining techniques. *Expert Systems with Applications*, 90, 374-393. <https://doi.org/10.1016/j.eswa.2017.08.030>
- Elliott, R. K. (1992). The Third Wave Breaks on the Shores of Accounting. *Accounting Horizons*, 6(2), 26.
- Fisher, I. E. (2018). A Perspective on Textual Analysis in Accounting. *Journal of Emerging Technologies in Accounting*, 15(2), 11-13. <https://doi.org/10.2308/jeta-10639>
- Fisher, I. E., Garnsey, M. R., & Hughes, M. E. (2016). Natural Language Processing in Accounting, Auditing and Finance: A Synthesis of the Literature with a Roadmap for Future Research: NLP in Accounting, Auditing and Finance. *Intelligent Systems in Accounting, Finance and Management*, 23(3), 157-214. <https://doi.org/10.1002/isaf.1386>
- Frydman, H., Altman, E. I., & Kao, D.-L. (1985). Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress. *The Journal of Finance*, 40(1), 269-291. <https://doi.org/10.1111/j.1540-6261.1985.tb04949.x>
- Gaganis, C. (2009). Classification techniques for the identification of falsified financial statements: A comparative analysis. *Intelligent Systems in Accounting, Finance & Management*, 16(3), 207-229. <https://doi.org/10.1002/isaf.303>
- Gepp, A., Kumar, K., & Bhattacharya, S. (2010). Business failure prediction using decision trees. *Journal of Forecasting*, 29(6), 536-555. <https://doi.org/10.1002/for.1153>
- Goel, S., & Gangolly, J. (2012). Beyond the Numbers: Mining The Annual Reports For Hidden Cues Indicative of Financial Statement Fraud. *Intelligent Systems in Accounting, Finance and Management*, 19(2), 75-89. <https://doi.org/10.1002/isaf.1326>
- Goel, S., Gangolly, J., Faerman, S. R., & Uzuner, O. (2010). Can Linguistic Predictors Detect Fraudulent Financial Filings? *Journal of Emerging Technologies in Accounting*, 7(1), 25-46. <https://doi.org/10.2308/jeta.2010.7.1.25>
- Goel, S., & Uzuner, O. (2016). Do Sentiments Matter in Fraud Detection? Estimating Semantic Orientation of Annual Reports: Do Sentiments Matter in Fraud Detection: Estimating Semantic Orientation of Annual Reports. *Intelligent Systems in Accounting, Finance and Management*, 23(3), 215-239. <https://doi.org/10.1002/isaf.1392>
- Granlund, M. (2011). Extending AIS research to management accounting and control issues: A research note. *International Journal of Accounting Information Systems*, 12(1), 3-19. <https://doi.org/10.1016/j.accinf.2010.11.001>
- Hájek, P., & Olej, V. (2015). *Word Categorization of Corporate Annual Reports for Bankruptcy Prediction by Machine Learning Methods* (P. Král & V. Matoušek, Ed.; C. 9302, ss. 122-130). Springer International Publishing. https://doi.org/10.1007/978-3-319-24033-6_14
- Hajek, P., Olej, V., & Myskova, R. (2014). Forecasting Corporate Financial Performance Using Sentiment in Annual Reports For Stakeholders' Decision-Making. *Technological and Economic Development of Economy*, 20(4), 721-738. <https://doi.org/10.3846/20294913.2014.979456>
- Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed). Springer.

- Humpherys, S. L., Moffitt, K. C., Burns, M. B., Burgoon, J. K., & Felix, W. F. (2011). Identification of fraudulent financial statements using linguistic credibility analysis. *Decision Support Systems, 50*(3), 585-594. <https://doi.org/10.1016/j.dss.2010.08.009>
- Issa, H., Sun, T., & Vasarhelyi, M. A. (2016). Research Ideas for Artificial Intelligence in Auditing: The Formalization of Audit and Workforce Supplementation. *Journal of Emerging Technologies in Accounting, 13*(2), 1-20. <https://doi.org/10.2308/jeta-10511>
- Jans, M., Alles, M. G., & Vasarhelyi, M. A. (2014). A Field Study on the Use of Process Mining of Event Logs as an Analytical Procedure in Auditing. *The Accounting Review, 89*(5), 1751-1773. <https://doi.org/10.2308/accr-50807>
- Jo, N., & Shin, K. (2016). Bankruptcy Prediction Modeling Using Qualitative Information Based on Big Data Analytics. *Journal of Intelligence and Information Systems, 22*(2), 33-56. <https://doi.org/10.13088/JIIS.2016.22.2.033>
- Johnson, R. A., & Wichern, D. W. (2007). *Applied multivariate statistical analysis* (6th ed). Pearson Prentice Hall.
- Kelleher, J. D. (2019). *Deep learning*. The MIT Press.
- Kim, S. Y. (2011). Prediction of hotel bankruptcy using support vector machine, artificial neural network, logistic regression, and multivariate discriminant analysis. *The Service Industries Journal, 31*(3), 441-468. <https://doi.org/10.1080/02642060802712848>
- Kim, S. Y., & Upneja, A. (2014). Predicting restaurant financial distress using decision tree and AdaBoosted decision tree models. *Economic Modelling, 36*, 354-362. <https://doi.org/10.1016/j.econmod.2013.10.005>
- Kirkos, E. (2009). Modeling the auditors' opinions by using association rules. İçinde *Data Mining and Management* (s. 16). Nova Science Publishers, Inc.
- Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007a). *Applying Data Mining Methodologies for Auditor Selection*. 165-178.
- Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007b). Data Mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications, 32*(4), 995-1003. <https://doi.org/10.1016/j.eswa.2006.02.016>
- Kirkos, E., Spathis, C., & Manolopoulos, Y. (2008). Support vector machines, Decision Trees and Neural Networks for auditor selection. *Journal of Computational Methods in Sciences and Engineering, 8*, 213-224.
- Kirkos, E., Spathis, C., & Manolopoulos, Y. (2009). Audit-firm group appointment: An artificial intelligence approach. *Intelligent Systems in Accounting, Finance & Management, 17*(1), 1-17. <https://doi.org/10.1002/isaf.310>
- Kirkos, E., Spathis, C., Nanopoulos, A., & Manolopoulos, Y. (2007). Identifying Qualified Auditors' Opinions: A Data Mining Approach. *Journal of Emerging Technologies in Accounting, 4*(1), 183-197. <https://doi.org/10.2308/jeta.2007.4.1.183>
- Koh, H. C. (1991). Model Predictions and Auditor Assessments of Going Concern Status. *Accounting and Business Research, 21*(84), 331-338. <https://doi.org/10.1080/00014788.1991.9729848>
- Koskivaara, E. (2004). Artificial neural networks in analytical review procedures. *Managerial Auditing Journal, 19*(2), 191-223. <https://doi.org/10.1108/02686900410517821>
- Kotsiantis, S., Koumanakos, E., Tzelepis, D., & Tampakas, V. (2006). Forecasting Fraudulent Financial Statements using Data Mining. *INTERNATIONAL JOURNAL OF COMPUTATIONAL INTELLIGENCE, 3*(2), 104-110.
- KPMG. (2017). *Audit 2025: The Future is Now*. KPMG. <https://assets.kpmg/content/dam/kpmg/us/pdf/2017/03/us-audit-2025-final-report.pdf>
- KPMG. (2018). *KPMG Clara*. KPMG. <https://assets.kpmg/content/dam/kpmg/qm/pdf/kpmg-clara-a-smart-audit-platform.pdf>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444. <https://doi.org/10.1038/nature14539>
- Li, H., & Sun, J. (2011). Predicting business failure using support vector machines with straightforward wrapper: A re-sampling study. *Expert Systems with Applications, 38*(10), 12747-12756. <https://doi.org/10.1016/j.eswa.2011.04.064>

- Liang, D., Tsai, C.-F., Lu, H.-Y. (Richard), & Chang, L.-S. (2020). Combining corporate governance indicators with stacking ensembles for financial distress prediction. *Journal of Business Research*, 120, 137-146. <https://doi.org/10.1016/j.jbusres.2020.07.052>
- Lombardi, D. R., Bloch, R., & Vasarhelyi, M. A. (2014). The Future of Audit. *Journal of Information Systems and Technology Management*, 11(1), 21-32. <https://doi.org/10.4301/S1807-17752014000100002>
- Lopatta, K., Gloger, M. A., & Jaeschke, R. (2017). Can Language Predict Bankruptcy? The Explanatory Power of Tone in 10-K Filings. *Accounting Perspectives*, 16(4), 315-343. <https://doi.org/10.1111/1911-3838.12150>
- Loughran, T., & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35-65. <https://doi.org/10.1111/j.1540-6261.2010.01625.x>
- Loughran, T., & McDonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research*, 54(4), 1187-1230. <https://doi.org/10.1111/1475-679X.12123>
- Luo, J., Meng, Q., & Cai, Y. (2018). Analysis of the Impact of Artificial Intelligence Application on the Development of Accounting Industry. *Open Journal of Business and Management*, 06(04), 850-856. <https://doi.org/10.4236/ojbm.2018.64063>
- Magnusson, C., Arppe, A., Eklund, T., Back, B., Vanharanta, H., & Visa, A. (2005). The language of quarterly reports as an indicator of change in the company's financial status. *Information & Management*, 42(4), 561-574. <https://doi.org/10.1016/j.im.2004.02.008>
- Makrygiannis, K. (2019). *EY announces the first solution designed to help gauge impact and trustworthiness of artificial intelligence systems*. https://www.ey.com/en_gl/news/2019/04/ey-announces-the-first-solution-designed-to-help-gauge-impact-and-trustworthiness-of-artificial-intelligence-systems
- Manita, R., Elommal, N., Baudier, P., & Hikkerova, L. (2020). The digital transformation of external audit and its impact on corporate governance. *Technological Forecasting and Social Change*, 150, 119751. <https://doi.org/10.1016/j.techfore.2019.119751>
- Manyika, J., Chui, M., Miremadi, M., Bughin, J., George, K., Willmott, P., & Dewhurst, M. (2017). *A Future That Works: Automation, Employment and Productivity*. McKinsey&Company. <https://www.mckinsey.com/~media/mckinsey/featured%20insights/Digital%20Disruption/Harnessing%20automation%20for%20a%20future%20that%20works/MGI-A-future-that-works-Executive-summary.ashx>
- Martens, D., Bruynseels, L., Baesens, B., Willekens, M., & Vanthienen, J. (2008). Predicting going concern opinion with data mining. *Decision Support Systems*, 45(4), 765-777. <https://doi.org/10.1016/j.dss.2008.01.003>
- Mecca, G., Raunich, S., & Pappalardo, A. (2007). A new algorithm for clustering search results. *Data & Knowledge Engineering*, 62(3), 504-522. <https://doi.org/10.1016/j.datak.2006.10.006>
- Melendez, C. (2016). *Melendez.pdf*. <https://www.infoworld.com/article/3044468/Artificial-Intelligence-Gets-into-Auditing-Whats-next.html>. <https://www.infoworld.com/article/3044468/artificial-intelligence-gets-into-auditing-whats-next.html>
- Meredith, K., Blake, J., Baxter, P., & Kerr, D. (2020). Drivers of and barriers to decision support technology use by financial report auditors. *Decision Support Systems*, 139, 113402. <https://doi.org/10.1016/j.dss.2020.113402>
- Olson, D. L., Delen, D., & Meng, Y. (2012). Comparative analysis of data mining methods for bankruptcy prediction. *Decision Support Systems*, 52(2), 464-473. <https://doi.org/10.1016/j.dss.2011.10.007>
- Ozdogoglu, G., Ozdogoglu, A., Gumus, Y., & Kurt Gumus, G. (2016). The Application of Data Mining Techniques in Manipulated Financial Statement Classification: The Case of Turkey. *Journal of AI and Data Mining, Online First*. <https://doi.org/10.22044/jadm.2016.664>
- PwC. (2016). *Technology in the PwC Audit Driving innovation*. PwC. <https://www.pwchk.com/en/audit-assurance/technology-in-pwc-audit.pdf>
- PwC. (2019). *Audit Explorer—The best people empowered by market-leading technologies*. PwC. <https://www.pwc.com/gx/en/audit-services/assets/pdf/audit-explorer-at-a-glance-on-screen.pdf>

- Qiu, X. Y., Jiang, S., & Deng, K. (2013). Automatic Assessment of Information Disclosure Quality in Chinese Annual Reports. İçinde G. Zhou, J. Li, D. Zhao, & Y. Feng (Ed.), *Natural Language Processing and Chinese Computing* (C. 400, ss. 288-298). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-41644-6_27
- Rahman, R. A., Masrom, S., Zakaria, N. B., & Halid, S. (2021). Auditor Choice Prediction Model Using Corporate Governance and Ownership Attributes: Machine Learning Approach. *International Journal of Emerging Technology and Advanced Engineering*, 11(7), 87-94. https://doi.org/10.46338/ijetae0721_11
- Rahrovi Dastjerdi, A., Foroghi, D., & Kiani, G. H. (2019). Detecting manager's fraud risk using text analysis: Evidence from Iran. *Journal of Applied Accounting Research*, 20(2), 154-171. <https://doi.org/10.1108/JAAR-01-2018-0016>
- Ramlukan, R. (2015). *How big data and analytics are transforming the audit*. Ernst & Young Global. https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/assurance/assurance-pdfs/ey-reporting-big-data-transform-audit.pdf?download#:~:text=The%20transformed%20audit%20will%20expand,and%20more%20relevant%20business%20insights.
- Raphael, J. (2015). *How Artificial Intelligence Can Boost Audit Quality*. <https://www.cfo.com/Auditing/2015/06/Artificial-Intelligence-Can-Boost-Audit-Quality/>. <https://www.cfo.com/auditing/2015/06/artificial-intelligence-can-boost-audit-quality/>
- Saif, S. M., Sarikhani, M., & Ebrahimi, F. (2012). Finding Rules for Audit Opinions Prediction Through Data Mining Methods. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2185919>
- Saif, S. M., Sarikhani, M., & Ebrahimi, F. (2013). An Expert System with Neural Network and Decision Tree for Predicting Audit Opinions. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 2(4), 151-158. <https://doi.org/10.11591/ij-ai.v2i4.3950>
- Sánchez-Serrano, J. R., Alaminos, D., García-Lagos, F., & Callejón-Gil, A. M. (2020). Predicting Audit Opinion in Consolidated Financial Statements with Artificial Neural Networks. *Mathematics*, 8(8), 1288. <https://doi.org/10.3390/math8081288>
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85-117. <https://doi.org/10.1016/j.neunet.2014.09.003>
- Serrano-Cinca, C., & Gutiérrez-Nieto, B. (2013). Partial Least Square Discriminant Analysis for bankruptcy prediction. *Decision Support Systems*, 54(3), 1245-1255. <https://doi.org/10.1016/j.dss.2012.11.015>
- Sevilengül, O. (2016). *Genel Muhasebe* (2016. bs). Gazi Kitabevi.
- Shirata, C. Y., Takeuchi, H., Ogino, S., & Watanabe, H. (2011). Extracting Key Phrases as Predictors of Corporate Bankruptcy: Empirical Analysis of Annual Reports by Text Mining. *Journal of Emerging Technologies in Accounting*, 8(1), 31-44. <https://doi.org/10.2308/jeta-10182>
- Smola, A., & Vishwanathan, S. V. N. (2008). *Introduction to Machine Learning* (1. bs). Cambridge University Press.
- Stanišić, N., Radojević, T., & Stanić, N. (2019). Predicting the type of auditor opinion: Statistics, machine learning, or a combination of the two? *The European Journal of Applied Economics*, 16(2), 1-58. <https://doi.org/10.5937/EJAE16-21832>
- Sun, J., Li, H., Fujita, H., Fu, B., & Ai, W. (2020). Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting. *Information Fusion*, 54, 128-144. <https://doi.org/10.1016/j.inffus.2019.07.006>
- Sun, T., & Vasarhelyi, M. A. (2017). How an Evolving Technology Could Transform Analysis and Improve Judgment. *The CPA Journal*, 87(6), 24-29.
- Sutton, S. G., Reinking, J., & Arnold, V. (2011). On the Use of Grounded Theory as a Basis for Research on Strategic and Emerging Technologies in Accounting. *Journal of Emerging Technologies in Accounting*, 8(1), 45-63. <https://doi.org/10.2308/jeta-10207>
- Tang, F., Norman, C. S., & Vendrzyk, V. P. (2017). Exploring perceptions of data analytics in the internal audit function. *Behaviour & Information Technology*, 36(11), 1125-1136. <https://doi.org/10.1080/0144929X.2017.1355014>

- Tsai, C.-F., & Chiou, Y.-J. (2009). Earnings management prediction: A pilot study of combining neural networks and decision trees. *Expert Systems with Applications*, 36(3), 7183-7191. <https://doi.org/10.1016/j.eswa.2008.09.025>
- Tung, W. L., Quek, C., & Cheng, P. (2004). GenSo-EWS: A novel neural-fuzzy based early warning system for predicting bank failures. *Neural Networks*, 17(4), 567-587. <https://doi.org/10.1016/j.neunet.2003.11.006>
- Veganzones, D., & Séverin, E. (2018). An investigation of bankruptcy prediction in imbalanced datasets. *Decision Support Systems*, 112, 111-124. <https://doi.org/10.1016/j.dss.2018.06.011>
- Volkov, A., Benoit, D. F., & Van den Poel, D. (2017). Incorporating sequential information in bankruptcy prediction with predictors based on Markov for discrimination. *Decision Support Systems*, 98, 59-68. <https://doi.org/10.1016/j.dss.2017.04.008>
- Wagner, J. M. (2016). *Continuous Auditing—The Future of Internal Audit?* 3244-3252.
- Wang, G., Chen, G., & Chu, Y. (2018). A new random subspace method incorporating sentiment and textual information for financial distress prediction. *Electronic Commerce Research and Applications*, 29, 30-49. <https://doi.org/10.1016/j.elerap.2018.03.004>
- Werner, M., & Gehrke, N. (2015). Multilevel Process Mining for Financial Audits. *IEEE Transactions on Services Computing*, 8(6), 820-832. <https://doi.org/10.1109/TSC.2015.2457907>
- Werner, M., & Gehrke, N. (2019). Identifying the Absence of Effective Internal Controls: An Alternative Approach for Internal Control Audits. *Journal of Information Systems*, 33(2), 205-222. <https://doi.org/10.2308/isys-52112>
- Witten, I. H. (2003). *Computer Science, University of Waikato, Hamilton, New Zealand email ihw@cs.waikato.ac.nz*. 23.
- Wu, D., Liang, L., & Yang, Z. (2008). Analyzing the financial distress of Chinese public companies using probabilistic neural networks and multivariate discriminate analysis. *Socio-Economic Planning Sciences*, 42(3), 206-220. <https://doi.org/10.1016/j.seps.2006.11.002>
- Xu, X., & Wang, Y. (2009). Financial failure prediction using efficiency as a predictor. *Expert Systems with Applications*, 36(1), 366-373. <https://doi.org/10.1016/j.eswa.2007.09.040>
- Yan, J., Wang, X., Wang, B., & Zhang, Y. (2019). Research on Application of Data Mining Technology in Risk Assessment Process of Audit. *2019 International Conference on Economic Management and Model Engineering (ICEMME)*, 487-491. <https://doi.org/10.1109/ICEMME49371.2019.00102>
- Yaşar, A., Yakut, E., & Gutnu, M. (2015). Predicting Qualified Audit Opinions Using Financial Ratios: Evidence from the Istanbul Stock Exchange. *International Journal of Business and Social Science*, 6(8), 57-67.
- Yeh, C.-C., Chi, D.-J., & Lin, Y.-R. (2014). Going-concern prediction using hybrid random forests and rough set approach. *Information Sciences*, 254, 98-110. <https://doi.org/10.1016/j.ins.2013.07.011>
- Yigitbasioglu, O. M., & Velcu, O. (2012). A review of dashboards in performance management: Implications for design and research. *International Journal of Accounting Information Systems*, 13(1), 41-59. <https://doi.org/10.1016/j.accinf.2011.08.002>