T.C.

BARTIN ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

İŞLETME ANABİLİM DALI

YÜKSEK LİSANS TEZİ

İŞLETMELERDE FİNANSAL BAŞARISIZLIK RİSKİNİN YAPAY ZEKÂ TEKNİKLERİYLE TAHMİN EDİLMESİNE DAİR BİR UYGULAMA

ÇAĞATAY SİNOPLU

DANIŞMAN

DR. ÖĞR. ÜYESİ İSMAİL FATİH CEYHAN

BARTIN-2024

T.C.
BARTIN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
İŞLETME ANABİLİM DALI

İŞLETMELERDE FİNANSAL BAŞARISIZLIK RİSKİNİN YAPAY ZEKÂ TEKNİKLERİYLE TAHMİN EDİLMESİNE DAİR BİR UYGULAMA

YÜKSEK LİSANS TEZİ

Çağatay SİNOPLU

JÜRİ ÜYELERİ

Danışman	:	Dr. Öğr. Üyesi İsmail Fatih CEYHAN
Üye	:	Doç. Dr. Yaşar ÖZ
Üye	:	Dr. Öğr. Üyesi Abdullah BAYRAM

BARTIN-2024

KABUL VE ONAY

Çağatay SİNOPLU tarafından hazırlanan "İŞLETMELERDE FİNANSAL BAŞARISIZLIK RİSKİNİN YAPAY ZEKÂ TEKNİKLERİYLE TAHMİN EDİLMESİNE DAİR BİR UYGULAMA" başlıklı bu çalışma, 25.10.2024 tarihinde yapılan savunma sınavı sonucunda oy birliği ile başarılı bulunarak jürimiz tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

Başkan	:	Doç. Dr. Yaşar ÖZ	……
Üye	:	Dr. Öğr. Üyesi İsmail Fatih CEYHAN	……
Üye	:	Dr. Öğr. Üyesi Abdullah BAYRAM	……

Bu tezin kabulü Lisansüstü Eğitim Enstitüsü Yönetim Kurulu'nun ../../202.. tarih ve ……. sayılı kararıyla onaylanmıştır.

Prof. Dr. Mustafa Sabri GÖK
Enstitü Müdürü

BEYANNAME

Bartın Üniversitesi Lisansüstü Eğitim Enstitüsü tez yazım kılavuzuna göre Dr. Öğr. Üyesi İsmail Fatih CEYHAN danışmanlığında hazırlamış olduğum "İŞLETMELERDE FİNANSAL BAŞARISIZLIK RİSKİNİN YAPAY ZEKÂ TEKNİKLERİYLE TAHMİN EDİLMESİNE DAİR BİR UYGULAMA" başlıklı yüksek lisans tezimin bilimsel etik değerlere ve kurallara uygun, özgün bir çalışma olduğunu, aksinin tespit edilmesi halinde her türlü yasal yaptırımı kabul edeceğimi beyan ederim.

25.10.2024

Çağatay SİNOPLU

ÖN SÖZ

Bu tez çalışmasında işletmelerde finansal başarısızlık riskinin yapay zekâ teknikleriyle tahmin edilmesi konusu ele alınmıştır. Öncelikle, tez konusunun seçiminden yazım ve sonuç aşamasına kadar bana desteğini esirgemeyen danışman hocam Dr. Öğr. Üyesi İsmail Fatih CEYHAN'a teşekkürlerimi sunarım. Yüksek lisans öğrenimim boyunca yardımlarını esirgemeyen Arş. Gör. Ferhat DEMİRCİ'ye ve tüm hayatım boyunca her konuda beni destekleyen aileme teşekkürlerimi bir borç bilirim.

.

Çağatay SİNOPLU

ÖZET

Yüksek Lisans Tezi

İŞLETMELERDE FİNANSAL BAŞARISIZLIK RİSKİNİN YAPAY ZEKÂ TEKNİKLERİYLE TAHMİN EDİLMESİNE DAİR BİR UYGULAMA

Çağatay SİNOPLU

Bartın Üniversitesi
Lisansüstü Eğitim Enstitüsü
İşletme Anabilim Dalı

Tez Danışmanı: Dr. Öğr. Üyesi İsmail Fatih CEYHAN

Bartın-2024, sayfa: 124

Bu çalışma, işletmelerde finansal başarısızlık riskinin belirlenmesi amacıyla kullanılan yapay zekâ tekniklerinin etkinliğini araştırmaktadır. Yapay sinir ağları, destek vektör makineleri ve diğer yapay zekâ modelleri kullanılarak Borsa İstanbul (BİST) Ulusal Pazar'daki rasgele seçilen 14 firmanın ve Yakın İzleme Pazarı'ndaki firmaların tamamının mali verileri analiz edilmiştir. Bu firmaların 2013-2022 yılları arasındaki üçer aylık finansal tablolarında yer alan 43 adet finansal oran analizde kullanılmıştır. M5Prime, Random Forest, Random Tree, M5Rules, Random Committee, Lazy Instance-Based K (IBk) ve Locally Weighted Learning (LWL) kullanılan regresyon modelleri arasında yer almaktadır. Çalışma sonuçlarına göre, M5Rules ve LWL en iyi performans gösteren modeller olarak öne çıkmıştır. M5Prime, yüksek korelasyon katsayısı ve düşük hata metrikleri ile başarılı bulunmuştur. Random Forest ve Random Committee modelleri de yüksek korelasyon katsayıları (0.99 civarında) ve makul hata metrikleri ile iyi performans sergilemiştir. Ancak, Random Tree ve IBk modelleri diğer modellere kıyasla daha düşük performans göstermiş, özellikle IBk modeli en yüksek hata metriklerine ve en düşük korelasyon katsayısına sahip olmuştur. Bu bulgular, yapay zekâ tabanlı modellerin finansal başarısızlık riskini tahmin etmede geleneksel yöntemlere kıyasla daha yüksek doğruluk ve genelleme kabiliyeti sunduğunu ve farklı alanlar için bu uygulamaların genişletilebileceğini göstermektedir.

Anahtar Kelimeler: Derin Öğrenme, Finansal Başarısızlık, Makine Öğrenmesi, Veri Madenciliği, Yapay Sinir Ağları, Yapay Zekâ.

ABSTRACT

M.Sc. Thesis

AN APPLICATION ON ESTIMATING FINANCIAL FAILURE RISK IN BUSINESSES WITH ARTIFICIAL INTELLIGENCE TECHNIQUES

Çağatay SİNOPLU

Bartin University
Graduate School
Department of Business Administration

Thesis Advisor: Assist. Prof. Dr. İsmail Fatih CEYHAN
Bartın-2024, pp: 124

This study investigates the effectiveness of artificial intelligence techniques in determining the risk of financial failure in businesses. The financial data of 14 randomly selected companies from the Borsa Istanbul (BIST) National Market and all companies listed on the Watchlist Market were analyzed using artificial neural networks, support vector machines, and other AI models. The analysis utilized 43 financial ratios derived from the quarterly financial statements of these firms for the period 2013-2022. The regression models employed include M5Prime, Random Forest, Random Tree, M5Rules, Random Committee, Lazy Instance-Based K (IBk), and Locally Weighted Learning (LWL). According to the study's findings, M5Rules and LWL emerged as the best-performing models. M5Prime was found to be successful due to its high correlation coefficient and low error metrics. The Random Forest and Random Committee models also performed well, with high correlation coefficients (approximately 0.99) and reasonable error metrics. However, the Random Tree and IBk models demonstrated lower performance compared to the other models, with the IBk model in particular having the highest error metrics and the lowest correlation coefficient. These findings suggest that AI-based models offer higher accuracy and generalization capabilities in predicting financial failure risk compared to traditional methods, and that these applications can be extended to different fields.

Keywords: Artificial Intelligence, Artificial Neural Networks, Data Mining, Deep Learning, Financial Failure, Machine Learning.

İÇİNDEKİLER

ŞEKİLLER DİZİNİ

Şekil No **Sayfa No**

TABLOLAR DİZİNİ

KISALTMALAR

ANOVA : Analysis of Variance

AUC : Alan Altındaki Alan

BMU : Best Matching Unit

BİST : Borsa İstanbul

CPT : Condition Possibility Table

FAVÖK : Faiz ve Vergi Öncesi Kar

FCM : Fuzzy C-Means

FK : Fiyat Kazanç

İMKB : İstanbul Menkul Kıymetler Borsası

k-NN : k-En Yakın Komşu

LBK : Lazy Instance-Based K

LWL : Locally Weighted Learning

MAE : Ortalama Mutlak Hata

M5P : M5Prime

MSE : Ortalama Kare Hata

PD : Piyasa Değeri

ROIC : Return of Invested Capital

RUL : Remaining Useful Life

SMOTE : Synthetic Minority Oversampling Technique

t-SNE : t-Stokastik Komşu Gömme

TBA : Temel Bileşen Analizi

VAFÖK : Vergi, Amortisman ve Faiz Öncesi Kar

VNÖ : Vektör Nicelemeyi Öğrenme

YDG : Yerel Doğrusal Gömme

YÖK : Yüksek Öğretim Kurumu

YSA : Yapay Sinir Ağları

YZ : Yapay Zekâ

1. GİRİŞ

İşletmelerin finansal başarısızlık riski, ekonomik belirsizliklerin arttığı ve rekabetin yoğunlaştığı günümüz iş dünyasında kritik bir konu haline gelmiştir. Bir işletmenin finansal başarısızlığa uğrama olasılığını önceden tahmin edebilmek, yatırımcılar, hissedarlar ve şirket yöneticileri için büyük önem taşır. Finansal başarısızlık, işletmelerin iflası veya finansal yükümlülüklerini yerine getirememe durumu olarak tanımlanır. Bu tür başarısızlıkların önceden öngörülebilmesi, potansiyel risklerin zamanında tespit edilmesi ve gerekli önlemlerin alınması açısından oldukça değerlidir.

Geleneksel finansal analiz yöntemleri, işletmelerin mali tablolarını ve finansal oranlarını kullanarak finansal başarısızlık riskini belirlemeye çalışır. Ancak, bu yöntemler her zaman doğru ve güvenilir sonuçlar vermez. İşte bu noktada, yapay zekâ teknikleri devreye girerek finan1.sal başarısızlık riskinin daha etkili ve isabetli bir şekilde tahmin edilmesini sağlar. Yapay zekâ, büyük veri setlerinden anlamlı bilgileri çıkarmak için güçlü bir araçtır ve karmaşık ilişkileri analiz edebilir.

Bu çalışma, işletmelerde finansal başarısızlık riskini tahmin etmek için yapay zekâ tekniklerinin nasıl kullanılabileceğini incelemektedir. Çalışmanın amacı, yapay zekâ tabanlı modellerin finansal başarısızlık riskini tahmin etme konusunda ne kadar etkili olduğunu ortaya koymaktır. Çalışma kapsamında, belirli işletmelerin mali verileri analiz edilerek, finansal başarısızlık riskini tahmin etmek için yapay sinir ağları, destek vektör makineleri ve diğer yapay zekâ modelleri uygulanacaktır.

Bu çalışmada kullanılan regresyon modelleri arasında M5Prime, Random Forest, Random Tree, M5Rules, Random Committee, Lazy Instance-Based K (IBk) ve Locally Weighted Learning (LWL) bulunmaktadır. M5Prime, karar ağaçlarını ve regresyon analizini birleştiren güçlü bir modeldir. Random Forest, birden fazla karar ağacı oluşturarak ve bunları birleştirerek tahmin doğruluğunu artıran bir yöntemdir. Random Tree, veri setinin rastgele alt kümeleri üzerinde karar ağaçları oluşturarak çalışan bir algoritmadır. M5Rules, karar kurallarını çıkarmak için M5 algoritmasını kullanır ve bu kurallar üzerinden tahminler yapar. Random Committee, farklı modellerin çıktılarının birleştirilmesiyle daha doğru tahminler elde etmeyi amaçlayan bir yöntemdir. IBk, benzer veri noktalarına dayalı olarak tahminler

yapan tembel öğrenme algoritmasıdır. Locally Weighted Learning (LWL) ise her bir tahmin için verinin yerel bir alt kümesini kullanarak ağırlıklı bir öğrenme süreci uygular.

Weka programında yapılan analizler sonucunda, M5Rules ve LWL en iyi performans gösteren modeller olmuştur. M5Prime, oldukça yüksek bir korelasyon katsayısı ve düşük hata metrikleri ile iyi performans sergilemektedir. RandomForest ve Random Committee modelleri, nispeten iyi performans göstermektedir. Bu modellerin korelasyon katsayıları 0.99 civarında olup, hata metrikleri makul seviyelerdedir. Bu modeller, M5Rules ve M5Prime kadar olmasa da yine de güçlü tahmin yeteneklerine sahiptir. Ancak, Random Tree ve IBk modelleri diğer modellere kıyasla daha düşük performans göstermektedir. Özellikle IBk modeli, en yüksek hata metriklerine ve en düşük korelasyon katsayısına sahiptir.

Bu yapay zekâ tabanlı modeller, işletmelerin mali durumlarına ilişkin çeşitli değişkenleri kullanarak finansal başarısızlık riskini tahmin eder. Geleneksel yöntemlere kıyasla, bu modeller daha yüksek doğruluk oranları ve genelleme kabiliyetleri sunar. Özellikle büyük veri setleri üzerinde etkili bir şekilde çalışabilen bu modeller, işletmelerin mali sağlığını değerlendirmede ve olası finansal riskleri öngörmede önemli bir avantaj sağlar. Çalışmanın bulguları, yapay zekâ tekniklerinin finansal başarısızlık riskini tahmin etmede önemli bir rol oynayabileceğini ve bu alandaki uygulamalarının genişletilebileceğini göstermektedir.

İkinci bölümde, finansal başarısızlık kavramı detaylı bir şekilde tanımlanmış ve finansal başarısızlığın başlıca nedenleri üzerinde durulmuştur. Bu nedenler arasında yönetim hataları, yetersiz sermaye, ekonomik durgunluk, rekabet baskıları ve piyasa koşullarındaki değişiklikler gibi faktörler yer almaktadır. Finansal başarısızlığı öngörmenin önemi, işletmelerin hayatta kalabilmesi ve stratejik kararlar alabilmesi açısından vurgulanmıştır. Ayrıca, finansal başarısızlığın sonuçları, işletmelerin iflası, itibar kaybı ve ekonomik etkiler gibi çeşitli açılardan ele alınmıştır. Finansal başarısızlığı ölçen modeller hakkında bilgi verilmiş ve bu modellerin işletmeler için nasıl kullanıldığı açıklanmıştır. Finansal sıkıntı, ödeme güçlüğü ve iflas kavramları arasındaki farklar da ayrıntılı bir şekilde incelenmiştir.

Üçüncü bölümde, yapay zekâ uygulamaları ve bu uygulamaların işletmelerde nasıl kullanıldığı ele alınmıştır. Makine öğrenmesi, yapay sinir ağları ve derin öğrenme gibi yapay zekâ teknikleri hakkında kapsamlı bilgi verilmiş ve bu tekniklerin finansal analizlerde nasıl

kullanılabileceği açıklanmıştır. Veri madenciliği konusuna da değinilmiştir

Dördüncü bölümde, finansal başarısızlığı öngörmek için kullanılabilecek regresyon modelleri açıklanmıştır. Bu modellerin temel prensipleri üzerinde durulmuştur ve veri madenciliğinde karşılaşılan sorunlar ile bu sorunları önleme teknikleri üzerinde durulmuştur. Özellikle büyük veri setleri ile çalışırken karşılaşılan zorluklar ve bu zorlukların üstesinden gelme yolları incelenmiştir ve son olarak konu ile ilgili literatür kapsamlı bir şekilde incelenmiş ve mevcut çalışmalar özetlenmiştir.

Beşinci bölümde, çalışmada kullanılan veri setleri ve analiz yöntemleri hakkında ayrıntılı bilgiler sunulmuştur. Kullanılan veri toplama yöntemleri, analiz teknikleri ve bu tekniklerin doğruluğu ile güvenilirliği üzerinde durulmuştur ve yapılan uygulamalar sonucunda elde edilen bulgular tartışılmıştır. Bulguların analizleri yapılmış ve elde edilen sonuçların mevcut literatür ile karşılaştırmaları yapılmıştır.

Son bölüm olan altıncı bölümde ise genel sonuçlar yer almaktadır. Çalışmanın genel bir değerlendirmesi yapılmış, elde edilen bulgular özetlenmiş ve gelecekte yapılabilecek çalışmalar için önerilerde bulunulmuştur. Ayrıca, çalışmanın kısıtları ve bu kısıtların nasıl aşılabileceği konusunda fikirler sunulmuştur.

Bu çalışma, yapay zekâ teknikleri ile finansal başarısızlık riskinin tahmin edilmesi alanındaki literatüre katkı sağlamayı hedeflemektedir. Aynı zamanda, işletmelerin finansal risklerini daha iyi anlamalarına ve erken uyarı sistemleri oluşturma konusunda yol göstermeyi amaçlamaktadır. Bu giriş bölümü, çalışmanın arka planını ve amacını açıklayarak, finansal başarısızlık riskinin tahmin edilmesinde yapay zekâ tekniklerinin neden önemli olduğunu vurgulamaktadır.

2. FİNANSAL BAŞARISIZLIK

Mali başarısızlık terimi İngilizcede "Financial Failure" kavramı ile ifade edilir (Terzi, 2011). Finansal başarısızlık, işletmelerin mali durumunun kötüleştiği ve finansal taahhütlerini yerine getiremeyecek duruma geldiği bir durumu ifade eder. İşletmelerdeki başarısızlık genellikle iki temel kategoriye ayrılır: ekonomik başarısızlık ve finansal başarısızlık. Ekonomik başarısızlık, işletmenin yatırılan sermayeden beklenen getiriyi sağlayamaması durumudur. Mali başarısızlık ise kısa ve uzun vadeli finansal yükümlülüklerin yerine getirilememesi, tahvil faizlerinin ve anaparanın ödenemez hale gelmesi, çeklerin karşılanamaması, işletmeye kayyum atanması gibi durumları içerir. Bu durumlar, işletmenin mali güçlüğe düştüğünü ve varlıklarının yükümlülüklerini karşılayamayacak hale geldiğini gösterir (Karadeniz ve Öcek, 2020).

İşletmelerin, özellikle de tüzel kişiliği haiz şirketlerin süresiz var olduğu teorik olarak kabul edilir. Ancak, genel olarak sınırsız ömürlü olarak kabul edilen şirketlerin bir kısmı, belirlenen amaçlara ulaştıktan sonra varlıklarını sonlandırabilirken, diğer bir kısmı çeşitli nedenlerle başarısızlığa uğrayabilir. Özellikle ekonomik çevredeki değişikliklere uyum sağlayamayan işletmelerin, pazarı terk etmeleri gibi durumlar sıkça yaşanmaktadır (Uzun, 2005).

2.1. Finansal Başarısızlık Tanımları

Literatürdeki bazı finansal başarısızlık tanımları;

Tablo 1.1: Finansal başarısızlık tanımları

Yazar	Tanım
(Altman, 1968)	Yasal prosedürler çerçevesinde iflas etmiş ve bir denetçi atanmış veya ABD Ulusal İflas Yasası'na göre yeniden yapılanma talep edilmiş işletmeleri ifade etmektedir.
(Elarm, 1972)	İflas yasalarına göre resmi olarak iflas etmiş kabul edilen işletmeler.
(Taffler, 1982)	Tasfiye, alacaklıların talebi üzerine veya mahkeme kararıyla işletmenin faaliyetlerine son verme işlemidir.
(Blum, 1974)	Vadesi dolmuş borçlarının karşılığını bulamama, iflas sürecine

	girmiş olma ve kredi verenler ile borçların bir kısmının silinmesi konusunda uzlaşma.
(Kulalı, 2016)	Faal olan firmaların, borçlarını ödemek için elde ettikleri nakit akışı yetersiz kalmaktadır.
(Selimoğlu ve Orhan, 2015)	İşletmenin mali taahhütlerini yerine getirirken zorluk çekmesi veya taahhütleri yerine getirememesi.

2.2. Finansal Başarısızlığın Nedenleri

Finansal başarısızlığın nedenleri çeşitli faktörlerin etkileşimiyle ortaya çıkar. Yanlış stratejilerin belirlenmesi, piyasa koşullarının yanlış değerlendirilmesi, yetersiz likidite yönetimi, kötü yönetim ve kontrol mekanizmaları, düşük verimlilik ve rekabet gücü gibi iç faktörlerin yanı sıra, hızlı büyüme, ani pazar değişimleri gibi dış faktörler de rol oynar. Bu nedenlerle finansal performans göstergeleri önem kazanır. Gelirler, kar marjları, nakit akışı, varlık verimliliği, borçluluk oranları gibi göstergeler, şirketin sağlığını, karlılığını ve sürdürülebilirliğini değerlendirmek için kullanılır. Ancak, bu göstergelerin yanlış yorumlanması veya göz ardı edilmesi, şirketin gerçek finansal durumunu anlamakta güçlük çekilmesine ve yanlış kararlar alınmasına yol açabilir. Dolayısıyla, finansal başarısızlığı önlemek için finansal performans göstergelerinin dikkatlice izlenmesi ve doğru bir şekilde yorumlanması hayati öneme sahiptir.

Bir işletme, başarısızlık sürecine giderken genellikle belirli belirtiler ve göstergeler gösterir. Finansal durumunun kötüleşmeye doğru ilerlediğine dair işaretler şunlar olabilir: (Akgüç, 1985);

- İşletmelerin oranlarında olumsuz eğilimler ve kötüye gidiş,
- Şirketin halka açık olması durumunda hisse senedi fiyatlarında sürekli ve hızlı düşüşler,
- İşletmelerin bankalardan kullandığı kredi limitlerini maksimum seviyede kullanması ve kredi hesaplarında hareketsizlik, banka mevduatlarının minimum seviyeye düşmesi ve bu durumun uzun süre devam etmesi,
- Ödemelerde gecikmelerin yaşanması.

İşletmelerin finansal başarısızlıkla karşılaşması, çeşitli şekillerde ve süreçlerde ortaya çıkabilir. Başarısızlık, işletmenin özel hedeflerine ulaşmada yetersiz kaldığı özel amaçlarla sınırlı olabileceği gibi, genel hedeflerine eşit derecede ulaşmakta yetersiz kaldığı genel amaçlarla ilgili de olabilir. Başarısızlık, genellikle içsel ve dışsal sebepler olmak üzere iki ana kategoriye ayrılabilir. İçsel başarısızlık sebepleri, işletmenin kendi faaliyetlerinden kaynaklanan ve yönetimin kontrol edebileceği içsel yönetimle ilgili nedenlerdir. Dışsal başarısızlık sebepleri ise, işletmenin çevresindeki faktörlerden kaynaklanan ve yönetimin etkisiz olduğu dışsal nedenlerdir (Baş, 2010).

İşletmelerin finansal açıdan başarısızlıkları genellikle içsel ve dışsal etkenlerden kaynaklanır. İşletmeler, kendi faaliyetlerinin yanı sıra çevresel etkilere de maruz kalır ve bu etkilerden etkilenir. İşletmeleri finansal açıdan başarısızlığa sürükleyen bazı faktörler, işletmenin kontrolü dışında gerçekleşebilir. Bu faktörlerin tamamen önlenebilmesi mümkün olmasa da işletmeler, etkilerini minimize etmeyi amaçlar. Bu etkenler arasında işletmenin bulunduğu sosyal ortam, yasal ve politik çevre, ekonomik koşullar, teknolojik gelişmeler ve doğal çevre gibi unsurlar yer alır (Ural, Gürada, vd., 2015).

İşletme içi başarısızlıklar, işletme riski olarak adlandırılır ve bunlar arasında yetersiz işletme sermayesi, aşırı kısa vadeli borç yükümlülükleri ve zayıf bütçe kontrolü gibi etkenler öne çıkar. İşletmenin içsel başarısızlığı, finansal olmayan faktörlere de bağlı olabilir. Şirketlerin başarısızlığında etkili olan finansal olmayan nedenler arasında yönetim zaafları, örgütsel çatışmalar, koordinasyon eksiklikleri, yeni ürünlerin geliştirilememesi ve yeni pazarlara açılamama gibi unsurlar bulunmaktadır (Ural, Gürada, vd., 2015).
Bunların yanı sıra yaşanan finansal krizler de şirketlerin finansal başarısızlığını tetikleyebilir (Kulalı, 2016).

2.3. Finansal Başarısızlığın İçsel Nedenleri

İşletmelerde finansal başarısızlığın içsel nedenleri arasında öne çıkanlar, yetersiz işletme sermayesi ve aşırı derecede kısa vadeli borçlanma şeklinde sıralanabilir (Uzun, 2005). Ayrıca, finansal içsel nedenler arasında yetersiz sermaye, aşırı derecede kısa vadeli borç yükümlülüklerin yanında ve bütçe kontrolündeki zayıflıklar gibi mali unsurlar öne çıkar (Ertan ve Ersan, 2019). Mali yapıda ve işletme sermayesinde devamlı ortaya çıkan

değişiklikler, işletmeyi bir avantajla ya da bir tehditle karşı karşıya getirebilir.

Şirketlerin finansal ve finansal olmayan konularda aldığı yanlış kararlardan kaynaklanan içsel başarısızlık nedenlerin şu şekilde sıralanabilir (Söylemez ve Türkmen, 2017):

- Firmanın yetersiz sayış yapması,
- Faaliyet giderlerinin yüksek olması,
- Alacak yönetimi konusu üzerinde etkili stratejiler belirleyememe,
- Stok devir hızında beklenmeyen değişiklikler,
- Âtıl üretim kapasitesi
- Finansal kaldıracın yüksek seviyelerde olması,
- İşletmenin bulunduğu konum,
- Rakiplerle rekabet edememe,
- Şirket satın almalarında yapılan hatalar,
- Likiditeyi yönetememe.

2.3.1. Yönetsel Hatalar

İşletmelerin finansal açıdan başarısız olmasının en büyük içsel nedenlerinden biri de kötü yönetimdir. İşletme yönetiminin sermaye yönetimi, alacak tahsil politikası, stok yönetimi, yatırım kararları ve kar dağıtım politikası gibi konularda yanlış kararlar alması, mali başarısızlığın önünü açabilir. Bu nedenle, kötü yönetim, sermaye yönetimi ve borçlanma kararları gibi faktörlerin etkisi daha da önemlidir (Selimoğlu ve Orhan, 2015).

2.3.2. İşletmenin Mali Durumu

İşletmelerin aşırı derecede borçlanma faktöründe, borcun vadesi kısa veya uzun fark etmeksizin, vade geldiğinde faiziyle beraber ödenmesi gereken yükümlülükler göz önünde bulundurulmalıdır. Ancak, bu borcun verimli bir şekilde kullanılmaması durumunda,

işletmenin borçlanma maliyeti artar ve bu da daha fazla faiz yükü getirir. Bu durum işletmelerin pazar fırsatlarından yararlanmasını engelleyebilir. Faiz ödemeleri, işletmelerin likiditesini önemli ölçüde azaltabilir ve borçlarını ödeme konusunda yetersiz hale getirebilir (Şahin ve Yangil, 2023).

Borçlanma kararı almadan önce oran analizi ve nakit akış analizi gibi değerlendirmeler yapılmalıdır. Bu analizler, işletmenin mevcut finansal durumunu anlamak için önemlidir. Bu bilgilerin yanı sıra işletmenin ihtiyaçlarına ve hedeflerine uygun bir borçlanma yapısı belirlenmelidir. Aşırı borçlanmadan kaçınılmalı ve borçların işletmenin ödeme kapasitesini aşmamasına dikkat edilmelidir. Bu şekilde, işletme finansal sürdürülebilirliğini koruyabilir ve olası riskleri en aza indirebilir (Selimoğlu ve Orhan, 2015).

2.4. Başarısızlığın Dışsal Nedenleri

Dışsal başarısızlık nedenleri, yöneticilerin müdahale edemediği ve işletmenin kontrolü dışındaki faktörlerden kaynaklanır (Baş, 2010). Bu faktörler genellikle yönetimin etkisi dışında olan ve işletmenin normal işleyişini olumsuz etkileyen sebeplerdir. Örneğin, ekonomik durgunluklar, politik istikrarsızlık, doğal afetler gibi dışsal faktörler işletmenin performansını etkileyebilir ve başarısızlığa yol açabilir. Bu tür dışsal faktörler genellikle yöneticilerin kontrolü dışındadır ve işletmenin her sürecinde etkili olabilir.

2.4.1. Toplumsal Çevre

İşletmeler, faaliyet gösterdikleri toplumun değerleri, normları, beklentileri ve talepleriyle etkileşim içindedirler. Bu etkileşim, işletmelerin ürün ve hizmetlerinin kabul edilmesini, marka imajını, müşteri sadakatini ve genel olarak finansal performanslarını etkileyebilir.

Başarılı işletmeler, toplumun yöresel ve bölgesel tercihlerini dikkate alarak ürün geliştirirler ve bu tercihleri karşılayacak kalıplarda ürünler sunarlar. İşletmeler, kâr amacı güderken ürettikleri ürünlerde tüketicilerin beklentilerini karşılamayı hedeflerler. Tüketiciler, satın alacakları ürünlerde kalite, fiyat, çekicilik, estetik, moda ve yenilik gibi özellikleri ararlar. Başarılı işletmeler, bu beklentilere uygun ürünler sunarak müşteri memnuniyetini artırır ve pazar paylarını güçlendirirler (Bakhshıyev, 2009).

2.4.2. Yasal ve Politik Çevre

İşletmeler genellikle belirli bir hukuki statüye sahip organizasyonlardır ve faaliyet gösterdikleri ülkelerin hukuki çerçevesine uymak zorundadırlar (Ural, 2014). Bu, işletmelerin kuruluş aşamasından başlayarak günlük iş operasyonlarına kadar çeşitli yasal düzenlemelere tâbi olmaları anlamına gelir. İşletmelerin faaliyet gösterdikleri sektöre, boyuta ve coğrafi konumlarına bağlı olarak farklı hukuki gereksinimlerle karşılaşabilirler. Hukuki düzenlemeler genellikle vergi yasaları, ticaret kuralları, çalışma ilişkileri, tüketici hakları ve çevre koruma gibi geniş bir yelpazeyi kapsar. İşletmeler, bu hukuki zorunluluklara uyum sağlamak ve yasal riskleri azaltmak için genellikle hukuk uzmanlarından destek alır veya hukuki danışmanlık hizmetleri kullanır. Bu şekilde, işletmeler hem yasal olarak uygun hareket etmeyi sağlar hem de hukuki sorunlardan kaynaklanan potansiyel zararları en aza indirmeye çalışır.

Bir şirketin başarısızlığa uğraması, sadece iş sahiplerini ve çalışanları değil, aynı zamanda toplumu ve genel ekonomiyi de etkiler. Geniş bir müşteri ve tedarikçi ağına sahip olan bir şirketin iflası, maliyetlerin yayılması nedeniyle tüm ekonomiyi olumsuz şekilde etkileyebilir.

Devletler, ekonomiye veya piyasalara doğrudan veya dolaylı olarak müdahale edebilirler (Öner, 2018). Bu müdahaleler genellikle ekonomik politika aracılığıyla gerçekleştirilir. Doğrudan müdahaleler arasında vergi politikası, para politikası, hükümet harcamaları ve düzenleyici tedbirler gibi ekonomik araçların kullanımı yer alır. Örneğin, hükümetler, faiz oranlarını belirleyerek para arzını kontrol edebilir veya belli sektörlere teşvikler sağlayabilir. Dolaylı müdahaleler ise genellikle yasal düzenlemeler, standartlar ve teşvikler gibi politika araçlarının kullanımını içerir. Bu müdahalelerde devletin amacı genellikle ekonomik istikrarı sağlamak, işsizliği azaltmak, enflasyonu kontrol altında tutmak ve genel refahı artırmaktır. Bu müdahaleler işletmeleri finansal yönden etkileyebilir.

2.4.3. Ekonomik Çevre

İşletmeler çevrelerindeki ekonomik koşullarla doğrudan bağlantılıdır ve işletme kararları bu koşulların etkisi altında şekillenir. İşletmeler, faaliyet gösterdikleri ekonomik sistemin bir

parçası olarak, genel ekonomik yapıdan, uygulanan ekonomik politikalardan ve finansal piyasalardan önemli ölçüde etkilenmektedir (Akkoç, 2007).

Genel ekonomik koşulların belirsizliği, tüketici harcamalarında azalmaya, talep düşüşlerine ve gelirlerde gerilemeye yol açabilir. Bu durum, işletmelerin kar marjlarını düşürebilir ve nakit akışlarını olumsuz etkileyebilir, bu da finansal zorluklarla karşı karşıya kalmalarına neden olabilir. Ayrıca, yüksek faiz oranları, artan maliyetler ve rekabetin şiddetlenmesi gibi ekonomik faktörler, işletmelerin finansal durumunu zayıflatabilir.

2.5. Finansal Başarısızlığı Öngörmenin Önemi

Finansal başarısızlığın öngörülmesi, ekonomik sistemde oluşabilecek olumsuz etkileri önlemek veya en aza indirmek açısından büyük önem taşır. Bu tahminlerin amacı, yatırımcılar, alacaklılar ve diğer paydaşlar için zararlı olabilecek işletmeleri belirlemektir. Erken uyarı sistemleri, işletmelerin kriz yaşamadan önce potansiyel sorunlarını tespit edebilmeleri için önemli bir araçtır. Bu sayede işletmeler, riskleri önceden görebilir ve gerekli önlemleri alarak finansal istikrarlarını koruyabilirler (Kılınç ve Seyrek, 2012).

Önceden finansal sorunların tespit edilmesi, işletmelerin kriz durumlarına karşı daha hazırlıklı olmalarını sağlar ve bu da finansal istikrarlarını korumalarına yardımcı olur. Ayrıca, finansal başarısızlığı öngörebilmek, yatırımcıların ve kredi verenlerin güvenini kazanmaya yardımcı olur ve işletmenin dış finansman kaynaklarına erişimini kolaylaştırır. Bunun yanı sıra, çalışanların işletme içindeki güvenlerini ve motivasyonlarını artırır ve rekabetçiliklerini sürdürülebilir kılar. Sonuç olarak, finansal başarısızlığı önceden tahmin etmek, işletmelerin uzun vadeli başarısı ve sürdürülebilirliği için kritik bir faktördür.

2.6. Finansal Başarısızlığın Sonuçları

- Şirketin kendi isteğiyle faaliyetlerine son vermesi
- Şirketin haciz, kayyum gibi yasal süreçlerle karşı karşıya kalması
- Şirketin borç yapılandırması için alacaklılarıyla uzlaşmaya varması.
- Şirketin faaliyetlerini sonlandırması veya iflas etmesi

2.7. Finansal Başarısızlığı Ölçen Modeller

Son zamanlarda, yapay zekâ yaklaşımları ve veri madenciliği teknikleri olan sinir ağları, genetik algoritmalar ve uzman sistemler, kurumsal finansal başarısızlık tahmininde yaygın olarak kullanılmaktadır. Bu teknikler, evrensel bir yaklaşım benimseme özelliğine ve geniş veri kümelerinden ve alan uzmanlarından yararlı bilgiler çıkarma yeteneğine sahiptir. Vapnik (1995) tarafından geliştirilen destek vektör makineleri, birçok çekici özelliği ve çok çeşitli problemlerde mükemmel genelleme performansı nedeniyle popülerlik kazanmıştır (Xu ve Wang, 2009).

Başlıca finansal başarısızlık tahmin modelleri aşağıdaki gibidir (Yıldız ve Gürkan, 2022);

- Beaver Modeli
- Tamari Modeli
- Altman Z Skoru Modelleri
- Meyer ve Pifer Modeli
- Deakin Modeli
- Weibel Modeli
- Sinkey Modeli
- Daniel Martin Modeli
- Springate Modeli
- Ohlson Skor Modeli
- Fulmer H Skor Modeli
- Diskriminant Analizi

2.7.1. Beaver Modeli

Beaver 1966 yılında yaptığı çalışmasında başarısızlığı, işletmelerin mali yükümlülüklerini yerine getirmede yetersiz kalmaları olarak tanımlamıştır. Araştırması için, 1954-1964 döneminde başarısız olan 79 işletmeyi seçmiştir. Bu işletmeler, Moody’s ve Dun & Bradstreet'in oluşturduğu başarısız işletme listelerinden seçilmiştir. Ayrıca, Moody's tarafından varlık büyüklükleri ve endüstri alanlarına göre hazırlanmış olan 12.000 başarılı işletme listesinden 79 işletme seçilmiştir. İşletmelerin seçiminde, endüstri alanlarına ve

büyüklüklerine göre sıralama ve gruplama yapılmıştır (Karadeniz ve Öcek, 2019).

Beaver, incelediği oranları literatürdeki en yaygın tercih edilme ve kullanım amaçlarına göre altı başlık altında gruplamıştır;

- Nakit Akış/Toplam Borçlar
- Net Kar/Toplam Varlıklar
- Toplam Borçlar/Toplam Varlıklar
- Dönen Varlıklar/Kısa Vadeli Yabancı Kaynaklar
- Net Çalışma Sermayesi/Faaliyet Giderleri
- Net Çalışma Sermayesi/Toplam Varlıklar

Başarısız işletmelerin finansal başarısızlıkları, geriye dönük olarak 5 yılı kapsayan ve finansal tablolardan elde edilen verilere dayanan bir analizle ortaya konmuştur. Beaver, bu süre zarfında başarılı işletmelerin finansal tablo verileri ile karşılaştırmıştır (Karadeniz ve Öcek, 2020).

Son olarak, Beaver işletmelerin finansal başarısızlıklarını önceden tahmin etmede en etkili oranın işletmelerin nihai olarak nakit akışı ile toplam borçları arasındaki oran olduğunu belirlemiştir. Bu oran, başarılı sonuçlar elde etmiştir (Aktaş, 1991).

2.7.2. Tamari Modeli

Tamari, 1966'daki çalışmasında, 28 iflas etmiş işletmeyle 28 başarılı işletmeyi incelemiş ve finansal zorluk yaşayan işletmelerin finansal oranlarının, beş yıl öncesinin endüstri ortalamalarından farklı olduğunu tespit etmiştir. Bu farkın, iflas dönemine yaklaşıldıkça daha da belirginleştiğini bulmuştur (Kul, 2012). Tamari modeli, subjektif değerlendirmeler içerdiği için kullanımı kolay olsa da güvenilirliği şüpheli olarak kabul edilir. (Yıldız ve Gürkan, 2022).

Tamari, işletmelerin risk durumlarını bir dizi farklı değişken yerine tek bir değişkene dayalı değerleme yöntemine göre daha etkin olduğunu ileri sürerek "risk indeksini" tavsiye etmiştir. Bu indeksi oluşturmak için, genel kabul gören finansal göstergeler arasından altı tanesini

seçmiştir.

Tamari'nin önerdiği finansal rasyolar ve ağırlıkları aşağıdaki gibidir;

Tablo 2.2: Tamari modelinde önerilen finansal rasyolar

Finansal Gösterge	Ağırlık
(Ana Sermaye + Yedekler) / Toplam Borçlar	25
Kar Trendi	25
Cari Oran	20
Üretim Değeri / Stoklar	10
Satışlar/ Kısa Vadeli Alacaklar	10
Üretim Değeri/ Çalışma Sermayesi	10

Tamari'nin tavsiye ettiği katsayılar, oranın en yüksek değerini göstermek için verilecek puanı belirtmektedir. Bu nedenle, diğer şirketlere kıyasla tüm oranlar açısından daha iyi durumda olan bir firmanın endeksi 100 olacaktır.

Buradan çıkacak sonuçlara göre 30 puandan daha düşük puan alan işletmelerin %50'si iflas ederken 30 puandan daha yüksek alan işletmelerin %3'ü iflas etmiştir.

2.7.3. Altman Z-Skor Analizi

Altman'ın çalışmaları, modelin Amerikan piyasalarında ve gelişmekte olan pazarlarda finansal başarısızlığı yüksek bir doğrulukla tahmin ettiğini ortaya koymuştur. Model, geleneksel oran analizinin yerini almış ve akademik alanda geniş kabul görmüştür. Altman'ın modeli, iflas eden ve sağlıklı firmaları içeren bir örnekleme üzerinde geliştirilmiş ve finansal oranlar temel alınarak oluşturulmuştur. Bu modelin uygulanabilirliği, çeşitli sektörlerde ve farklı veri türlerinde gösterilmiştir. Altman Z-Skor modeli, finansal başarısızlık riskini ve potansiyel iflas olasılığını tahmin etme amacıyla kullanılmaktadır (Kulalı, 2016).

Altman Z-Skoru, bir şirketin likidite, kârlılık, öz sermaye karlılığı, finansal kaldıraç ve işletme verimliliği gibi beş farklı finansal oranı kullanır. Bu oranlar, şirketin genel mali

sağlığı hakkında bilgi sağlar ve bir iflas durumunda ne kadar risk altında olduğunu değerlendirir. Altman Z-Skoru, aşağıdaki formülle hesaplanır:

$$Z = 1{,}2X_1 + 1{,}4X_2 + 3{,}3X_3 + 0{,}6X_4 + 1{,}0X_5 \quad \text{(Denklem 1)}$$

Burada:

- X_1 = İşletmenin işletme varlıklarının net çalışma sermayesine oranı (Çalışma Sermayesi / Toplam Varlıklar)
- X_2 = İşletmenin net karlılığının net satışlara oranı (Net Kar / Toplam Varlıklar)
- X_3 = İşletmenin işletme gelirlerinin öz sermayesine oranı (Öz Sermaye / Toplam Varlıklar)
- X_4 = İşletmenin net satışlarının toplam varlıklarına oranı (Net Satışlar / Toplam Varlıklar)
- X_5 = İşletmenin piyasa değerinin borçlarını ödeme kapasitesinin öz sermayeye oranı (Piyasa Değeri / Borçlar)

Elde edilen Z-Skoru, bir şirketin finansal durumu hakkında bilgi sağlar ve genellikle aşağıdaki kategorilere göre yorumlanır (Koç ve Ulucan, 2016):

- 3 ve üstü: Finansal durumu iyi
- 1.8 ile 3 arası: Finansal durumu normal
- 1.8'den düşük: Finansal zorlukla karşılaşma ihtimali yüksek

Altman Z-Skoru, yatırımcılar, kredi verenler ve şirket yöneticileri tarafından şirketin finansal durumunu değerlendirmek için yaygın olarak kullanılan bir araçtır. Ancak, herhangi bir finansal oran gibi, tek başına kesin bir iflas tahmini yapmak için yeterli değildir ve diğer faktörler de dikkate alınmalıdır.

2.7.4. Meyer ve Pifer Modeli

Meyer ve Pifer, Amerika Birleşik Devletleri'nde gerçekleştirdikleri çalışmada çoklu regresyon modelini kullanarak bankaların mali başarısızlığını öngörmek için bu yöntemi uygulamışlardır (Aktaş vd., 2003). Meyer ve Pifer'ın çalışmasında, 1948-1965 yılları arasında ABD'de faaliyetine son veren 39 banka ile benzer niteliklere sahip ancak başarılı olan 39 banka üzerinde bir inceleme yapılmıştır. Bu incelemede, eşleştirme kriterleri olarak aynı şehirde yer alma, benzer büyüklük ve yaş, verilerin aynı tarihte olması esas alınmıştır. Çalışmada, dönem sonu bilanço, gelir tablosu ve banka müfettişlerinin raporlarından elde edilen 32 oran bağımsız değişken olarak kullanılmıştır. Bağımlı değişken olarak ise (0,1) kukla değerlerini alan lineer regresyon fonksiyonu oluşturulmuştur. Araştırma sonuçlarına göre, iflas tarihinden bir ya da iki yıl öncesinde bankaların %80'inin doğru gruplara ayrılabildiği gözlenmiş ve belirlilik katsayısı (R2) olarak %70 civarında bir değer elde edilmiştir. Ancak, modelin tahmin yeteneği geçmiş yıllara doğru azalmıştır. Ayrıca, modeli oluşturan dokuz değişkenden yalnızca biri finansal rasyo iken, diğer sekizi ekonomik trendlere, değişimlere ve beklentilere dayalı oranlardan oluşmaktadır (Kul, 2012).

2.7.5. Deakin Modeli

Deakin'ın 1972'de gerçekleştirdiği çalışmada, 1964-1970 yılları arasında 32 başarılı işletme ile 32 başarısız işletme mali açıdan incelenmiştir. Araştırmada, başarısızlık öncesindeki üç yıl için sınıflandırma hatası oldukça düşük bulunmuştur (Söylemez ve Türkmen, 2017). İşletmelerin mali başarısızlık riskini tahmin etmek için kullanılan bir modeldir. İngiliz ekonomist ve istatistikçi Edward S. Deakin tarafından geliştirilmiştir. Bu model, finansal oranları ve diğer mali verileri kullanarak bir işletmenin finansal sağlığını değerlendirir ve potansiyel olarak mali sıkıntıya düşme riskini belirlemeye çalışır.

2.7.6. Weibel Modeli

Weibel (1973), İsviçre'de banka müşterileri arasında yer alan küçük ölçekli işletmeler üzerinde bir çalışma yapmış ve Weibel modelini ortaya koymuştur (Karadeniz ve Öcek, 2019). Kötü ve iyi işletme çiftlerini seçerken iş kolu, işletme büyüklüğünü, işletmenin yaşını, hukuki yapısını, kuruluş yerini, ekonomik durumu ve taşınmaz mülk sahipliği gibi kriterleri

kullanarak Wilcoxon testi uygulayarak 42 oranı analiz etmiştir (Dizgil, 2018).

2.7.7. Sinkey Modeli

Sinkey (1975) tarafından gerçekleştirilen çalışmada, eşleştirilmiş örnekleme yöntemi kullanılarak banka başarısızlıkları sınıflandırılmıştır. Araştırmanın neticesine göre, modelin iflas tarihinden bir sene öncesine kadar mali başarısızlığı yüzde 80 oranında doğru bir şekilde sınıflandırdığı belirlenmiştir (Tekin ve Gör, 2022).

Sinkey, araştırmasında sorunlu bankaları tanımlamak için bir model geliştirmiştir. FDIC'nin sorunlu olarak belirlediği 62 bankayı, iki gurup şeklinde; sorunlu grup ve sorunsuz grup olarak bu işletmelerin finansal verilerini karşılaştırmış ve işletme faaliyetleri ile finansal davranışları arasındaki farklılıkları incelemiştir (İloğlu, 2020). Sinkey, başarılı ve başarısız işletmeler arasındaki farklılıkları belirlemek için analiz yöntemi olarak çok değişkenli diskriminant analizi yöntemini kullanmıştır (Öcek, 2018). Sonuçlar doğrultusunda, bankaları başarılı veya başarısız olarak sınıflandırılmasında kullanılmıştır.

Araştırmanın sonucunda, modelin iflas olasılığını 6 yıl öncesinden %50 ve 1 yıl öncesinden %80 doğruluk oranıyla tahmin ettiği belirlenmiştir. Ayrıca, geliştirilen modelin bankalar için erken uyarı sistemi olarak avantaj sağlayabileceği şu şekilde sıralanmıştır:

- Bankacılık düzenleyici kuruluşların kaynaklarını daha etkin bir şekilde kullanmalarını teşvik etmek.
- Banka finansal verilerinin daha etkin bir şekilde kullanılmasını sağlamak.
- Belirli objektif kriterlerin belirlenmesini kolaylaştırmak.
- Denetim ve inceleme performanslarını değerlendirmek için düzenleyici kurumların imkanlarını artırmak.
- Mevduat sigorta primlerinin belirlenmesinde temel oluşturmak için yardımcı olmak.

2.7.8. Daniel Martin Modeli

Daniel Martin, çalışmasında, finans sektöründeki 5.600 işletmeyi analiz etmiş ve bankalar

için bir tahmin ve erken uyarı modeli oluşturmuştur. Çalışmasını 1970-1976 yıllarında gerçekleştirmiştir. Martin, çalışmasında, finansal oranları 25 bağımsız değişken olacak şekilde kullanmış ve işletmelerin iflasını iki yıl öncesine kadar tahmin etmeyi başarmıştır. Ayrıca, başarılı olan işletmeleri %91,3 oranında, başarısız olan işletmeleri ise %91,1 oranında doğru bir şekilde sınıflandırmıştır (Ural, 2014).

2.7.9. Springate Modeli

Gordon L.V. Springate, 1978 yılında yapmış olduğu çalışmada, kademeli çok değişkenli diskriminant analizi kullanarak bir model ortaya koymuştur. Kanada'da faaliyet gösteren imalat işletmelerinde gerçekleştirdiği araştırmada, işletmeleri başarılı ve başarısız olarak ayırt etmek için dört finansal rasyoyu temel almıştır (Öcek, 2018).

$$S = 1{,}03A + 3{,}07B + 0{,}66C + 0{,}4D \qquad \text{(Denklem 2)}$$

A = Çalışma Sermayesi / Toplam Varlıklar Oranı
B = Faiz ve Vergiden Önceki Kar / Toplam Varlıklar Oranı
C = Faiz ve Vergiden Önceki Kar / Kısa Vadeli Borçlar Oranı
D = Satışlar / Toplam Varlıklar Oranı

Yukarıdaki modelde S skor değeri 0,862'den küçük ise işletme başarısız sayılmaktadır.

2.7.10. Ohlson Skor Modeli

Ohlson, çalışmasında logit modelini kullanarak, 1970-1976 yılları arasındaki finansal verilerden yararlanarak 105 başarısız ve 2058 başarılı işletmenin durumunu incelemiştir. Bu modelin diğerlerinden farklılığı, başarılı ve başarısız işletme sayısının eşit olmamasıdır. Ayrıca, genellikle Moodys gibi kurumlardan sağlanan verilerin aksine, bu çalışmada ABD borsalarında faaliyet gösteren işletmelerin finansal performansını yansıtan 10-K mali durum raporları kullanılmıştır. Yapılan analiz, iflas olasılığını bir yıl öncesi için %96,12, iki yıl öncesi için ise %95,55 oranında doğru bir şekilde tahmin etmiştir (Çelik ve Dursun, 2021).

2.7.11. Fulmer H Skor Modeli

Fulmer tarafından geliştirilen H Skor modeli, işletmelerin finansal başarılarını veya başarısızlıklarını tahmin etmek amacıyla kullanılan bir analiz aracıdır. Model, işletmelerin mali durumunu değerlendirirken belirli bir aralık kullanır ve bu aralık, işletmenin başarılı ya da başarısız olarak sınıflandırılmasına yardımcı olur (Medetoğlu ve Tutar, 2023).

Araştırmada, ortalama aktif büyüklüğü 455 milyon dolar olan 60 işletmenin verileri analiz edilmiştir. Bu 60 işletmenin yarısı finansal olarak başarılı, diğer yarısı ise başarısız olarak sınıflandırılmıştır. H Skor modelinin temel prensibi, hesaplanan H Skoru'nun pozitif ya da negatif olmasına dayanır. Eğer H Skoru 0'dan küçükse, işletme başarısız olarak kabul edilir. H Skoru 0'dan büyükse, işletme başarılı olarak kabul edilir.

Bu sınıflandırma yöntemi, işletmelerin mali durumunu değerlendirmek ve potansiyel sorunları veya başarılı stratejileri belirlemek için kullanışlı bir araçtır.

Model;

$$H = (5{,}528 \times Y_1) + (0{,}212 \times Y_2) + (0{,}073 \times Y_3) + (1{,}270 \times Y_4) - (0{,}12 \times Y_5) + (2{,}335 \times Y_6) + (0{,}575 \times Y_7) + (1{,}083 \times Y_8) + (0{,}894 \times Y_9) - (6{,}075) \qquad \text{(Denklem 3)}$$

Denklemdeki değişkenler;

Y_1 = Dağıtılmamış Kâr / Toplam Varlıklar
Y_2 = Satışlar / Toplam Varlıklar
Y_3 = Vergi Öncesi Kâr / Öz Sermaye
Y_4 = Nakit Tutarı / Toplam Borçlar
Y_5 = Toplam Borçlar / Toplam Varlıklar
Y_6 = Kısa Vadeli Yabancı Kaynaklar / Toplam Varlıklar
Y_7 = Maddi Duran Varlıklar
Y_8 = Çalışma Sermayesi / Toplam Borçlar
Y_9 = Faiz ve Vergi Öncesi Kâr / Faiz

2.7.12. Diskriminant Analizi

Bu yöntemi uygularken, başlangıçta analize dahil edilecek işletmeler finansal açıdan başarılı olanlar ve başarısız olanlar olarak iki ana gruba ayrılır. Bu ayrım, belirlenen mali başarısızlık kriterine dayanır. Ardından, analiz için kullanılacak finansal oranlar hesaplanır. İşletmelerin mali başarısızlık durumu bağımlı değişken olarak belirlenirken, finansal oranlar bağımsız değişkenlerdir. Diskriminant analizi sonucunda, belirlenen anlamlılık seviyesinde (%1, %5, %10 gibi), gruplar arasında farklılık gösteren finansal oranlar saptanır. Diskriminant analizinin sınıflandırma fonksiyonu kullanılarak, modelin doğru sınıflandırma performansı ölçülebilir (Selimoğlu ve Orhan, 2015).

2.8. Finansal Sıkıntı, Ödeme Güçlüğü ve İflas Arasındaki Farklar

Finansal sıkıntı, mevcut harcamaları ve borçları karşılamak için yeterli para akışının olmaması durumu olarak ifade edilebilir (Coşkun ve Sayılgan, 2008). Bu durum genellikle artan borç, gelir kaybı, kötü finansal yönetim veya ekonomik durgunluk gibi nedenlerle ortaya çıkar. Finansal sıkıntı, zamanında müdahale edilmezse iflas veya tasfiye gibi ciddi sonuçlara yol açabilir.

Ödeme güçlüğü, şirketlerin mali yükümlülüklerini zamanında ve tam olarak karşılayamama durumudur.

Finansal sıkıntının sonucu olan iflas, borçların ödenememesi durumunda uygulanan yasal süreci ifade eder (Bilir, 2015). Finansal sıkıntı ile iflas arasındaki fark; iflas finansal sıkıntının sonucunda ortaya çıkar (Kuızınıenė vd., 2022).

Bir işletmenin finansal sıkıntıya girmesi sonucunda ödeme güçlüğü ortaya çıkar. Ödeme güçlüğü nakit akışı sorunları, gelir kaybı, yüksek borç yükü veya beklenmedik mali krizler gibi nedenlerden kaynaklanır ve faturaların gecikmesi, kredi notunun düşmesi, faiz cezaları gibi sorunlara yol açabilir sonrasında ciddiyetine bağlı olarak temerrüt veya iflasla sonuçlanabilir.

2.9. Literatür Taraması

Bir araştırma, İstanbul Menkul Kıymetler Borsası'nda (İMKB) işlem gören gıda sektörü şirketlerinin finansal başarısızlık riskini belirlemek için güvenilir bir model geliştirmeyi hedeflemiştir. Bu çalışmada, şirketlerin finansal performansını değerlendirmek için Altman Z Skoru kullanılmıştır. Bu kriter, şirketlerin mali durumunu ölçmek ve mali başarısızlık riskini belirlemek için yaygın olarak kullanılan bir yöntemdir. Analizler sonucunda, gıda sektöründe faaliyet gösteren işletmelerin finansal başarısını belirleyen önemli faktörlerin aktif karlılık oranı ve borç-özkaynak oranı olduğu sonucuna varılmıştır. Aktif karlılık oranı, şirketin varlıklarını ne kadar etkin kullandığını gösterirken, borç-özkaynak oranı, şirketin finansal kaldıraç seviyesini ve borçluluğunu ölçer. Bu iki oranın finansal başarısızlık riskini öngörmede kritik olduğu tespit edilmiştir. (Terzi, 2011).

(Büyükarıkan ve Büyükarıkan, 2014) Borsa İstanbul'da işlem gören bilişim sektöründeki şirketlerin finansal performanslarını Altman Z-Score ve Springate modelleriyle incelemişlerdir. 2008-2013 dönemleri arasındaki altı hesap döneminin konsolide mali tablolarındaki verilere dayanarak bu analizi gerçekleştirmişlerdir. Elde edilen sonuçları yorumlarken Altman Z-Score ve Springate finansal başarısızlık modellerinden gelen verileri dikkate almışlardır. Ardından, modellerdeki bileşenler arasındaki istatistiksel ilişkiyi belirlemek için regresyon analizi ve ANOVA testi uygulanmış, bu bileşenler arasındaki ilişkiyi daha iyi anlamak için korelasyon analizi yapılmıştır. Bilişim sektöründe faaliyet gösteren firmaların Altman Z-Score ve Springate modellerinden elde edilen verilere dayanarak, her iki modelin de finansal başarısızlığı belirlemede benzer sonuçlar verdiği tespit edilmiştir. Ancak, iflas öngörüsünün gerçekleşmemesi, bu işletmelerin finansal risk taşımadığı anlamına gelmemelidir.

(Bağcı ve Sağlam, 2020) Bir araştırma, Borsa İstanbul'da işlem gören sağlık ve spor sektörlerindeki işletmelerin finansal başarısızlık risklerini incelemiştir. Araştırmada, 2014-2018 yılları arasındaki beş yıllık dönemde toplam altı işletme incelenmiştir: dört spor işletmesi ve iki sağlık işletmesi. Finansal başarısızlık riskini tahmin etmek için Altman, Springate ve Fulmer modelleri kullanılmıştır. Araştırmanın sonuçlarına göre, sağlık işletmelerinin finansal yapısı sağlamdır ve performansları olumlu bir seyir izlemektedir. Bu nedenle, sağlık işletmelerinin iflas riskiyle karşılaşma olasılığı düşüktür. Öte yandan, spor

işletmelerinin finansal yapısının zayıf olduğu ve bu nedenle finansal başarısızlık riski taşıdığı belirlenmiştir. Araştırma, spor işletmelerinin her an iflas tehlikesiyle karşılaşabilecekleri sonucuna ulaşmıştır.

Bir araştırmada, Borsa İstanbul'daki (BİST) imalat sektöründe şirketlerin finansal başarısızlığını tahmin etmek için yapay zekâ yöntemleri incelenmiştir. Araştırmada, şirketlerin 12 aylık finansal tablolarından ve raporlarından elde edilen veriler kullanılmıştır. Çalışmada kullanılan 26 finansal oran, dört ana başlık altında sınıflandırılmıştır: likidite oranları, finansal yapı oranları, faaliyet oranları ve karlılık oranları. Finansal başarısızlık kriterleri olarak, şirketlerin iki yıl üst üste zarar etmesi, BİST gözaltı pazarında bulunması, iflas etmesi, faaliyetlerin durdurulması ve toplam aktiflerde %10'luk bir kayıp yaşaması gibi durumlar dikkate alınmıştır. Araştırma sonucunda, çeşitli yapay zekâ modelleri arasından yapay sinir ağının, destek vektör makinesine göre daha iyi performans gösterdiği belirlenmiştir. Bu sonuç, yapay sinir ağlarının finansal başarısızlık tahmininde güçlü bir araç olduğunu gösterir. Özellikle karmaşık veri yapıları ve ilişki ağı bulunan durumlarda, yapay sinir ağları daha etkin sonuçlar verebilir. Bu çalışma, finansal başarısızlık riskinin daha iyi tahmin edilmesine yönelik yapay zekâ uygulamalarının önemini vurgulamaktadır (Yürük ve Ekşi, 2019).

(Yıldız, 2022) Finansal Teknoloji kavramı çerçevesinde yapılan çalışmaları ampirik ve ampirik olmayan çalışmalar şeklinde sınıflandırarak incelemiştir. Bu çalışmaların mevcut durumu belirlemesi ve gelecekteki değişimleri tahmin etmesi hedeflenmiştir. Bu amaçla, Türkçe olarak yazılmış ve Google Akademik, Dergipark ve YÖK veri tabanlarında bulunan çalışmalar üzerinde bir literatür taraması yapılmıştır. Ampirik çalışmalar, yıl, sektör, yatırım aracı ve kullanılan teknik gibi betimsel istatistiksel analizlerle yapılmıştır. Ampirik olmayan çalışmalardan elde edilen bilgilerle genel sonuçlar çıkarılmıştır. Yapılan incelemelere göre, çalışmaların çoğunun hisse senetleri, altın vb. yatırım araçlarının tahminine odaklandığı ve yapay sinir ağları tekniğinin sıkça kullanıldığı görülmektedir. Ancak, son zamanlarda blokzincir ve bitcoin fiyat tahmini gibi daha ileri düzey analizlerde derin öğrenme gibi tekniklerin kullanımında hızlı bir artış gözlenmiştir.

(Altınöz, 2013) banka başarısızlıklarını önceden belirlemek için çoklu istatistiksel yöntemlerden biri olan yapay sinir ağları modeli incelemiştir. Bu yöntem, yapay zekâ ve

makine öğrenimi tekniklerinin finansal istatistiklerde nasıl kullanılabileceğini gösterir. Araştırmada, yapay sinir ağları modelinin hem bir yıl öncesi hem de iki yıl öncesi için banka başarısızlıklarını başarılı bir şekilde tahmin edebildiği gözlemlenmiştir. Bu sonuç, yapay sinir ağlarının karmaşık finansal verilerdeki kalıpları öğrenme ve gelecekteki sonuçları tahmin etme konusundaki potansiyelini göstermektedir. Yapay sinir ağları, istatistiksel modelleme alanında literatürdeki diğer çalışmalarla uyumlu olarak test edilmiştir ve özellikle finansal başarısızlık tahmini gibi alanlarda etkili olduğu kanıtlanmıştır. Bu bulgu, bankaların risk yönetimi ve önceden önlem alma stratejileri geliştirmesi için yapay zekâ temelli modellerin değerli olabileceğini göstermektedir.

Yapay sinir ağlarının finansal başarısızlığı tahmin edilmesini ele alan makalede (Akkaya, Demireli, ve Yakut, 2009) Tekstil ve Kimya, Petrol ve Plastik sektörlerinde faaliyet gösteren işletmelerin mali başarısızlıklarını bir yıl öncesinden tahmin etmeye odaklanılmıştır. Araştırmanın sonucuna göre, YSA modeli, 11 başarılı işletmeden 9'unu doğru bir şekilde sınıflandırmıştır. Bu, modelin başarılı işletmelerin yaklaşık %82'sini doğru tahmin ettiği anlamına gelir. Ayrıca, test setinde yer alan 10 başarısız işletmeden %80 oranında doğru sınıflandırma yapılmıştır. Bu sonuç, 10 başarısız işletmeden 8'inin doğru olarak sınıflandırıldığını, 2'sinin ise yanlış sınıflandırıldığını gösterir. Genel sınıflandırma doğruluğunu incelediğimizde, eğitim setindeki tüm örneklerin doğru sınıflandırıldığı, test setinin toplam sınıflandırma doğruluğunun ise yaklaşık %81 olduğu tespit edilmiştir.

(Çelik, 2010) yapmış olduğu çalışmada Türkiye'deki bankaların finansal başarısızlıklarını öngörmek için erken uyarı modelleri oluşturmayı amaçlamıştır. Bu kapsamda, yaygın olarak kullanılan Diskriminant Analizi ve Yapay Sinir Ağları modellerinin tahmin güçleri karşılaştırılmıştır. Araştırma, 36 özel sermayeli ticaret bankasına ait finansal oranları kullanarak çalışmaya konu olan işletmelerin mali başarısızlığa düşme ihtimalini 1 ve 2 yıl öncesinden tahmin etmeye çalışmıştır.

Araştırmanın sonuçlarına göre:

- Bir yıl öncesi için en başarılı model, %100 genel başarı oranı ile yapay sinir ağı modelidir. Bu model, hem başarılı hem de başarısız bankaları en doğru şekilde tahmin etmiştir.

- İki yıl öncesi için, başarılı bankalar için en iyi model %88,9 başarı oranı ile diskriminant analizi olmuştur. Başarısız bankalar için ise %100 başarı oranı ile yapay sinir ağı modeli en başarılı olmuştur. İki yıl öncesi için genel başarı ortalaması ise %91,7 ile diskriminant analizi modelidir.

Bu sonuçlar, yapay sinir ağlarının finansal başarısızlık tahmininde güçlü bir araç olduğunu ve özellikle bankacılık sektöründe erken uyarı sistemleri oluşturmak için kullanılabileceğini göstermektedir. Aynı zamanda, diskriminant analizi de belirli durumlarda etkili bir tahmin yöntemi olarak karşımıza çıkmaktadır.

(Akpınar ve Akpınar, 2017) mali başarısızlık riskinin belirleyicilerinin tespit etmek için bir araştırma yapmışlar. Borsa İstanbul'da işlem gören 82 imalat işletmesini 2010-2014 periyodunda incelemişler. Altman Z skor modelinde kullanılan bağımsız değişkenlerle mali başarısızlık riski arasında anlamlı bir ilişkiye rastlamışlardır. Araştırma sonucunda ortaya çıkardıkları sonuçlara göre entelektüel sermaye Z skor üzerinde pozitif bir etki yaratmaktadır.

(Aktümsek ve Göker, 2018) çalışmalarında farklı sektörlerdeki işletmelerin finansal başarısızlık tahminlerinde kullanılacak değişkenlerin sektöre göre değişiklik gösterebileceğini araştırmışlardır. Borsa İstanbul'da işlem gören 3 farklı sektördeki şirketlerin 2008-2017 periyodundaki finansal verilerini kullanmışlardır. Lojistik regresyon analizi ile yaptıkları çalışmada her yılın finansal verilerini bir önceki yıl ile kıyaslayarak finansal başarısızlık öngörüsü yapmışlardır. Araştırma sonucunda seçtikleri 3 sektörde de finansal başarısızlığın habercisi olan oran farklı çıkmıştır.

Yapılan bir araştırmada kriz dönemlerinde işletmelerin finansal başarısızlık nedenleri incelenmiştir. Bir işletmede kriz zamanlarında, yönetim kalitesi, kaynak temini ve kullanımındaki verimlilik ve kaliteyi korurken maliyetleri en aza indirme gibi faktörlerin finansal başarıyı etkilediği ortaya konmuştur (Karacan ve Savcı, 2011).

(Salur, 2021) Borsa İstanbul'da işlem gören şirketlerin mali tablolarındaki verileri kullanarak mali durumlarını tahmin eden ve tahmin gücünü ölçen bir model geliştirmek istemiştir. Bu modelde yapay sinir ağlarından faydalanmıştır. Araştırmada, işletmelerdeki

finansal başarısızlığın bir yıl önceden tahmin edilmesi için geliştirilen yapay sinir ağı modelinin test setindeki 24 iflas eden işletmenin 22'sini ve 24 başarılı işletmenin tamamını doğru bir şekilde sınıflandırdığı belirlenmiştir. Model, başarılı işletmeleri %100 doğrulukla sınıflandırırken, başarısız olanlardan yalnızca 2'sini başarılı olarak etiketlemiştir.
Araştırma sonuçlarına göre, yapay sinir ağlarının işletmelerdeki finansal başarısızlığın tahmininde yüksek doğrulukla sınıflandırma yaparak etkili bir model olduğu ortaya konmuştur.

Türkiye'de yapılan bir çalışmada 2008-2010 yılları arasında İMKB'de işlem gören 130 sanayi şirketinin mali başarısızlık durumu finansal verileri kullanılarak analiz edilmiştir. Şirketler mali başarısızlık açısından işlem sırasının kapatılması ve 3 sene art arda zarar etmiş olanlara göre seçilmiştir. Altman Z Skor modelinde kullanılan oranlar bağımsız değişken olarak kullanılmıştır. Çalışmanın sonucunda şirketlerin mali başarısızlık durumu 1, 2 ve 3 yıl önceden tahmin edilmiştir (Akgün, 2013).

(Sakız ve Ünkaya, 2018) yaptıkları çalışmada Borsa İstanbul'da işlem gören Türk Hava Yolları Anonim Ortaklığı şirketinin iflas riskini havacılık sektörüne özgü Airscore modelini ve Yapay Sinir Ağlarını kullanarak öngörmeyi amaçlamışlardır. Airscore modeline göre tüm değerler 0,03'ün üzerinde olduğundan finansal başarı açısından olumlu olduğu tespit edilmiştir. Yapay Sinir Ağları kullanılarak geleceğe dönük yapılan tahminlerde de ileriki 3 sene içerisinde bu değerlerin 0,03'ün üzerinde olacağı öngörülmüştür.

Sağlık işletmelerinin finansal başarısızlık tahmini üzerine yapılan çalışmada 2008-2012 yılları arasında Zonguldak ilinde bulunan sağlık işletmelerinin mali başarısızlıklarının ölçülmesi amaçlanmıştır. Toplam 10 adet kamu sağlık işletmesinin finansal verileri Altman Z Skor modeli ve Yapay Sinir Ağları ile analiz edilmiştir. Sonuçlara göre Yapay Sinir Ağları modeli finansal başarısızlık tahmininde başarılı tahminlerde bulunduğu ortaya konulmuştur (Civan ve Dayı, 2014).

Geleneksel ve modern tahmin yöntemlerini kullanarak mali başarısızlık öngörüsü yapılan bir araştırmada 1997-2017 yılları arasında Borsa İstanbul'da işlem gören imalat sektöründeki bir şirketin Mali başarısızlıklarının 3 ay önceden tespit edilmesi amaçlanmıştır. Analizlerde işletmenin finansal rasyoları kullanılmıştır. Analizler sonucunda en iyi tahmin

yapan modeller, sırasıyla Yapay Sinir Ağı, Rastgele Ormanlar ve Karar Ağaçları olup, başarısız olan tahmin modeli ise Lojistik Regresyon modeli olduğu ortaya çıkmıştır (Selçik, 2019).

(Anandarajan vd., 2001) yaptıkları çalışmada iflası tahmin etmek için yapay sinir ağlarını kullanarak iki farklı teknik geliştirmişlerdir. Modellerin tahmin kabiliyetleri, çoklu ayrım analizi (MDA) modeline kıyasla gerçek dünya örnekleri kullanılarak test edildi. Sonuçlar, genel olarak yapay sinir ağı (ANN) modellerinin, MDA modellerine kıyasla daha yüksek tahmin doğruluğuna sahip olduğunu ortaya koydu.

Yapılan bir araştırma, bankaların başarısızlığını tahmin etmeye yönelik geleneksel istatistik yöntemleri ile makine öğrenimi yöntemlerinin doğruluğu karşılaştırılmaktadır. İncelemeye 3000 Amerikan bankasından oluşan bir örneklem dahil edilmiştir (1438'i başarısız olmuş ve 1562'si halen aktif olan bankalar). İki geleneksel istatistiksel yöntem (Ayrım Analizi ve Lojistik Regresyon) ve üç makine öğrenimi yöntemi (Yapay Sinir Ağı, Destek Vektör Makineleri ve En Yakın K-Komşu) kullanılarak analizler gerçekleştirilmiştir. Her bir banka için, bankalar başarısız olmadan önceki 5 yıllık döneme ait veriler toplanmıştır. Bankaların finansal raporlarından elde edilen 31 oran, kredi kalitesi, sermaye kalitesi, operasyonel verimlilik, kârlılık ve likidite gibi 5 temel alanı kapsamaktadır. Deneysel sonuçlar, yapay sinir ağı ve en yakın k-komşu yaklaşımlarının en yüksek doğruluk oranına sahip olduğunu göstermiştir (Lee ve Viviani, 2018).

Finansal krizleri yapay sinir ağları ile tahmin etmeyi amaçlayan bir çalışmada finansal krizleri etkileyen faktörler, finansal kriz teorilerinden elde edilen ekonomik temeller, gerçek döviz kuru rejimi ve bulaşma etkilerine odaklanılarak değerlendirilmiştir. 1990 ile 2014 yılları arasında Türk ekonomisinin 7 temel makroekonomik ve finansal göstergesinin aylık verileri kullanılarak, yapay sinir ağlarının tahmin gücünün oldukça etkileyici olduğu ortaya konmuştur (Aydın ve Çavdar, 2015).

Çin'de yapılan bir çalışmada halka açık şirketlerin 2007-2020 dönemine ait finansal verileri kullanılarak finansal sıkıntı tahmin modellerinin doğruluğunu ve açıklanabilir olmasını artırmak için bir yapay zeka yaklaşımı önerilmektedir. Çalışma, özellik seçimi ve tahminci inşasını içeren bir topluluk yöntemi ile yerel ve küresel açıklamaları sağlayan bir yorumlama

çerçevesi tasarlamıştır. Araştırmada, çeşitli temel tahminciler ve entegre modeller karşılaştırılmış, LightGBM modelinin %92 ile en yüksek doğruluk oranını sağladığı tespit edilmiştir (Zhang vd., 2022).

(Kasztelnik, 2020) yaptığı araştırmada 2008 Büyük Ekonomik Durgunluğunun ABD'deki ulusal bankaların sermaye yatırımlarının değerlemesi üzerindeki etkisini incelemek ve bankaların finansal başarısızlığını tahmin etmek için bir model oluşturmayı amaçlamaktadır. Bu doğrultuda 2009-2012 döneminde ulusal bankaların fiyat/kazanç oranı ve temettü getirisi ile öz sermaye getirisi arasındaki ilişkiyi değerlendirmiştir. İstatistiksel modelleme ve makine öğrenimi teknikleri kullanılarak bazı bulgular elde edilmiştir:

- 2012'deki ortalama hisse başına kazanç ve 2009'daki temettü getirisi, 2009 ve 2012'deki ortalama öz sermaye getirisinden daha büyük bulunmuştur.
- 2012'deki temettü getirisi, aynı yılın ortalama öz sermaye getirisinden daha küçük çıkmıştır.

Yapılan bir çalışma, finansal faktörlere dayalı iflas tahminini sınıflandırmak için yeni bir makine öğrenimi algoritması olan CatBoost'u kullanmayı öneriyor. Fransa'daki 2014-2016 dönemine ait kurumsal başarısızlık verilerini inceleyen çalışma, CatBoost'un sekiz diğer yaygın algoritma ile karşılaştırıldığında iflas tahmininde daha iyi performans gösterdiğini belirledi. Çalışma ayrıca, iflası öngörmek için önemli olan finansal değişkenleri belirlemek amacıyla kısmi bağımlılık grafiklerini kullandı. Araştırma, önerilen yaklaşımın bankalar ve yatırımcılar için finansal sıkıntıyı erken tespit etmeye yardımcı olabileceğini ve şirketlerde kontrolü iyileştirebileceğini gösterdi. Çalışma, CatBoost'un iflası tahmin etmek için etkili bir araç olduğunu ortaya koyarken, banka ve alacaklılara kendi modellerini karşılaştırma ve değerlendirme imkânı sunuyor. Ayrıca, CatBoost'un başarılı ve başarısız firmaları ayırt etmek için de kullanılabileceğini ve böylece bankaların daha iyi kararlar almasına yardımcı olabileceğini gösteriyor (Jabeur vd., 2021).

3. YAPAY ZEKÂ

Yapay zekâ terimi, ilk kez 1956 yılında John McCarthy tarafından bir akademik konferansta kullanılarak literatüre girmiştir. Vannevar Bush'un "As We May Think" adlı çalışması, yapay zekâ kavramının öncüllerinden biri olarak kabul edilir. 1945'te yayımlanan bu makale, insan zekâ ve algısını güçlendiren yardımcı sistemleri öngörerek, bilgi işleme konusunda önemli bir vizyon sunmuştur. Alan Turing'in çalışmaları ise yapay zekâ alanında büyük bir etki yaratmıştır. Turing, makinelerin insan zekâsını taklit edebileceği ve insan zekâsı gerektiren oyunlarda kararlar alabileceği fikrini ortaya atmıştır. Özellikle "Turing testi" adı verilen insan-makine etkileşimini değerlendiren bir metot geliştirmesiyle bilinir (Arda ve Küçükkocaoğlu, 2021).

Yapay zekâ, bilgisayar sistemlerine insana benzeyen zekâ özelliklerini kazandırmayı amaçlayan bir teknoloji alanıdır. Karmaşık problemleri çözebilme, öğrenme yeteneği, karar alma, dil işleme, algılama, öngörü ve adaptasyon gibi insana özgü yetenekleri bilgisayarlara kazandırmak amacıyla geliştirilmiştir. Bu teknoloji, Endüstri 4.0 ve dijital dönüşüm gibi Çağdaş endüstriyel ve toplumsal felsefelerin temelinde yer alır (Gacar, 2019). Bilgisayarların, veri analizi, örüntü tanıma, otomatik karar verme gibi karmaşık görevleri yerine getirebilme kabiliyeti, yapay zekânın ilerlemesiyle birlikte giderek artmaktadır. Bu, endüstriyel süreçlerden sağlık sektörüne, finanstan ulaşıma kadar birçok alanda insan yaşamını ve işleyişini büyük ölçüde etkilemektedir. Yapay zekâ, bilgisayarların sadece veri işleme araçları olmaktan çıkıp, düşünen ve öğrenen birer araç haline gelmelerine olanak tanımaktadır.

Yapay zekâ, bilgisayar bilimlerinde önemli bir alt dal olarak kabul edilir ve zaman içinde birçok alt dalı ortaya çıkmıştır. Bunlar arasında arama algoritmaları, makine öğrenmesi, doğal dil işleme gibi birçok farklı alan bulunmaktadır.

Yapay zekânın kullanımı, bazen insanların beklentileriyle çelişebilir. Genellikle filmlerde gördüğümüz gibi konuşabilen makineler veya uzay görevleri yapabilen robotlar yerine, günlük hayatta farklı ve fark edilmeyen işlevlerde kullanılabilir. Örneğin, kişiye özel pazarlama, öneri sistemleri, arama motorları gibi alanlarda yapay zekâ yoğun bir şekilde kullanılmaktadır. Bu tür sistemler, genellikle insanlar tarafından fark edilmeden veri analizi

yapar ve belirli işlevleri gerçekleştirir.
Yapay zekâ, günümüzde ve gelecekte pek çok farklı alanda kullanılabilir ve gelişmeye devam edecektir. Her geçen gün yeni uygulamalar ve teknolojilerle birlikte insan yaşamına daha fazla entegre olmaktadır.

Yapay zekâ 4 kısımda incelenmektedir (Gacar, 2019);

- Tip 1: Tip 1 yapay zekâ uygulamaları, reaktif makineler olarak bilinir. Bu yapay zekâlar, çevresel değişikliklere tepki göstererek ve anlık verilere dayanarak kararlar alırlar. Öğrenme ve bellek kapasiteleri olmadığından, geçmiş deneyimler gelecekteki hareketlerini etkilemez. IBM firmasının yapmış olduğu satranç oyunu Deep Blue tip 1 yapay zekâ örneğidir.
- Tip 2: Tip 2 yapay zekâ uygulamaları, sınırlı belleğe sahiptir. Bu kategoriye örnek olarak, otomobillerde sıkça görülen şerit değişimlerinde sürücüyü uyarabilen uygulamalar gösterilebilir. Bu tür sistemler, belirli bir veri setini hafızalarında tutar ve bu veri seti üzerinden sınırlı bir öğrenme ve karar alma yeteneğine sahiptirler. Genellikle belirli bir amaca yönelik olarak tasarlanmışlardır ve geniş ölçekli veri setleri üzerinden karmaşık kararlar alamazlar.
- Tip 3: Zihin teorisi olarak bilinmektedir. Henüz bu tip 3 yapay zekâ uygulamalarına rastlanılmamaktadır.
- Tip 4: Bu tür yapay zekâ, kişisel bilgiye dayalı olup, makinelerin spesifik bir bilinçleri olduğu düşünülen türdedir. Bununla birlikte, bu tip yapay zekâ uygulamaları henüz gerçekleşmemiştir.

En çok kullanılan yapay zekâ teknolojileri uzman sistemler, yapay sinir ağları, genetik algoritmalar, bulanık mantık ve zeki etmenlerdir.

Yapay zekâ teknolojisinin en büyük gelişmeleri arasında ses tanıma, veri işleme hızındaki artış, bulut depolama ve işleme kapasitesindeki gelişmeler, makine öğrenimi ve görüntü tarama teknolojilerindeki ilerlemeler yer almaktadır (Tas ve Mert, 2019).

Yapay zekanın çok çeşitli uygulama alanları bulunmaktadır.;

- Sağlık: Tanı koyma, hasta yönetimi ve tıbbi görüntüleme.
- Finans: Risk analizi, ticaret algoritmaları ve dolandırıcılık tespiti.
- Otomasyon: Üretim süreçlerinde robotik ve otomatikleştirilmiş sistemler.
- Müşteri Hizmetleri: Sohbet botları ve sanal asistanlar.
- Eğlence: Oyun tasarımı ve içerik önerileri.

Yapay zekanın etik ve toplumsal sorunlara çözümleri;

- Gizlilik ve Güvenlik: Yapay zekâ sistemlerinin kişisel verileri nasıl kullandığı ve koruduğu.
- İş Gücü: Otomasyonun iş piyasası üzerindeki etkileri.
- Yanlılık ve Adalet: Yapay zekânın, öğrenme süreçlerinde veri yanlılığından kaynaklanan adaletsizlikleri çoğaltması riski.

Yapay zekâ ile ilgili teknolojik gelişmeler ve gelecek beklentileri;

- Makine Öğrenimi ve Derin Öğrenme: Yapay zekânın temelini oluşturan teknolojiler.
- Yapay Zekâ Güvenliği ve Riskleri: Yapay zekânın yanlış ellerde kullanılma olasılığı.
- Yapay Zekâ ve İstihbarat: Güvenlik ve istihbarat alanında yapay zekâ uygulamaları.

Yapay zekâ hakkında akademik ve endüstri perspektifler;

- Araştırma ve Geliştirme: Üniversiteler ve araştırma laboratuvarlarında yapay zekâ üzerine çalışmalar.
- Endüstri Uygulamaları: Teknoloji şirketlerinin ve girişimcilerin yapay zekâ projeleri.
- Hükümet ve Düzenleme: Yapay zekâ teknolojileri için yasal ve etik düzenlemeler.

Yapay zekâ finans sektörü için ele alınırsa üç temel bölümce incelenebilir; bunlar sermaye piyasaları, bankacılık ve sigortacılıktır (Kozan, 2023).

Yapay zekânın finanstaki bazı kullanım alanları;

- Hisse Senedi Fiyat Öngörüleri
- Finansal Başarısızlık Öngörüleri
- Para Krizlerinin Tahmini
- Finansal Teknoloji
- Bütçe Planlama
- Bankacılık
- Sigortacılık
- Risk Yönetimi

Finans alanında kullanımı giderek artan bireysel ve kurumsal danışmanlık sistemlerinde, yapay zekâ teknolojileri aktif olarak kullanılmaktadır. Yapay zekâ asistanları, büyük miktardaki veriyi izleyerek müşterilere yatırım önerileri sunacak şekilde programlanmıştır. Ayrıca, organizasyonel risklerin azaltılmasına yönelik programlar, portföy oluşturma ve takip sistemleri gibi uygulamalar da giderek daha fazla yapay zekâ tarafından desteklenmektedir (Sayım, 2022).

Yapay zekâ teknolojisi, farklı alanlarda uygulanır. Örneğin, makine öğrenimi, sistemlerin verilerden otomatik olarak öğrenmesini sağlayarak, büyük miktarda bilgiyi hızlı bir şekilde analiz eder ve tahminlerde bulunur. Derin öğrenme ise, yapay sinir ağlarını kullanarak daha karmaşık ve katmanlı veri analizini mümkün kılar. Bu teknolojiler, otonom araçlardan sağlık teşhis sistemlerine, müşteri hizmeti sohbet botlarından öneri sistemlerine kadar birçok uygulamada yer bulur.

Yapay zekâ, topluma ve iş dünyasına birçok fayda sağlarken, beraberinde etik ve güvenlik sorunlarını da getirir. Gizlilik ihlalleri, yapay zekâ sistemlerinde yanlılık ve işgücündeki otomasyonun sosyal etkileri gibi konular, yapay zekâ teknolojilerinin sorumlu ve etik bir şekilde kullanılmasını sağlamak için dikkatle ele alınması gereken alanlardır. Ayrıca, yapay zekânın hızlı gelişimi, toplumun bu teknolojiyi anlama ve düzenleme konusunda çaba göstermesini gerektirir.

Yapay zekâ, gelecekte büyük değişikliklere yol açabilecek potansiyele sahiptir. Ancak bu potansiyelin farkında olarak, bu teknolojiyi insanların yaşamını iyileştirmek için kullanırken, olası riskleri ve etik sorunları da göz önünde bulundurmalıyız. Yaratıcı

uygulamaları, iş dünyasında verimliliği artırabilir ve insan yaşam kalitesini yükseltebilir, ancak yapay zekâ kullanımında dikkatli ve etik bir yaklaşım benimsemek de aynı derecede önemlidir.

3.1. Makine Öğrenmesi

Makine öğrenimi, bilgisayar sistemlerinin, gözlem ve gerçek dünya etkileşimi verilerini kullanarak, insan benzeri öğrenme ve davranış sergileyebilme yeteneğini kazanmalarını ve zaman içinde performanslarını geliştirebilmelerini sağlayan bir bilimdir (Arda ve Küçükkocaoğlu, 2021). Başka bir ifadeyle makine öğrenimi, veri analizi için algoritmaların kullanıldığı, verilerdeki yapıları belirleyip öğrenme süreciyle bu yapıları anlama ve bir durumu belirleme veya tahmin etme sürecidir (Yıldız. , 2022).

Makine öğrenmesi genellikle denetimli, denetimsiz ve pekiştirmeli öğrenme olarak üç ana kategoriye ayrılır (Gökalp, 2022). Denetimli öğrenmede, model, etiketlenmiş veriler kullanılarak eğitilir; yani giriş verileri ve karşılık gelen doğru çıktılar mevcuttur. Model, bu verilere dayanarak belirli bir görevi yerine getirmeyi öğrenir ve daha sonra bilinmeyen veriler üzerinde doğru tahminler yapmayı amaçlar. Bu yaklaşım, sınıflandırma ve regresyon problemlerinde yaygın olarak kullanılır.

Denetimsiz öğrenmede, model etiketlenmemiş verilerden örüntüleri ve ilişkileri keşfetmeye çalışır. Bu yöntem, verilerdeki gizli yapıları bulmak için kullanılır ve kümeleme veya boyut azaltma gibi uygulamalar için uygundur. Örneğin, müşteri segmentasyonu veya öneri sistemleri gibi alanlarda denetimsiz öğrenme kullanılarak verilerdeki benzerlikler ve farklılıklar keşfedilebilir.

Pekiştirmeli öğrenme, bir modelin bir ortamda etkileşimde bulunarak ve ödüller veya cezalar alarak öğrendiği bir yaklaşımdır. Bu tür öğrenme, oyun oynayan yapay zekalarda, robotikte ve otonom sistemlerde kullanılır. Model, hedeflenen ödülü maksimize etmek için en iyi stratejiyi belirlemeye çalışır.

Makine öğrenmesi, günümüz teknolojisinin önemli bir bileşenidir ve birçok sektörde devrim yaratmıştır. Tıp, finans, pazarlama, ulaşım, güvenlik ve daha birçok alanda makine

öğrenmesi, verilerden bilgi çıkarmak ve karar destek sistemleri oluşturmak için kullanılır. Ancak, bu teknolojiyle birlikte gelen etik ve yanlılık sorunları da dikkatle ele alınmalıdır. Modelin eğitildiği verilerin yanlılık içermemesi ve sonuçların adil olması, makine öğrenmesinin doğru ve etik bir şekilde kullanılmasını sağlamak için kritik öneme sahiptir.

3.2. Yapay Sinir Ağları

Yapay Sinir Ağları (YSA), 1890 yılında Alman bilim adamı Heinrich Wilhelm Gottfried von Waldeyer-Hartz tarafından insan beynindeki sinir hücrelerine atfen "nöron" olarak adlandırılmış bu hücrelerden esinlenerek geliştirilmiş bir hesaplama ve modelleme yöntemidir (Korkmaz, 2020). Nöronlar, birbiriyle bağlantılı düğümlerden oluşan bir sistemdir. Her düğüm, diğer düğümlerden ve dış çevreden gelen sinyalleri alır, bu sinyalleri bir aktivasyon veya dönüşüm süreci aracılığıyla işler ve daha sonra bu işlenmiş çıktıyı diğer nöronlara veya dış dünyaya iletir. Düğümler arasındaki bu bağlantılar, ağın genel davranışını ve bilgi akışını belirler (Yıldız, 2022).

Aktivasyon fonksiyonu, bir nöronun aldığı girdiye nasıl tepki vereceğini ve bu girdinin bir çıktı sinyali olarak nasıl dönüştürüleceğini belirleyen önemli bir bileşendir. Çeşitli aktivasyon fonksiyonları kullanılarak, nöronların girdi ve çıktıları arasındaki ilişkiler kontrol edilebilir. Bu fonksiyonlar, ağın doğrusal mı yoksa doğrusal olmayan mı çalışacağını ve sinyallerin iletilmesi veya engellenmesi konusunda nasıl davranacağını belirler.

Nöronlar arasındaki bağlantılar, ağın nasıl organize olduğunu ve bilgi iletim yolunu oluşturur. Bu bağlantılar aynı zamanda bir ağın öğrenme kapasitesini de etkiler; ağırlıklandırılmış bağlantılar, bir nöronun diğerine olan etkisini artırabilir veya azaltabilir. Sonuç olarak, nöronların birbiriyle nasıl etkileşime girdiği ve hangi aktivasyon fonksiyonlarının kullanıldığı, ağın genel performansını ve sonuçta ürettiği çıktıları büyük ölçüde etkiler.

Diğer bir deyişle, Yapay Sinir Ağları, insan beyninin öğrenme yeteneğini taklit eden bilgisayar sistemleridir (Öztemel, 2012).

Yapay sinir ağlarının önemli bir özelliği, deneyimlerden öğrenme yeteneğidir. Yapay sinir

ağları, deneyimlerden yararlanarak yeni bilgiler türetebilir, oluşturabilir ve keşifler yapabilir. Bu yetenekler, dış yardım almadan gerçekleştirilebilir. Ayrıca yapay sinir ağları, öğrenme yeteneğinin yanı sıra bilgileri analiz edip, bu bilgiler arasında ilişki kurma kabiliyetine de sahiptir (Yürük ve Ekşi, 2019).

Yapay sinir ağları, ağın yapısını belirleme, parametre seçimi ve eğitim sürecini tanımlama gibi konularda belirli bir standardın olmaması, sorunları sadece sayısal verilerle ifade etme zorunluluğu ve ağın davranışlarını tam olarak açıklayamaması gibi zorluklara rağmen, günümüzde giderek artan bir ilgiyle karşılanmaktadır. Bu artan ilginin sebeplerinden biri, yapay sinir ağlarının çeşitli uygulama alanlarında gösterdiği başarıdır (Öztemel, 2012).

Yapay sinir ağları, sınıflandırma, örüntü tanıma, sinyal işleme, veri sıkıştırma ve optimizasyon gibi konularda en etkili tekniklerden biri olarak kabul edilir. Özellikle büyük veri ve karmaşık örüntülerin analizi söz konusu olduğunda, bu ağlar benzersiz bir esneklik ve öğrenme yeteneği sunar. Yapay sinir ağları, çok katmanlı ve geniş ölçekli yapılarına rağmen, öğrenme kapasitesi ve hız gibi avantajlar sağlar.

Günlük yaşamda yapay sinir ağlarının başarılı uygulamalarına sıkça rastlanır. Örneğin, veri madenciliği, optik karakter tanıma, optimum rota belirleme, parmak izi tanıma, malzeme analizi, iş çizelgeleme ve kalite kontrol gibi alanlarda yapay sinir ağları yaygın olarak kullanılır. Bu teknolojilerin çeşitli endüstrilerde uygulanması, yapay sinir ağlarının potansiyelini ve gelecekteki uygulama alanlarını genişletmektedir.

Sonuç olarak, yapay sinir ağlarının çeşitli dezavantajlarına rağmen, onların sunduğu öğrenme esnekliği, güçlü modelleme yetenekleri ve çeşitli alanlarda gösterdikleri başarılı performans, teknoloji ve endüstri alanında artan ilgiyle karşılanmalarını sağlamaktadır. Yapay sinir ağlarının önemi, büyük veri çağında ve karmaşık sorunların çözümlerinde daha da artacaktır.

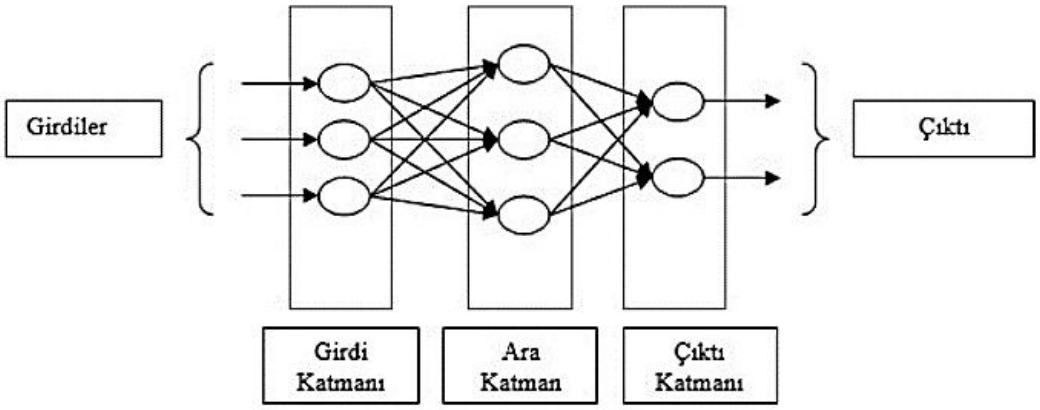
3.2.1. Yapay Sinir Ağlarının Yapısı

Yapay Sinir Ağları (YSA), yapay sinir hücrelerinin bir araya gelmesiyle oluşur. Bu hücreler ağ içinde iletişim halindedir. Her bir yapay sinir hücresinin çıktısı, diğer hücreler için bir girdi olarak işlev görür (Yürük ve Ekşi, 2019). Yapay sinir ağı genellikle katmanlardan

oluşur. En temel yapısı üç katmanlı olan bir yapay sinir ağı şu şekildedir:

1. **Girdi Katmanı:** Bu katman, verilerin ağa girdisinin sağlandığı katmandır. Örneğin, bir resmin pikselleri veya metin verileri gibi girdiler buraya verilir.
2. **Gizli Katmanlar (Ara Katman):** Yapay sinir ağının öğrenme işlemlerinin gerçekleştiği ve girdilerin işlendiği katmanlardır. Bu katmanlarda, nöronlar girdileri alır, belirli ağırlıklarla çarpar, aktivasyon fonksiyonunu kullanarak işler ve çıktı üretirler.
3. **Çıktı Katmanı:** Son katmandır ve ağın sonuçlarını verir. Örneğin, bir sınıflandırma probleminde, bu katman belirli bir sınıfa ait olasılıkları veya bir tahmini verebilir.

Yapay sinir ağları, öğrenme sürecinde girdi ve çıktı arasındaki ilişkileri belirlemek ve bu ilişkileri öğrenerek sonuçları tahmin etmek için ağırlık ayarlamaları yapar. Bu ağırlıklar, veri üzerinden tekrar tekrar ayarlanarak ağın belirli bir görevde daha iyi performans göstermesi sağlanır. Bu öğrenme süreci, genellikle gerçekleştirilen deneme-yanılma işlemleri yoluyla gerçekleşir (Yürük ve Ekşi, 2019).



Şekil 3.1: Yapay sinir ağı modeli

3.2.2. Yapay Sinir Ağlarının Özellikleri

Yapay sinir ağları, birçok benzersiz özelliğiyle karmaşık sorunların çözümünde önemli bir

rol oynar. Bu özellikler, ağın çeşitli durumlara uyum sağlama ve karmaşık verileri analiz etme yeteneklerini artırır.

Birincisi, yapay sinir ağlarının doğrusal olmama özelliği, onların karmaşık ilişkileri ve örüntüleri tanımlayabilmesini sağlar. Doğrusal olmayan yapılar, doğrusal modellere göre daha karmaşık sorunları ele alabilir ve daha zengin veri setlerinde etkili olabilir.

İkincisi, paralel çalışma yetenekleri, sinir ağlarının aynı anda birden fazla işlemi gerçekleştirmesine olanak tanır. Bu özellik, özellikle büyük veri setlerinde ve derin öğrenme uygulamalarında önemli bir avantajdır.

Üçüncüsü, yapay sinir ağları öğrenme yeteneğine sahiptir. Eğitim süreci boyunca ağ, girdi-çıktı ilişkilerini öğrenir ve bu ilişkileri genelleme yeteneğiyle yeni durumlara uygulayabilir. Genelleme, eğitilmiş bir modelin daha önce karşılaşmadığı durumlarda bile doğru tahminler yapabilme kapasitesini ifade eder.

Dördüncüsü, hata toleransı ve esneklik, yapay sinir ağlarının olası hatalara veya eksik verilere karşı dayanıklı olmasını sağlar. Bu, ağın sağlamlığını artırarak, gerçek dünya uygulamalarında hatalara veya belirsizliklere rağmen güvenilir sonuçlar üretmesine olanak tanır.

Beşincisi, yapay sinir ağları, eksik verilerle çalışma yeteneğine sahiptir. Bu da onları eksik veya eksikliği düzeltilebilen veri setlerinde kullanışlı kılar.

Altıncısı, çok sayıda değişken ve parametre kullanma özelliği, ağın karmaşık veri setlerinde çalışabilmesini sağlar. Sinir ağları, birden fazla değişken ve parametre ile çalışarak karmaşık sorunlara çözüm sunar.

Son olarak, uyarlanabilirlik, yapay sinir ağlarının değişen koşullara ve yeni bilgiye göre kendini yeniden ayarlayabilmesi anlamına gelir. Bu, sinir ağlarının çeşitli endüstrilerde ve değişen veri yapılarında esnek ve etkili olmasını sağlar.

Tüm bu özellikler bir araya geldiğinde, yapay sinir ağları çok yönlü, esnek ve karmaşık

sorunları çözmede etkili bir araç olarak ortaya çıkar. Bu nedenle, yapay sinir ağları, veri analizi, makine öğrenimi ve yapay zekâ gibi alanlarda yaygın olarak kullanılır. Bunlara ek olarak Yapay Sinir Ağları, birçok alanda en yaygın olarak tahmin, sınıflandırma, veri ilişkilendirme, veri yorumlama ve veri filtreleme gibi işlemlerde kullanılmaktadır (Öztürk ve Şahin, 2018).

3.2.3. Yapay Sinir Ağlarında Öğrenme

Yapay sinir ağlarının önemli bir özelliği, öğrenme yeteneğine sahip olmalarıdır. Öğrenme süreci, mevcut örnekler arasındaki ilişkileri başarılı bir şekilde modelleyebilecek bağlantı ağırlıklarının hesaplanmasıdır. Yapay sinir ağları, bu öğrenme sürecinde elde ettikleri bilgileri, sinir hücreleri arasındaki bağlantıların güçlüklerinde saklarlar. Bu bağlantı ağırlıkları, yapay sinir ağlarının verileri etkili bir şekilde işleyebilmesi için gereken önemli bilgileri içermektedir (Ataseven, 2013).

Yapay sinir ağlarında öğrenme türleri;

- **Denetimli Öğrenme (Supervised Learning):** Bu türde, algoritma giriş verilerini ve bu verilere karşılık gelen çıkışları öğrenir. Bir model, belirli bir girişle ilişkilendirilmiş doğru çıkışı tahmin etmeye çalışır. Denetimli öğrenme, yapay sinir ağlarının eğitim sürecidir. Bu süreç, sinir ağına giriş değerleri sağlanarak çıkış bilgilerinin belirlenmesini içerir. Ağ, ürettiği çıkışı, beklenen değerlerle karşılaştırarak ağırlıkların ayarlanması için gereken bilgiyi elde eder. Ağın çıktıları ile hedeflenen çıktılar arasındaki fark hesaplanır ve bu farka göre ağın ağırlıkları güncellenir. Bu güncelleme süreci, belirlenen bir hata sınırlamasına ulaşılana kadar devam eder (Kurt vd., 2017).
- **Danışmansız Öğrenme (Unsupervised Learning):** Bu türde, algoritma giriş verilerini ve bu verilerdeki desenleri keşfetmeye çalışır. Model, veriler arasındaki ilişkileri anlamaya çalışır, ancak çıkışlar önceden bilinmez.

  Ağa sadece giriş veri seti sağlanır ve ağın, bu veri setine uygun bir çıkış değeri üretebilmesi için kendi ağırlıklarını otomatik olarak ayarlaması istenir (Yazıcı vd.,

2007).

- **Destekleyici (Reinforcement Learning) Öğrenme:** Bu yöntemde, ağ her iterasyonun sonunda elde ettiği sonuçların başarısını değerlendirir. Bu değerlendirme sonucunda, ağın performansıyla ilgili geri bildirim alır ve bu geri bildirimlere dayanarak kendisini yeniden düzenler.

- Destekleyici öğrenme, denetimsiz öğrenmenin bir türü olan Yapay Sinir Ağı (YSA) düğümlerinin girdi veri alt kümeleriyle etkileşime girerek rekabet ettiği, Hebbian öğrenme prensibine dayalı bir yaklaşımdır. Bu yöntem, her bir düğümün uzmanlığını artırarak çalışır ve veri içindeki kümeleri tanımlamak için oldukça etkili bir yöntemdir. Destekleyici öğrenme temelli modeller ve algoritmalar, vektör nicelleştirme ve kendini organize eden haritaları içermektedir (Şanlı, 2018).

3.3. Derin Öğrenme

Hinton'un yaptığı çalışmalar ve yayınladığı makaleler, yapay sinir ağlarına yeni bir perspektif getirmiştir. Derin Öğrenme, beynin yapısından ve Yapay Sinir Ağları olarak adlandırılan işlevlerinden ilham alan bir makine öğrenme algoritmasıdır. Bu algoritma, biyolojik nöronlara benzer şekilde, yapay nöronlar tarafından giriş sinyallerinin alındığı, bu sinyallerin toplandığı, işlendiği ve ardından çıkışlara iletildiği bir süreci taklit eder (Şişmanoğlu vd., 2020).

Konvolüsyonel sinir ağları, çok katmanlı yapılarıyla bilinen sinir ağları türleridir. Bu tür ağlar, özellikle görüntü işleme ve bilgisayarla görme gibi alanlarda önemli başarılar elde etmişlerdir. Derin konvolüsyonel sinir ağları, bu başarının üzerine inşa edilerek, daha karmaşık sorunları çözme kapasitesine sahip olmuştur. Derin öğrenme, çok katmanlı yapısıyla aynı anda birçok işlemi gerçekleştirerek sonuca ulaşma yeteneği sunar. Bu yaklaşım, özellikle ön işlem ve özellik çıkarımı gibi adımların sinir ağı içinde otomatik olarak yapılmasına olanak tanır. Yani, ağ, ham verileri alır ve içindeki çeşitli katmanlar aracılığıyla bu verileri analiz ederken, gereken özellikleri otomatik olarak çıkartır. Derin konvolüsyonel sinir ağları, bir görüntü veya veri kümesindeki önemli özellikleri kendi içinde tanımlar ve bu özellikler katmanlar arasında ilerlerken otomatik olarak tespit edilir. Bu, daha hızlı ve etkili sonuçlar elde etmeyi sağlar. Özetle, derin konvolüsyonel sinir ağları, veri işleme ve analiz sürecini büyük ölçüde otomatikleştirerek, karmaşık sorunların çözümünde

büyük bir avantaj sağlar (Doğan ve Türkoğlu, 2019).

Derin öğrenme, makine öğreniminin bir dalı olup, sinir ağları, yapay zekâ, grafik modelleme, optimizasyon, örüntü tanıma ve sinyal işleme gibi farklı alanların birleşiminden doğmuştur. Derin öğrenme ağları, geleneksel sinir ağlarının daha karmaşık ve derin versiyonlarıdır ve daha güçlü tahminler elde etmek için tasarlanmıştır. Derin öğrenme, çok katmanlı yapılar kullanarak verilerden denetimli veya denetimsiz öğrenmeyi gerçekleştirir. Bu modellerdeki katmanlar, doğrusal olmayan veri dönüşümlerini bir dizi ardışık aşama şeklinde uygular. Her katman, verileri daha yüksek ve soyut seviyelerde temsil eder, böylece daha karmaşık örüntüler ve ilişkiler yakalanabilir. Derin öğrenmenin gücü, verilerin çok katmanlı ve karmaşık bir şekilde işlenmesinden gelir, bu da ona görüntü işleme, dil tanıma ve diğer karmaşık görevlerde büyük avantajlar sağlar. Özellikle büyük veri ve karmaşık örüntülerin analizi için derin öğrenme, makine öğreniminde devrim niteliğindedir (Sakarya ve Yılmaz, 2019).

Derin öğrenme modelleri, özellikle karmaşık yapıdaki verilerin otomatik olarak çıkarılmasını ve işlenmesini sağlamak için kullanılır. Bu modeller, derin sinir ağları vasıtasıyla verilerin anlamlı temsillerini elde etmek veya sınıflandırma işlemlerini gerçekleştirmek için geliştirilmiştir. Bu sayede, dış yardıma minimum düzeyde ihtiyaç duyularak veri analizi ve yorumlama süreçleri daha etkin bir şekilde gerçekleştirilebilir (Kilimci, 2020). Ayrıca derin öğrenme, bilgiye erişim sürelerini kısaltmak ve en doğru bilgiye hızlıca ulaşmak için bir araç olarak da kullanılmaktadır. Bu teknoloji, bilgiye erişimi hızlandırarak ve optimize ederek daha verimli kararlar alınmasına olanak tanır (Akbulut ve Adem, 2023).

Derin öğrenme, verilerin farklı özellik seviyelerini ve temsillerini öğrenerek karmaşık yapıları anlamamıza olanak tanıyan bir yapıya sahiptir. Bu yapının içinde, daha karmaşık ve soyut özellikler, daha temel ve düşük seviyeli özelliklerden türetilir, böylece verilerin hiyerarşik bir şekilde temsil edilmesi sağlanır. Bu yöntem, verilerin daha derinlemesine analiz edilmesine ve daha kapsamlı sonuçlar elde edilmesine olanak sağlar (Şeker vd., 2017).

6 tane derin öğrenme yöntemi vardır:

- Derin Sinir Ağları: Bu ağda, bir giriş katmanı bulunur ve doğrudan çıktıya bağlanır. Bu yapı, lineer olarak ayrılabilir desenlerin sınıflandırılması için kullanılabilen sınıflandırıcıları oluşturmak amacıyla geliştirilebilir. Daha karmaşık problemler için, bu algoritma bir veya daha fazla gizli katman ile genişletilebilir ve her katmanın ağırlıklarını ayarlamak için delta kuralı adı verilen bir öğrenme yöntemi kullanılabilir (Küçük ve Arıcı, 2018).
- Derin Oto-Kodlayıcılar: Derin Oto-kodlayıcılar, girdi olarak aldıkları veriyi çıktı olarak üretmeye çalışan, birden fazla katmandan oluşan bir derin öğrenme algoritmasıdır. Bu algoritmaların temel amacı, verilen girdiye en yakın çıktıyı üretebilen ve veri içindeki önemli yapısal bilgileri çıkaran bir fonksiyonu bulmaktır (Kayaalp ve Süzen, 2018).
- Derin İnanç Ağları: Derin İnanç Ağları, özellikle görüntü tanıma uygulamalarında kullanılan bir türdür ve sınırlı Boltzman Makineleri'nin yığılmış ve oto-kodlayıcılarla birleşimi olarak kabul edilir (Gündüz ve Cedimoğlu, 2019).
- Derin Boltzmann Makinesi: Öncelikle, derin inanç ağları gibi, Derin Boltzmann Makineleri, içsel temsilleri giderek karmaşık hale getirme potansiyeline sahiptir; bu, nesne ve konuşma tanıma problemlerini çözme konusunda umut verici bir yöntem olarak kabul edilmektedir. Derin inanç ağlarından farklı olarak, yaklaşık çıkarım prosedürü, başlangıçta bir alttan yukarı geçişin yanı sıra üstten aşağı geri bildirim içerebilir, bu da derin Boltzmann makinelerinin belirsiz girişlerle daha iyi bir şekilde baş etmesine ve dolayısıyla daha sağlam bir şekilde işlemesine izin verir (Salakhutdınov ve Hinton, 2009).
- Yinelenen Sinir Ağları: Yinelenen sinir ağları 1990'lı yıllarda araştırma ve geliştirmenin önemli bir odağı olmuştur. Sıralı veya zamanla değişen örüntüleri öğrenmek için tasarlanmışlardır. Yinelenen bir ağ, geri beslemeli (kapalı döngü) bağlantılara sahip bir sinir ağıdır (Medsker ve Jain, 2001).
- Evrişimsel Sinir Ağları: Genellikle bilgisayarlı görü alanında ve görsel veriler üzerinde kullanılan evrişimsel sinir ağları, büyük boyutta işaretli veri gerektirmesine rağmen başarılı sonuçlar vermektedir. Ayrıca, doğal dil işleme alanındaki problemlere de başarıyla uygulanmıştır.

Derin öğrenme, aşağıdaki durumları kapsayacak şekilde kullanılabilmektedir (Ceyhan, 2023):

- Ticaret Stratejisi
- Fiyat Tahmini
- Portföy Yönetimi
- Piyasa Simülasyonu
- Hisse Senedi Seçimi
- Risk Yönetimi
- Riskten Korunma

3.4. Veri Madenciliği

Veri madenciliği, büyük veri setlerinde gizli desenleri, bilgiyi ve anlamlı ilişkileri ortaya çıkarmak için istatistiksel ve matematiksel yöntemlerin kullanıldığı bir veri analizi sürecidir. Bu süreçte, büyük miktardaki veriyi anlamlı bir şekilde analiz etmek ve değerli içgörüler elde etmek için bir dizi teknik ve yöntem uygulanır (Baykal, 2006).

Veri madenciliği kapsamında yaygın olarak kullanılan teknikler şunlardır:

- **Veri kümelenmesi**: Benzer özelliklere sahip veri noktalarını bir araya getirerek gruplar oluşturma sürecidir.
- **Özetleme**: Büyük veri setlerini daha küçük, anlaşılır biçimlerde sunmak için kullanılan tekniktir.
- **Sınıflandırma kurallarının öğrenilmesi**: Veri setindeki öğeleri belirli kategorilere veya sınıflara ayırmak için kurallar geliştirme sürecidir.
- **Bağımlılık ağlarının belirlenmesi**: Veriler arasındaki ilişkileri ve bağımlılıkları ortaya çıkarır.
- **Değişkenlik analizi**: Verideki değişkenlik ve trendleri inceleyerek anlamlı sonuçlar çıkarmayı sağlar.
- **Anormal veri tespiti**: Normal veri örüntülerinden sapmaları ve aykırı değerleri tespit etme işlemidir.

Bu teknikler, farklı endüstrilerde ve uygulamalarda veri madenciliğinin geniş bir kullanım alanı olduğunu gösterir. Veri madenciliği, şirketlerin, araştırmacıların ve analistlerin büyük veri setlerinden değerli içgörüler elde etmelerini ve bu bilgiyi stratejik kararlar için kullanmalarını sağlar.

Veri madenciliği, istatistik, yapay zekâ ve makine öğrenmesi gibi farklı alanlardan türetilmiş çeşitli teknikleri ve yöntemleri içeren bir disiplindir. Ancak, sadece araçlar ve tekniklerle sınırlı değildir; aynı zamanda veri toplama, veri temizleme, model oluşturma, model testi ve uygulama gibi geniş bir süreci kapsar (Seyrek ve Ata, 2010).

Veri madenciliği, genellikle aşağıdaki adımları içerir:

- Veri temizleme
- Veri bütünleştirme
- Veri indirgeme
- Veri dönüştürme
- Değerlendirme

Veri Temizleme: Veri tabanında hatalı veriler olabilir. Bu hataların giderilmeye çalışıldığı adım veri temizleme adımıdır. Eksik değer içeren kayıtların atılması, kayıp değerlerin sabit bir değerle doldurulması, diğer verilerin ortalamasının hesaplanarak kayıp verilere bu değerin atanması veya verilere uygun bir tahmin yapılması ve eksik verinin bu tahminle yerine konulması veri temizlemede yapılan işlemlerdendir.

Veri Bütünleştirme: Farklı veri tabanları veya veri kaynaklarından gelen verilerin birleştirilerek tek bir formatta sunulması ve değerlendirilmesi için yapılan işlem, veri entegrasyonu olarak adlandırılır. Bu süreç, verilerin tutarlılığını ve bütünlüğünü sağlamak amacıyla gerçekleştirilir.,

Veri İndirgeme: Veri madenciliği projelerinde, analizin sonuçlarını etkilemeyecek şekilde, veri setinin boyutunu veya değişkenlerin sayısını azaltmak mümkündür. Bu sürece veri indirgeme adı verilir ve verilerin özünü korurken analizin karmaşıklığını azaltmayı amaçlar.
Veri Dönüştürme: Veri dönüşümü, verinin içeriğini koruyarak farklı bir formata

dönüştürülmesini ifade eder. Bu dönüşüm, verinin kullanılacağı modele uygun bir şekilde yapılmalıdır, çünkü model ve algoritmanın türü, verinin nasıl temsil edileceğinde önemli bir etkiye sahiptir.

Değerlendirme: Algoritmaların uygulanmasının ardından, elde edilen sonuçlar düzenlenir ve ilgili yerlere sunulur. Sunumda kullanılacak grafikler ve şekiller, modele uygun olarak seçilir ve yerleştirilir.

Gördükleri işleve göre veri madenciliği modelleri aşağıdaki gibidir;

- Sınıflandırma ve Regresyon
- Kümeleme
- Birliktelik Kuralları

Sınıflandırma ve Regresyon: Sınıflama ve regresyon, veri analizinde kullanılan iki temel yaklaşımdır ve farklı veri tiplerini ele alırlar. Sınıflama, belirli bir veri örneğini önceden tanımlanmış kategorilere atama işlemidir, regresyon ise sürekli değişkenlerin tahmininde kullanılan bir yöntemdir.

Örneğin, bir sınıflama modeli, müşteri profillerini belirli kategorilere ayırmak için kullanılabilir; örneğin, bir çevrimiçi mağazanın ziyaretçilerini potansiyel alıcılar ve sadece gezginler olarak ayırmak gibi. Diğer taraftan, bir regresyon modeli, bir otomobilin fiyatını tahmin etmek için kullanılabilir; bu, modelin araba özelliklerini ve pazar trendlerini analiz ederek bir fiyat tahmini yapmasıyla gerçekleşir (Özekes, 2003).

Kümeleme:

Sınıflandırma ve regresyon modelinde kullanılan teknikler;

- Karar Ağaçları
- Yapay Sinir Ağları
- Genetik Algoritmalar

- K-En Yakın Komşu
- Bellek Temelli Nedenleme
- Naive-Bayes

4. REGRESYON MODELLERİ

Regresyon modelleri, bağımlı bir değişken ile bir veya daha fazla bağımsız değişken arasındaki ilişkiyi tahmin etmek veya modellemek için kullanılan istatistiksel yöntemlerdir. Bu modeller, çeşitli disiplinlerde veri analizi, tahmin, modelleme ve ilişkilerin anlaşılması için kullanılır. Regresyon modelleri, istatistiksel yazılım ve programlama dilleri (örneğin R, Python, SAS, SPSS) kullanılarak uygulanır. Seçilen model, verinin doğası, bağımlı değişkenin türü, bağımsız değişkenlerin sayısı ve türü gibi faktörlere bağlıdır.

4.1. Doğrusal Regresyon

Doğrusal regresyon analizi, hedeflenen bir değişkeni tahmin etmek için kolayca ölçülebilen veya erken tespit edilebilen diğer değişkenlerden yararlanarak bir model kurma yöntemidir (Kılıç, 2013). İki bağımlı değişkenli bir sistem için kurulacak model örneği aşağıdaki gibidir;

$$\gamma = \beta_0 + \beta_1 \times x_1 + \beta_2 * x_2 + \varepsilon \qquad \text{(Denklem 13)}$$

Kurulan modelde;
y = bağımlı değişken
x = açıklayıcı değişken
β = çarpılan ağırlıklar
ε = hata terimini ifade etmektedir.

Doğrusal regresyon, normal dağılımlı veriler için uygundur ve genellikle ortalama değerler üzerinden çalışır. Bu nedenle, yüksek ve düşük değerler birbirini dengelediğinden, doğrusal regresyon uzun vadeli tahminlerde genellikle etkili olur. Bu etkinlik, çeşitli çalışmalarda da gözlemlenmiştir (Arda ve Küçükkocaoğlu, 2021).

4.2. Bayes Doğrusal Regresyon

Bayesci doğrusal regresyon, parametre tahmini için Bayesci yaklaşım kullanan bir regresyon türüdür. Bayesci yaklaşımla çalışırken, başlangıçtaki bilgi veya inançları temsil eden bir önsel dağılım, gözlemlenen verilerin parametrelere bağlı olasılığını gösteren olabilirlik

dağılımı ve sonrasında parametreler hakkında güncellenmiş bilgi veren sonuç dağılımı kullanılır. Doğrusal regresyonun klasik bir formu olan en küçük kareler hata teriminin normal dağılımı olduğu varsayılır (Permai ve Tanty, 2018). Doğrusal regresyon modelindeki tahminlerin aksine olasılık dağılımları ile bir model kurulur.

$$\gamma \approx N(\beta^T X, \sigma^2 1) \quad \text{(Denklem 14)}$$

Bu regresyon modeli bağımlı değişken için tek bir "en iyi" tahmin sunmaktan ziyade, bu değişkenin alabileceği olası değerlerin dağılımını tahmin etmeyi hedefler (Arda ve Küçükkocaoğlu, 2021). Bayes Doğrusal Regresyonda, sonuç bir olasılık dağılımı olarak ifade edilir ve model parametreleri de bir dağılım şeklinde kabul edilir. Dağılımlı modelin denklemi aşağıdaki gibidir;

$$P(\beta|y,X) = \frac{P(y|\beta,X)*P(\beta|X)}{P(y|X)} \quad \text{(Denklem 15)}$$

Yukarıdaki model çıktı olarak bir dağılım grafiği verir.

4.3. Karar Ağacı Regresyonu

Karar ağacı, makine öğrenimi teorisinde kök salmış olan ve sınıflandırma ve regresyon problemlerinin çözümü için verimli bir araçtır. Ortak bir kararlaştırma adımında bir dizi özelliği (veya bantları) kullanarak sınıflandırma yapan diğer sınıflandırma yaklaşımlarının aksine, karar ağacı, çok aşamalı veya hiyerarşik bir karar şeması veya ağaç benzeri bir yapıya dayanır. Ağaç, bir kök düğüm (tüm verileri içerir), bir dizi iç düğüm (ayrımlar) ve bir dizi terminal düğümden (yapraklar) oluşur. Karar ağacı yapısının her düğümü, bir sınıfı veya bazı sınıfları kalan sınıflardan ayıran ikili bir karar verir. İşlem genellikle yaprak düğüme ulaşana kadar ağaç boyunca aşağı doğru ilerleyerek gerçekleştirilir. Buna yukarıdan aşağı yaklaşımı denir (Xu vd., 2005).

Karar ağacı, tahminlerde bulunmak için belirli özelliklere dayanarak karar verdirici çıkarımlar yapan, denetimsiz bir makine öğrenimi yöntemidir. Modelin ismine uygun olarak, bu yaklaşım veriyi bölüp, her bir parçanın özelliklerine göre tahminler yapar.

4.4. Destekli Karar Ormanı Regresyonu

Destekli karar ormanı, regresyon analizinde kullanılan bir modeldir ve modeldeki alt segmentlerdeki kararların birleştirilmesiyle tahmin yürütür. Temel konsept, ardışık her bir ağacın, önceki ağacın tahmin ettiği hatalar temel alınarak oluşturulması ve bu nedenle aslında basit ağaçların bir araya getirilmesiyle ortaya çıkan hesaplamalar dizisidir.

$$g(x) = f_0(x) + f_1(x) + f_2(x) + \cdots \quad \text{(Denklem 16)}$$

Model yukarıdaki gibi bir g fonksiyonuyla gösterilebilir.

Rastgele orman algoritmasından farklı olarak, destekli karar ormanı regresyonu, "gradyan artırma" olarak bilinen özel bir modelleme tekniği kullanır ve bu teknik, her bir ağacın bir alt veri örneğiyle eğitilmesi sürecini içerir (Arda ve Küçükkocaoğlu, 2021).

4.5. Hızlı Orman Yüzdelik Regresyonu

Hızlı Orman Yüzdelik Regresyonu, parametrik olmayan dağılımları tahmin edebilen güçlü bir ağaç tabanlı yüzdelik regresyon modelidir. Hızlı orman yüzdelik regresyonu, rastgele orman yüzdelik regresyonunu uygulamak için karar ağaçlarını kullanır. Rastgele ormanlar, karar ağaçlarıyla aşırı uydurmayı önlemeye yardımcı olabilir. Rastgele ormanda bir ağaç topluluğu, bir alt küme rastgele örnek ve eğitim verisi özellikleri seçmek için çantalanma kullanılarak geliştirilir ve ardından her veri alt kümesine bir karar ağacı uyarlanır. Rastgele orman algoritmasından farklı olarak, bu yöntem tüm ağaçların çıktısını ortalamasından ziyade her bir alt kümeyi ayrı ayrı ele alır. Hızlı orman yüzdelik regresyonu, yüzdelik örnek sayısı parametresi tarafından belirtilen ağaçlarda tahmin edilen tüm etiketleri saklar. Bu, kullanıcının verilen bir örnek için yüzdelik değerlerini görebilmesi için dağılımı gösterir. Hızlı orman yüzdelik regresyonunun ana gücü, her ağacın her yaprağında, ortalama değil, ilgili tüm gözlemlerin saklanmasıdır, bu durum rastgele ormanda ortalama olarak gerçekleşir. Rastgele ormanda olduğu gibi koşullu ortalama yerine, belirli bir örneğin koşullu yüzdeliklerini tahmin etmeye yardımcı olur. Hızlı orman yüzdelik regresyonu gibi ağaç tabanlı yüzdelik regresyon modellerinin ek avantajı, parametrik olmayan dağılımları tahmin etmek için kullanılabilmeleridir. Genel olarak, rastgele orman, orijinal eğitim

verilerinden birçok bootstrap örneği çıkararak ve her bir örneği bir ağaçla uyarlayarak daha doğru regresyonlar oluşturmak için birkaç regresyon ağacını bir toplulukta birleştirir (Zahıd vd., 2020).

Modelin matematiksel denklemi aşağıdaki gibidir;

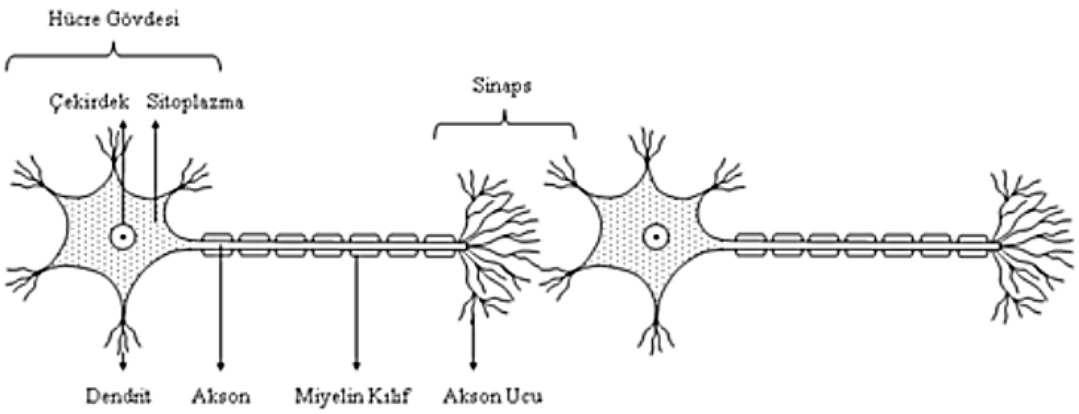
$$Y_m^p = \sum_{j=1}^{n} \omega_j(x_m) Y_j \qquad \text{(Denklem 17)}$$

4.6.Nöral Ağ Regresyonu

Nöral ağ regresyonu, yapay zekâ ve makine öğrenimi alanında kullanılan bir regresyon tekniğidir ve biyolojik sinir sistemlerinden esinlenmiştir. Biyolojik sinir sistemi, birbirine bağlı bir ağ oluşturan ve "nöron" adı verilen hücrelerden oluşur. Nöronlar arasında bilgi, "aksiyon potansiyeli" olarak bilinen elektriksel sinyallerle taşınır. Aksiyon potansiyeli, nöronun uzun uzantısı olan akson boyunca ilerler ve aksonun uç kısmındaki telodendritler adı verilen dallı bölgeye ulaşır (Arda ve Küçükkocaoğlu, 2021).

Telodendritlerde, sinaps denilen boşluklar aracılığıyla nörotransmitterler serbest bırakılır. Bu nörotransmitterler, sinaps sonrası nöronlarla etkileşime girerek sinyali iletir ve böylece sinir ağı boyunca bilgi taşınmış olur.

Yapay nöral ağlar, bu biyolojik yapıyı taklit eder. Veriler, katmanlar halinde düzenlenmiş nöronlar arasında işlenir ve bu yapı, karmaşık regresyon sorunlarını çözmek için kullanılır. Yapay nöral ağlar, tıpkı biyolojik sinir sistemi gibi, bir girdiden başlayarak çok sayıda işlem katmanından geçerek bir çıktıya ulaşır. Bu, onları makine öğrenimi ve yapay zekâ uygulamaları için çok güçlü ve esnek bir araç haline getirir.



Şekil 4.1: Nöron yapısı (Arda ve Küçükkocaoğlu, 2021)

Bilgisayar bilimlerinde sinir sistemi, nöronların elektriksel sinyalleri iletme şeklini basit bir matematiksel modelle açıklanır. Nöronlar, belirli bir eşik değerin üzerine çıkan bir aksiyon potansiyeli aldıklarında, bu sinyalin akson boyunca ilerlemesine ve akson ucundaki nörotransmitter keseciklerinin patlamasına neden olur. Keseciklerin patlamasıyla birlikte nörotransmitterler sinapsa salınır, bu da sinyalin bir sonraki nörona iletilmesini sağlar.

Eğer aksiyon potansiyeli eşik değerini aşamazsa, nörotransmitter kesecikleri patlamaz ve sinyal bir sonraki nörona iletilmez. Bilgisayar bilimlerindeki algoritmalar, nöral ağlar gibi yapılarla bu mekanizmayı taklit ederek bilgi işleme ve karar verme süreçlerini simüle eder. Yani, bir nöronun ateşlenip ateşlenmemesi, algoritmaların da benzer şekilde belirli eşik değerleri kullanarak kararlar almasını yansıtır.

4.7. Poisson Regresyonu

Poisson regresyon modeli, özellikle kesikli ve kategorik olmayan bağımlı değişkenlerin analizinde kullanılır (Deniz, 2005). Bu model genellikle nadir olayların veya belirli bir zaman aralığındaki sayısal verilerin analiz edilmesi için uygundur. Modelin temel varsayımı, olayların belirli bir zaman veya alan içinde rastgele ve bağımsız olarak meydana gelmesidir.

Poisson regresyonu, bir olayın belirli bir zaman diliminde veya belirli bir alanda meydana gelme sayısını tahmin eder. Örneğin, bir restorana gün boyunca gelen siparişleri düşünelim. Kahvaltı, öğle yemeği ve akşam yemeği saatlerinde yoğunluk görülürken, bu ana zaman

dilimleri arasında seyrelme yaşanır. Poisson regresyon modeli, bu tür olayların dağılımını anlamak için kullanılabilir. Benzer şekilde, borsa açılış ve kapanış saatlerinde işlem hacmi yoğunken, bu saatler arasında işlem hacmi düşer. Poisson regresyon modeli, bu tür zaman dilimleri boyunca olayların ne şekilde değiştiğini anlamak için kullanılır. Model, özellikle nadir olayların dağılımını analiz ederek belirli bir zaman veya alan diliminde olayların ortalama sayısını tahmin etmek için etkilidir (Arda ve Küçükkocaoğlu, 2021).

Poisson dağılım fonksiyonu aşağıdaki gibidir;

$$P_x(k) = \frac{e^{-(\lambda t)} * (\lambda t)}{k!} = Poisson\ (\lambda t) \qquad \text{(Denklem 18)}$$

$P_x(k)$= t sürede k olayının görülme ihtimali
λt= birim zamanda gerçekleşen olay
k= olay sayısı

4.8. Yüksek Veri Boyutluluğu

Boyutlanma, makine öğrenme algoritmalarının boyut arttıkça daha karmaşık ve maliyetli hale geldiğini belirtir. Yeni bir özellik eklenildiğinde veriler daha seyrek hale gelir ve bu, model doğruluğunu olumsuz etkileyebilir. Veri ön işleme, veri analitiği için kritik bir adımdır ve bu süreç, daha verimli ve basit modeller yaratmaya yardımcı olabilir (Kuızınıenė vd., 2022).

4.8.1. Yüksek Veri Boyutluluğu ile Başa Çıkma Yöntemleri

Verinin boyutları, her gözlem için ölçülen özelliklerin sayısını ifade eder. Yüksek boyutlu verileri analiz etmek giderek daha zorlaşır. İlgisiz veya tekrarlayan özellikleri ortadan kaldırmak için literatürde çeşitli boyut azaltma teknikleri bulunmaktadır. Uygun bir boyut azaltma tekniği seçmek, işlem hızını artırabilir ve değerli bilgileri elde etmek için gereken zaman ve çabayı azaltabilir (Ayesha vd., 2020).

4.8.2. Boyut Azaltma Teknikleri

Boyut indirgeme, yüksek boyutlu veri temsillerini düşük boyutlu temsillere dönüştürme işlemidir. Yüksek boyutlu verilerin hızla artması, farklı uygulama alanlarında çeşitli boyut indirgeme tekniklerinin kullanılmasını yaygınlaştırmıştır. Ayrıca, bu alanda sürekli yeni teknikler geliştirilmekte ve kullanılmaktadır. Boyut indirgeme teknikleri, yüksek boyutlu orijinal veri setlerini, verilerin temel anlamlarını koruyarak daha düşük boyutlu bir temsile dönüştürmeyi amaçlar. Orijinal verilerin düşük boyutlu temsilleri, boyutlanma laneti sorununu hafifletmek için etkili bir çözümdür. Düşük boyutlu veriler, daha kolay işlenebilir, analiz edilebilir ve görselleştirilebilir, bu da veri analitiği ve makine öğrenimi alanlarında pratik avantajlar sağlar.

Başlıca boyut azaltma teknikleri;

Boyut Azaltma Teknikleri	Doğrusal Boyut Azaltma	Temel bileşen analizi
		Tekil değer ayrıştırma
		Gizli anlamsal analiz
		Yerelliği koruyan projeksiyonlar
		Bağımsız bileşen analizi
		Doğrusal diskriminant analizi
		Projeksiyon takibi
	Doğrusal Olmayan Boyut Azaltma	Kernel temel bileşen analizi
		Çok boyutlu ölçekleme
		İsomap
		Yerel doğrusal gömme
		Kendi kendini düzenleyen harita
		Vektör nicelemeyi öğrenme
		T-Stokastic komşu gömme

Şekil 4.2: Boyut azaltma teknikleri

4.8.2.1.Doğrusal Boyut Azaltma

Lineer boyut indirgeme yöntemleri, yüksek boyutlu veriyi daha düşük boyutlu bir alana doğrusal olarak yansıtma fikrine dayanır. Bu yöntemlerde hedef, n adet örnek içeren d boyutlu veriyi, örnek sayısı n sabit kalacak şekilde r boyutlu bir yapıya dönüştürmek için dx_r

boyutunda bir dönüşüm matrisi P bulmaktır. P dönüşüm matrisi kullanılarak, yüksek boyutlu veri öğeleri, aşağıdaki gibi düşük boyutlu öğelere dönüştürülür (Yıldız ve Sevim, 2016):

$$y_i = P^T x_i \quad \text{(Denklem 4)}$$

Yüksek boyutlu veri matris şeklinde tanımlanırsa $X=[x_1,x_2\ldots,x_n]$, boyutu indirgenmiş veri rx_n matrisi $Y=[y_1,y_2\ldots,y_n]$ aşağıdaki denklem ile dönüştürülür:

$$Y = P^T X \quad \text{(Denklem 5)}$$

Çeşitli doğrusal boyut azaltma yöntemleri, farklı P dönüşüm matrisleri üretir. Yöntemler arasındaki fark, verinin hangi özelliklerinin yansıtılacağı ve korunacağına odaklanır.

Temel Bileşen Analizi

Temel bileşen analizi, bir veri setinin varyans ve kovaryans yapısını doğrusal bileşimler yoluyla açıklayarak, boyut indirgeme ve yorumlama sağlayan çok değişkenli bir istatistik yöntemidir. Bu yöntemde, karşılıklı bağımlılık gösteren p değişken, n adet ölçümle temsil edilir ve doğrusal, ortogonal (dik) ve birbirinden bağımsız özelliklere sahip k yeni değişkene dönüştürülür. Temel bileşen analizi, veri setindeki temel bilgileri çıkarmada etkili bir yoldur ve yüksek boyutlu verilerde boyut sayısını azaltarak veri sıkıştırması sağlar (Yıldız vd., 2010).

Boyut indirgeme sürecinde bazı özelliklerin kaybolması kaçınılmazdır, ancak hedeflenen, kaybolan özelliklerin veri seti hakkında az bilgi içermesidir. Temel bileşen analizinin amacı, verinin çeşitliliğini en iyi şekilde temsil edebilecek yeni bir boyut takımı oluşturmaktır. İlk bileşen, mümkün olduğunca fazla çeşitliliği yansıtacak şekilde seçilir. İkinci bileşen ise ilkine dik olacak şekilde, yine maksimum çeşitliliği temsil edecek şekilde belirlenir.

Tekil Değer Ayrıştırma

Tekil değer ayrıştırma, doğrusal boyut indirgeme tekniklerinden biridir ve Temel Bileşen Analizi ile benzerlik gösterir. Tekil değer ayrıştırma, matrisleri daha düşük boyutlu

temsillere ayırarak metrik denklemleri çözmek ve veri indirgeme problemlerini çözmek için kullanılır. Özellikle dijital görüntü işleme, biyolojik dizilerin taksonomik sınıflandırılması, desen tanıma, gen ekspresyon verileri, sinyal işleme, doğal dil işleme, biyoinformatik ve metin özetleme gibi birçok alanda yaygın olarak kullanılır (Ayesha vd., 2020).

Tekil değer ayrıştırma, gerçek dünyadaki herhangi bir matrise uygulanabilir ve matrisleri U, S ve V T şeklinde ayrıştırır. U ve V, ortogonal matrislerdir; S ise tekil değerleri içeren bir diyagonal matristir. Tekil değer ayrışımının önemli bir özelliği, orijinal matrisi bu bileşenlerden yeniden oluşturabilmesidir. Ancak, tekil değer ayrıştırma hesaplama açısından karmaşık olabilir ve aykırı değerlere veya veri içindeki doğrusal olmayanlara karşı hassas olabilir. Ayrıca, tekil değer ayrışımı sonuçlarının görselleştirilmesi ve yorumlanması her zaman kolay değildir.

Verimliliği artırmak için kısıtlı tekil değer ayrışımı ve çok seviyeli tekil değer ayrışımı gibi farklı versiyonlar geliştirilmiştir. Bu yöntemler, farklı veri türlerinin ön işlenmesi ve yönetimi için faydalı olabilir ve eğitim, tıp ve yaşam bilimleri gibi alanlarda kullanılabilir.

Gizli Anlamsal Analiz

Gizli anlamsal analiz, metinlerin geniş bir külliyatında istatistiksel hesaplamalar yoluyla kelimelerin bağlamsal kullanım anlamlarını çıkarmayı ve temsil etmeyi amaçlayan bir teori ve yöntemdir. Gizli anlamsal analizin temel konsepti, bir kelimenin bulunduğu ve bulunmadığı tüm kelime bağlamlarının toplamının, kelimeler ve kelime grupları arasındaki anlamsal benzerliği büyük ölçüde belirleyen karşılıklı ilişkileri ortaya çıkarmasıdır (Landauer vd., 1998).

Gizli anlamsal analiz, metinlerin anlamsal temsilini öğrenmek ve kelimeler arasındaki ilişkileri çıkarmak amacı taşır. Gizli anlamsal analiz, girdi olarak eş oluşum matrisi Y ∈ R$m \times p$ den m × p kelime alır. Her matris girdisi y_{ij}, belirli bir kelimenin i, belirli bir belge j için yerel sıklığını temsil eder. Eş oluşum sayıları, bir belgenin anlamı hakkında bilgi edinmek için ağırlıklara dönüştürülür. Ağırlıklı terimlerin dönüştürülmüş matrisinin boyut faktörleri vardır ve tekil değer ayrışımı hesaplaması ile azaltılabilir. Örneğin, gizli anlamsal analiz [USV T] = $SV\ D(\mathrm{Y},\ k)$ olup, burada U ve V ortogonal matrislerdir. Son olarak,

azaltılmış ve ayrıştırılmış matristen yeni bir matris yeniden oluşturulur ve benzerlik ölçüleri bu yeniden oluşturulan matristen hesaplanır. Gizli anlamsal analizde, tekil değer ayrışımı kullanılarak hesaplanan tekil değerler, kelimeler ve pasajlar için dildeki anlam boyutlarını temsil etmek için kullanılır. Bu anlamsal temsillerin boyutu, bazı tekil değerleri 0 ile değiştirerek azaltılabilir. Gizli anlamsal analiz, terimlerin ve belgelerin kategorize edilmesi, büyük belge koleksiyonlarının özetlenmesi, kelimelerin ve anlamlarının aranması, eş anlamlı ve zıt anlamlı çiftlerin bulunması, eğitimsel uygulamalar, başarılı iş yönetimi için profesyonellerin yetkinlik desenlerini analiz etme, hasta ve hekim arasındaki iletişimi inceleyerek ilaç reçete etme süreçlerine kadar birçok uygulamada kullanılmaktadır. Gizli anlamsal analiz, metin, görüntü ve videolar için de uygulanabilir ve tıbbi kayıtların tanımlanmasında kullanılabilir (Ayesha vd., 2020).

Yerelliği Koruyan Projeksiyonlar

Yerelliği koruyan projeksiyonlar, verinin yerel komşuluk bilgisini olabildiğince iyi korumayı amaçlayan doğrusal bir yöntemdir. Bu yöntemde, d boyutlu ve n örnekli bir veri kümesi olan $X=[x_1,x_2,...,x_n]$ için n düğümden oluşan bir komşuluk grafiği (G) oluşturulur. Burada, x_i veri örneği, grafikteki i. düğüme karşılık gelir. Komşuluk grafiği oluşturulurken, bir veri örneğinin ε-komşuluğu veya k en yakın komşusu kullanılır. Birbirine komşu olan veri örneklerinin düğümlerini bağlayan kenarlar eklenir. Oluşturulan komşuluk grafiği temel alınarak simetrik nxn boyutunda bir ağırlık matrisi (W) elde edilir. Ağırlık matrisinin elemanları, komşu olan düğümler arasındaki bağların gücünü temsil eder ve iki farklı yolla belirlenebilir (Yıldız ve Sevim, 2016).

Bu yöntemle, verinin yerel yapısı korunarak boyut indirgeme gerçekleştirilir, böylece düşük boyutlu temsillerde bile verinin orijinal komşuluk yapısı korunmuş olur.

Bağımsız Bileşen Analizi

Bağımsız bileşen analizi, verilerin doğrusal bir koordinat sisteminde istatistiksel olarak birbirinden bağımsız hale getirilmesini sağlar. Bu yöntem, veriler arasındaki istatistiksel bağımlılığı azaltmak için doğrusal dönüşümler uygular. Bu dönüşümler sonucunda, veri boyutu azaltılır ve yeni bir öznitelik vektörü oluşturulur. Bağımsız bileşen analizinin temel

amacı, karışık verilerden bağımsız kaynak bileşenlerini ortaya çıkarmaktır. Bu, genellikle ses karışımlarını ayırmak, görüntü işleme veya sinyal analizi gibi alanlarda kullanılır (Türk, 2022).

Aşağıdaki denklem ile karışıklık vektörü elde edilir;

$$X = K_x V \qquad \text{(Denklem 6)}$$

K: İkinci derece kare matris.
X: Karışık pixel vektörü.
V: Kaynak pixel vektörü.

Doğrusal Diskriminant Analizi

Doğrusal diskriminant analizi, etiketlenmiş bir veri kümesini kullanarak doğrusal bir projeksiyon hesaplar. Veriler, C farklı sınıfa ayrılır ve her sınıfın ortalama değeri sıfır olarak kabul edilir. Doğrusal diskriminant analizi, veri kümesini düşük boyutlu bir alana yansıtmak için doğrusal bir dönüşüm matrisi W'yi hesaplar (Wang vd., 2016).

Doğrusal diskriminant analizinde, üç tür dağılım matrisi tanımlanır: sınıf içi dağılım, sınıflar arası dağılım ve toplam dağılım. Sınıf içi dağılım, aynı sınıf içindeki verilerin çeşitliliğini, sınıflar arası dağılım, farklı sınıflar arasındaki çeşitliliği ve toplam dağılım ise tüm veri kümesindeki genel çeşitliliği gösterir. Doğrusal diskriminant analizinin temel amacı, sınıflar arası ayrımı maksimize ederken, sınıf içi çeşitliliği minimize etmektir. Bu, sınıfları ayırt etmek ve veriyi düşük boyutlu bir temsile dönüştürmek için kullanılır.

Projeksiyon Takibi

Projeksiyon takibi, veriyi k boyutlu bir projeksiyona dönüştürerek, önceden tanımlanmış bir hedef fonksiyonunu maksimuma çıkarmayı hedefler. Projeksiyon takibi, $A = [a_1, a_2, \cdots, a_k] \in Rp \times k(k < p)$ gibi bir matris kullanarak, projeksiyon takibi indeksi olarak bilinen fonksiyonun değerini artırmaya çalışır. Bu indeks, projeksiyonun ilginçlik derecesini ölçer. Projeksiyon takibi, esnek bir yapıya sahiptir ve küme analizi, sınıflandırma

gibi desen tanıma görevlerinde kullanılabilir. Bu nedenle, Projeksiyon takibi, veri analizi ve keşfi için düşük boyutlu projeksiyonlar oluşturmak için etkili bir yöntemdir (Espezua vd., 2015).

4.8.2.2.Doğrusal Olmayan Boyut Azaltma Teknikleri

Doğrusal olmayan boyut azaltma, yüksek boyutlu verilerdeki gizli düşük boyutlu yapıları keşfetmeyi amaçlar. İnsan yüzleri, konuşma dalga formları, el yazısıyla yazılmış karakterler ve doğal dil gibi veriler üzerinde çalışırken, bu teknik özellikle önemlidir. Doğrusal boyut azaltma algoritmaları, verilerin doğrudan doğrusal yapısını çözme konusunda oldukça iyi performans gösterir. Örneğin, Temel Bileşen Analizi, yüksek boyutlu verilerin ana eksenlerini belirleyerek boyut sayısını azaltır. Ancak, bu tür yöntemler genellikle doğrusal ilişkilere dayanır, yani verilerin altında yatan yapılar düz bir çizgi boyunca düzenlenmiş gibi kabul edilir (Saxena vd., 2004).

Gerçekte, yüksek boyutlu veriler genellikle karmaşık, eğri veya burulmuş şekiller alabilir. Bu tür durumlarda, doğrusal boyut azaltma algoritmaları verilerin altında yatan desenleri tam olarak yakalayamaz. Örneğin, yüz ifadeleri veya el yazısı gibi veriler, birçok içsel değişkenin karmaşık etkileşimine bağlı olabilir ve doğrusal olmayan özellikler sergileyebilir.

Bu nedenle, doğrusal olmayan boyut azaltma teknikleri, verilerin karmaşık ve eğri yapısını anlamak için geliştirilmiştir. Bu teknikler, verilerin doğal eğrisel manifoldlarını bulmak için daha karmaşık matematiksel yöntemler kullanır. Örneğin, IZOMAP veya yerel doğrusal gömme, veri noktaları arasındaki yerel ilişkileri dikkate alarak, verilerin daha doğru bir temsilini elde etmek için çabalar. Böylece, doğrusal olmayan boyut azaltma teknikleri, geleneksel yöntemlerin göremediği karmaşık desenleri ve ilişkileri açığa çıkararak, yüksek boyutlu verilerdeki zenginliği ve çeşitliliği daha iyi yakalayabilir.

Kernel Temel Bileşen Analizi

Kernel Temel Bileşen Analizi (Kernel TBA), doğrusal olmayan veri setlerinin düşük boyutlu temsilini elde etmek için kullanılan bir yöntemdir. Geleneksel Temel Bileşen Analizi (TBA), yüksek boyutlu verilerdeki ana bileşenleri belirleyerek boyut azaltma işlemini gerçekleştirir.

Ancak bu yaklaşım, genellikle doğrusal ilişkileri temel alır ve karmaşık, eğri veya doğrusal olmayan yapıdaki verileri tam olarak yakalayamaz. Kernel Temel Bileşen Analizi, bu sorunu çözmek için geliştirilmiştir. Kernel TBA, doğrusal olmayan verileri yüksek boyutlu bir doğrusal uzaya dönüştürmek için bir çekirdek fonksiyonu kullanır. Bu çekirdek fonksiyonu, verileri doğrusal olmayan bir uzaydan doğrusal bir uzaya haritalar. Haritalama işlemiyle, doğrusal olmayan ilişkileri daha iyi temsil etmek mümkün hale gelir. Çekirdek fonksiyonu, yüksek boyutlu veri noktalarının iç çarpımlarını hesaplar ve bu çarpımlar bir çekirdek matrisi oluşturur (V.s ve Surendran, 2015).

Kernel Temel Bileşen Analizi, çekirdek matrisi üzerinde özdeğer ayrıştırması yaparak, boyut azaltma işlemini gerçekleştirir. Özdeğerler ve özvektörler, yüksek boyutlu verilerin temel yapısını temsil eder. Kernel TBA, bu özvektörlerden en önemlilerini seçerek verilerin düşük boyutlu bir temsilini oluşturur. Bu temsil, doğrusal olmayan ilişkileri ve karmaşık desenleri daha iyi yakalayarak, verileri anlamayı ve analiz etmeyi kolaylaştırır.

Kernel TBA'nın avantajı, doğrusal olmayan veri setlerinde karmaşık yapıları yakalama yeteneğidir. Bu, özellikle yüz tanıma, gen ekspresyon analizi, metin madenciliği ve diğer karmaşık veri alanlarında faydalı olabilir. Kernel TBA, doğrusal olmayan verileri doğrusal olarak ayrılabilir hale getirerek, daha iyi veri analizi ve makine öğrenimi modelleri geliştirmeye yardımcı olur.

Buradaki fikir, doğrusal olmayan verileri doğrusal olarak ayrılabilir hale geldiği daha yüksek boyutlu bir uzaya haritalamaktır. Doğrusal olmayan haritalama işlevine çekirdek fonksiyonu denir ve örneğinin bir x haritası,$x \rightarrow \phi(x)$ şeklindedir. Çekirdek fonksiyonu, ϕ altında örneklerin x imgelerinin nokta çarpımını hesaplar.

$$_K(x_i, x_j) = \phi(x_i)^T \phi(x)_j \qquad \text{(Denklem 7)}$$

Çok Boyutlu Ölçekleme

Çok boyutlu ölçekleme, veri noktaları arasındaki benzerlik veya mesafeyi korumaya yönelik bir doğrusal olmayan boyut azaltma tekniğidir. Kruskal ve Wish tarafından tanıtılan çok boyutlu ölçekleme, veri noktaları arasındaki çiftler arası mesafeleri düşük boyutlu bir uzayda olabildiğince doğru bir şekilde korumayı hedefler. Bu teknik, çok değişkenli analiz, keşifsel

analiz ve veri görselleştirme için yaygın olarak kullanılmıştır. Çok boyutlu ölçekleme, bir mesafe matrisinden yararlanır ve düşük boyutlu bir temsil oluştururken, bu matris içindeki çiftler arası mesafelerin gömülü uzayda da korunmasını sağlar. Bu amaçla, çok boyutlu ölçekleme bir "stress fonksiyonu" kullanarak, benzerlikler ile gömülü uzaydaki mesafeler arasındaki farkı minimize etmeyi amaçlar. Optimal çözüm, bu farkı en aza indiren düşük boyutlu koordinatları belirler. Bu nedenle, çok boyutlu ölçekleme, çiftler arası mesafeleri koruyarak verileri daha düşük boyutlu bir uzaya dönüştürür (Ayesha vd., 2020).

ISOMAP

ISOMAP, yüksek boyutlu verilerin doğrusal olmayan bir şekilde düşük boyutlu uzaya indirgenmesi için kullanılan bir algoritmadır. Doğrusal olmayan boyut azaltma yöntemleri, verilerin karmaşık yapısını doğrusal bir şekilde temsil etmekte zorlandığında kullanılır. ISOMAP, özellikle verilerin bir manifold boyunca düzenlendiği durumlarda etkilidir. Bu algoritma, Çok Boyutlu Ölçekleme (ÇBÖ) ile benzer prensiplere sahiptir, ancak temel farkı, veri noktaları arasındaki mesafeleri hesaplarken Öklid uzaklığı yerine eğriçizgisel uzaklık (geodesic distance) kullanmasıdır. Eğriçizgisel uzaklık, bir manifold üzerindeki iki nokta arasındaki en kısa mesafe olarak tanımlanır. Bu mesafe, Öklid uzaklığına göre daha karmaşık ve doğrudan hesaplanması daha zordur çünkü eğri bir yüzey boyunca en kısa yolu bulmayı gerektirir (Bilgin vd., 2008).

ISOMAP, eğriçizgisel uzaklıkları hesaplamak için önce bir komşuluk çizgesi oluşturur. Bu çizge, veri noktaları arasındaki bağlantıları temsil eder ve bu bağlantılar genellikle belirli bir mesafe veya en yakın komşular gibi bir kriterle belirlenir. Ardından, çizge üzerinde "en kısa yol" algoritmalarından biri (genellikle Dijkstra algoritması) kullanılarak iki nokta arasındaki eğriçizgisel uzaklık hesaplanır.

ISOMAP algoritmasının etkinliği, komşuluk çizgesini oluştururken kullanılan parametrelere bağlıdır. Örneğin, komşuluk kriteri olarak kaç komşunun dikkate alınacağı veya belirli bir mesafenin ne kadar olacağı gibi faktörler, algoritmanın sonuçlarını etkileyebilir. Yanlış seçilmiş parametreler, manifoldun doğru temsilini bozabilir veya gürültüye neden olabilir.

Son olarak, ISOMAP, bu eğri çizgisel uzaklıkları kullanarak, Çok Boyutlu Ölçeklemenin

ilkelerine dayalı olarak düşük boyutlu bir temsil oluşturur. Bu süreç, veri setinin genel yapısını ve içsel geometrisini korurken, aynı zamanda boyutları azaltmayı sağlar. ISOMAP, bu özelliği nedeniyle karmaşık veri setlerinde ve eğri yapıya sahip manifoldlarda etkili bir doğrusal olmayan boyut azaltma yöntemi olarak kabul edilir.

Yerel Doğrusal Gömme

Yerel Doğrusal Gömme (YDG) algoritması, başlangıçta üç boyutlu verilerde kullanılmıştır. Kim ve Finkhel, hiperspektral verileri retinal spektral soğurma eğrisiyle çarparak RGB görüntüler elde etmişlerdir. Retinal soğurma eğrisi, insan gözünün soğurduğu spektral ışınımları belirler. Bu veri setinde anomali tespit algoritmaları uygulanmış ve K-ortalama ile YDG algoritması birlikte kullanıldığında etkin sınıflandırma yapıldığı gözlemlenmiştir (Aydoğdu vd., 2015).

Yerel Doğrusal Gömme (YDG) algoritması, gözetimsiz öğrenme yöntemleri arasında yer alır ve yüksek boyutlu verileri daha düşük boyutlara indirger. Bu algoritmanın temel amacı, yerel topolojiyi koruyarak ve ilişkileri mümkün olduğunca bozmadan, veri setinin daha kompakt bir temsilini elde etmektir. YDG, doğrusal olmayan bir yöntemdir ve özvektör tabanlı bir teknik kullanarak çalışır.

Algoritma üç ana adımda uygulanır:

1. **Komşuluk Hesabı**: İlk adımda, X veri setindeki her bir veri noktasının diğer noktalara olan uzaklıkları hesaplanır. Uzaklık hesaplamak için genellikle Öklit uzaklığı kullanılır. Ardından, belirlenen bir komşuluk kriterine göre, her bir veri noktasının en yakın komşuları (k-en yakın komşu) belirlenir. Bu adım, verinin yerel yapısını anlamak için kritik öneme sahiptir.
2. **Ağırlık Matrisi Hesabı**: İkinci adımda, belirlenen komşuluk ilişkilerine göre bir ağırlık matrisi oluşturulur. Bu matris, her bir veri noktasının komşularına olan bağımlılığını ve ilişkilerini temsil eder. Ağırlık matrisinin oluşturulması, verinin yerel topolojisini korumak için önemlidir, çünkü bu matris, hangi noktaların birbirine daha yakın olduğunu belirler.

3. **Boyutu Azaltılmış Verinin Elde Edilmesi**: Son adımda, ağırlık matrisini kullanarak yüksek boyutlu verilerden düşük boyutlu bir temsil elde edilir. Bu adım, özvektörler ve özdeğerler yardımıyla gerçekleştirilir. Ağırlık matrisinin özdeğer ayrıştırması yapılarak, yüksek boyutlu verilerin temel özellikleri çıkarılır ve bu özelliklere dayalı olarak düşük boyutlu bir temsil oluşturulur.

YDG'nin başarısı, yerel komşulukların doğru bir şekilde belirlenmesine ve ağırlık matrisinin hassas bir şekilde oluşturulmasına dayanır. Bu yaklaşım, verinin yerel yapısını korurken boyut azaltmayı sağladığı için özellikle karmaşık veri setlerinde etkilidir. YDG, karmaşık ilişkilerin bulunduğu yüksek boyutlu verilerde etkili bir boyut azaltma yöntemi olarak kabul edilir ve çeşitli uygulamalarda, özellikle makine öğrenimi ve veri madenciliğinde yaygın olarak kullanılır.

Kendi Kendini Düzenleyen Haritalar

Kendi kendini düzenleyen haritalar (Self-Organizing Maps, SOM), sinir ağlarının bir türüdür ve girdileri düşük boyutlu, genellikle iki boyutlu bir düzleme indirgemek için kullanılır. 1980'lerde Teuvo Kohonen tarafından geliştirilmiş olan SOM, denetimsiz öğrenme yöntemleri arasında yer alır ve genellikle veri madenciliği, kümeleme ve veri görselleştirme uygulamalarında kullanılır.

Kendi kendini düzenleyen haritaların temel amacı, yüksek boyutlu veri setlerini düşük boyutlu b”ir düzleme projekte etmek ve bu projeksiyon sırasında girdilerin topolojik ilişkilerini korumaktır. Bu, benzer özelliklere sahip verilerin haritada birbirine yakın bir şekilde konumlanmasını sağlar. Kendi kendini düzenleyen haritaların bu yeteneği, özellikle karmaşık veri setlerinin anlaşılması ve görselleştirilmesinde faydalıdır.

Kendi kendini düzenleyen haritalar, Kohonen özellik haritası olarak da bilinir ve doğrusal olmayan projeksiyon gerçekleştiren bir denetimsiz sinir ağı türüdür. Kendi kendini düzenleyen haritaların çıktıları, gruplar veya kümeler şeklinde organize edilir. Bu ağ, verileri iki boyutlu bir ızgara üzerindeki bir dizi birime (Kendi kendini düzenleyen harita düğümleri) projekte etmek için kullanılır. Kendi kendini düzenleyen haritalarda her birim, bir ağırlık vektörüne sahiptir ve bu ağırlık vektörleri, girdilere göre kendilerini uyarlayarak veri

kümesinin yapısını yansıtır (Kim vd., 2014). Projeksiyon işlemi, kendi kendini düzenleyen haritalar düğümünün ağırlık vektörü ile giriş vektörü arasındaki Öklit uzaklığına dayanan bir aktivasyon fonksiyonuyla gerçekleştirilir. Veriler, en yakın ağırlık vektörüne sahip düğüme doğru çekilir ve böylece girdiler, kendi kendini düzenleyen haritalar ızgarası üzerinde benzer özelliklere sahip gruplar veya kümeler halinde organize edilir. Kendi kendini düzenleyen harita, girdiler arasındaki ilişkileri ve benzerlikleri yakalayarak verileri görselleştirmek ve analiz etmek için yaygın olarak kullanılır.

Bir kendi kendini düzenleyen haritalar, genellikle iki boyutlu bir ızgara üzerine yerleştirilen bir dizi düğümden oluşur. Her düğüm, bir "ağırlık vektörü" ile ilişkilendirilir, bu da girdilerle karşılaştırmak için kullanılır. Kendi kendini düzenleyen haritaların eğitimi sırasında, giriş verileri tekrarlı bir şekilde ağ üzerinden geçirilir. Her giriş için, Kendi kendini düzenleyen haritalardaki düğümlerden hangisinin girişe en yakın olduğu belirlenir. Buna "en iyi eşleşen birim" (Best Matching Unit, BMU) denir. BMU bulunduktan sonra, BMU'nun ağırlık vektörü ve onun yakın çevresindeki diğer düğümlerin ağırlık vektörleri, girdi yönünde güncellenir. Bu süreç, ağdaki düğümlerin girdilere daha iyi uyum sağlamasını sağlar.

Kendi kendini düzenleyen haritaların bir başka önemli özelliği, "komşuluk" etkisidir. BMU'nun sadece kendisi değil, aynı zamanda komşu düğümler de güncellenir, bu da girdilerin topolojik yapısının korunmasına yardımcı olur. Eğitim ilerledikçe, komşuluk etkisi genellikle azalır, böylece ağ daha hassas bir şekilde girdilere uyum sağlar.

Kendi kendini düzenleyen haritalar, verilerin görselleştirilmesi, kümeleme ve boyut azaltma gibi birçok uygulamada kullanılır. Özellikle karmaşık veri setlerini düşük boyutlu bir düzlemde temsil etme yeteneği, Kendi kendini düzenleyen haritaları veri analizi ve keşfi için güçlü bir araç yapar.

Vektör Nicelemeyi Öğrenme

Vektör Nicelemeyi Öğrenme (VNÖ), sınıflandırma ve kümeleme amacıyla kullanılan rekabetçi tabanlı bir sinir ağıdır. 1995'te Teuvo Kohonen tarafından tanıtılan bu teknik, sınıf bölgelerini temsil eden prototipleri öğrenir ve bu prototipler arasındaki hiper-düzlemlerle Voronoi bölmeleri oluşturur. VNÖ, denetimli bir öğrenme yöntemi olarak Kendi Kendini

düzenleyen haritalara benzer, ancak Kendi kendini düzenleyen haritalardan farklı olarak, yalnızca kazanan düğümü günceller ve bu nedenle çıktı uzayı topolojik olarak düzenlenmez (Nanga vd., 2021).

VNÖ, sezgisel ve kolay anlaşılır yapısıyla avantajlıdır ve karmaşık sınıflandırma yöntemlerine göre daha basit bir alternatif sağlar. Ayrıca, VNÖ'nün çeşitli versiyonları geliştirilmiştir, örneğin VNÖ1, VNÖ2, VNÖ3 ve bunların gelişmiş sürümleri. VNÖ, konuşma tanıma, kontrol desen tanıma ve sinyal işleme gibi çeşitli alanlarda uygulanır. Bununla birlikte, yavaş yakınsama ve kararsız davranış gibi bazı sınırlamaları da vardır. Yakınsama sorununu çözmek için genetik algoritmalar kullanılmış ve sınıflandırma performansını artırmak için diğer yaklaşımlar geliştirilmiştir.

VNÖ'nin kullanım alanları oldukça çeşitlidir. Örneğin, Gabor filtresi ile birlikte yüz ifadelerini tanımak için kullanılmış ve ECG sinyallerinden aritmiyi tanımak için VNÖ tabanlı yapay sinir ağları geliştirilmiştir. VNÖ, çok sınıflı problemlerle başa çıkabilir ve farklı türdeki veriler için çeşitli VNÖ varyantları geliştirilmiştir.

T-Stokastic Komşu Gömme

t-Stokastik Komşu Gömme (t-SNE), yüksek boyutlu verileri düşük boyutlu bir uzaya indirmek için kullanılan denetimsiz bir doğrusal olmayan boyut azaltma tekniğidir. Hinton ve Roweis tarafından tanıtılan t-SNE, verileri görselleştirmek ve doğrusal olmayan yapılarını anlamak için oldukça uygundur. Bu teknik, veri setlerindeki yerel yapıyı yakalarken, aynı zamanda küresel bir yapı da sunar (Ayesha vd., 2020).

t-SNE, yüksek boyutlu verilerin yerel ve küresel yapılarını koruyarak, veri setlerini iki veya üç boyutlu bir alanda görselleştirilebilir hale getirir. Bunu, yüksek boyutlu veri noktaları arasındaki koşullu olasılıkları hesaplayarak ve bu olasılıkları düşük boyutlu uzayda eşleştirerek yapar. Bu sayede t-SNE, karmaşık ve doğrusal olmayan veri setlerinin görsel temsilini elde etmek için güçlü bir araçtır.

4.9. Sınıf Dengesizliği

Birçok gerçek problem verilerinde, veri setleri genellikle dengesizdir; bazı sınıflar

diğerlerine göre çok daha fazla örnek içerir. Dengesizlik oranı oldukça büyük olabilir, bazen 10^6 kadar yüksek seviyelere ulaşabilir. Sınıf dengesizliği, sınıflandırıcı tasarımında önemli bir sorun olarak kabul edilir ve sınıflandırıcıların performansı üzerinde ciddi etkiler yaratır. Sınıf dengesizliğini dikkate almayan öğrenme algoritmaları, çoğunluk sınıfına odaklanıp azınlık sınıfını göz ardı etme eğilimindedir. Örneğin, büyük bir dengesizlik oranına sahip bir problemde, hata oranını minimize etmeye çalışan bir algoritma, tüm örnekleri çoğunluk sınıfı olarak sınıflandırarak düşük hata oranı elde edebilir, ancak bu durumda azınlık sınıfındaki tüm örnekler yanlış sınıflandırılmış olur (Liu vd., 2009).

Sınıf dengesizliği ile başa çıkmak için çeşitli yöntemler vardır. Örnekleme yöntemleri, eğitim setinde azınlık ve çoğunluk sınıflarının boyutlarını değiştirerek dengesizliği ele alır. Maliyet duyarlı öğrenme, farklı sınıflardaki hatalar için farklı maliyetler uygulayarak sınıf dengesizliği ile başa çıkar. Bu yöntemler, dengesiz veri setlerinden öğrenme sırasında ortaya çıkan zorlukları ele alır. Sınıf dengesizliğiyle başa çıkmak için kullanılan bazı teknikler, sınıf oranını dengeleme veya karar sınırlarını değiştirme üzerine odaklanır. Tüm bu yaklaşımlar, sınıflandırıcıların azınlık sınıfını doğru şekilde tanımasına yardımcı olmayı amaçlar.

Son yıllarda, teknolojik ilerlemeler ve veri üretimindeki patlama, sınıflama problemlerinde dengesiz sınıf dağılımı ve kayıp gözlemlerin belirgin bir zorluk haline gelmesine neden oldu. Bu durum, özellikle büyük çaplı akıllı uygulamalar, elektronik cihazlar arasındaki etkileşim ve sosyal medya gibi çeşitli kaynaklardan elde edilen veri setlerinin analizi için mevcut yaklaşımların yetersiz kaldığını ortaya çıkardı.

Bu değişen veri yapılarına ve çeşitliliğe yanıt olarak, istatistik tabanlı makine öğrenmesi yaklaşımları önem kazandı. Bu yaklaşımlar, büyük veri kümelerini işlemek ve analiz etmek için daha esnek ve etkili bir çerçeve sunar. Özellikle derin öğrenme gibi karmaşık algoritmalar, bu tür veri setlerindeki desenleri tespit etme yetenekleriyle dikkat çekmektedir.

Bu bağlamda, geleneksel makine öğrenmesi yöntemlerinin yanı sıra, derin öğrenme ve benzeri teknikler giderek daha yaygın hale gelmektedir. Bu teknikler, büyük miktarda veriye dayalı karmaşık problemleri çözmek için güçlü araçlar sağlar. Ancak, bu yaklaşımların kullanımıyla ilgili belirli zorluklar ve etik endişeler de bulunmaktadır, bu nedenle dikkatli bir şekilde uygulanmaları gerekmektedir.

Dengesiz sınıf dağılımı, veri setinde bulunan sınıfların sayısal olarak önemli ölçüde farklılık gösterdiği durumları ifade eder. Bu sorun, finansal alanda kredi kartı sahtekarlığı gibi, sağlık alanında kanser teşhisi gibi ve doğal dil işleme gibi çeşitli alanlarda karşılaşılan yaygın bir zorluktur.

Dengesiz sınıf dağılımını ele alan çalışmalar genellikle üç grupta incelenir:

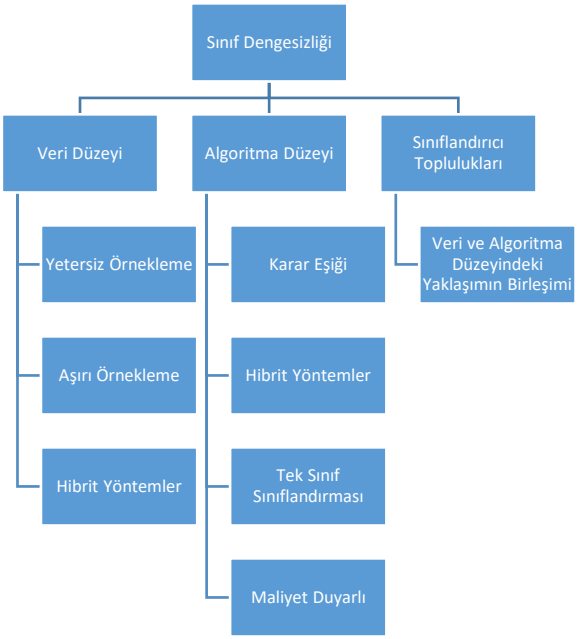
Veri Düzeyinde Yaklaşımlar**:** Bu yaklaşımlarda, veri setinin kendisi üzerinde değişiklikler yapılır. Rassal az örnekleme (random under sampling) ve rassal aşırı örnekleme (random over sampling) en yaygın olanlarıdır. Rassal az örneklemede, çoğunluk sınıfından örnekler rastgele çıkarılarak sınıflar arası denge sağlanırken, rassal aşırı örneklemede azınlık sınıfından örnekler çoğaltılarak denge sağlanır (Gümüştaş ve Pehlivanlı, 2023).

Algoritma Düzeyinde Yaklaşımlar: Bu yaklaşımlarda, sınıflandırma algoritmaları dengesizliği doğrudan ele alır. Örneğin, maliyet duyarlı sınıflandırma algoritmaları, farklı sınıfların hatalarının farklı maliyetlere sahip olduğunu dikkate alır.

Maliyete Duyarlı Yaklaşımlar: Bu yaklaşımlarda, sınıflar arasındaki maliyet dengesi göz önünde bulundurulur. Bu, özellikle nadir sınıfın daha fazla önem taşıdığı durumlarda kullanışlı olabilir.

Sentetik veri üretme, dengesiz sınıf dağılımını ele almak için veri düzeyinde kullanılan bir yaklaşımdır. En bilinen yöntemlerden biri, sentetik azınlık aşırı örnekleme yöntemi (SMOTE - Synthetic Minority Oversampling Technique) olarak adlandırılır. Bu yöntem, azınlık sınıfındaki gözlemleri sentetik olarak çoğaltarak dengesizliği azaltmaya çalışır.

Ancak, kullanılan yöntemlerin olumsuz yan etkilerine dikkat edilmelidir. Örneğin, rassal örnekleme yöntemleri bazı etkili gözlemleri kaybetme riski taşırken, sentetik veri üretme yöntemleri aşırı öğrenmeye yol açabilir. Bu nedenle, her bir yöntemin avantajları ve dezavantajları dikkate alınarak uygun bir yaklaşım seçilmelidir.



Şekil 4.3: Sınıf Dengesizliği

4.9.1. Yetersiz Örnekleme (Undersampling)

Undersampling yöntemleri, sınıf dengesizliği problemleriyle başa çıkmak için kullanılan bir grup tekniktir. Bu yöntemler, çoğunluk sınıfına ait örneklerin sayısını azaltarak, sınıflar arasındaki dengesizliği gidermeyi amaçlar. Bu azaltma işlemi, rastgele seçim yaparak veya belirli bir istatistiksel bilgi temelinde gerçekleştirilir (Shelke vd., 2017).

Rastgele undersampling, çoğunluk sınıfına ait örneklerin rastgele seçilerek azaltılmasıdır. Bu yöntem, veri setindeki dengesizliği doğrudan ele alır ve azınlık sınıfına ait örneklerle çoğunluk sınıfı arasındaki orantıyı düzeltir. Ancak, bu yöntemde bazı önemli bilgilerin kaybedilme riski vardır.

Bilgilendirilmiş undersampling ise, daha fazla bilgi kullanarak azaltma işlemini gerçekleştirir. Bu yöntemde, verinin istatistiksel özellikleri veya alan bilgisi gibi bilgilere dayanarak hangi örneklerin çıkarılacağına karar verilir. Bu sayede, azaltma işlemi daha kontrollü bir şekilde gerçekleştirilebilir ve önemli bilgilerin kaybedilme riski azaltılabilir.

Bazı bilgilendirilmiş undersampling yöntemleri ve iterasyon yöntemleri, veri setindeki çoğunluk sınıfına ait örnekleri daha da geliştirmek için veri temizleme tekniklerini

kullanabilir. Bu teknikler, gereksiz veya anormallik içeren örneklerin filtrelenmesi veya düzeltilmesi gibi adımları içerebilir. Bu şekilde, azaltma işlemi daha etkili hale getirilebilir ve sonuçlar daha dengeli bir şekilde elde edilebilir.

4.9.2. Aşırı Örnekleme (Oversampling)

Aşırı örnekleme yöntemi, veri kümesindeki sınıf dengesizliğini düzeltmek için kullanılan bir grup tekniktir. Bu yöntemler, özellikle azınlık sınıfına ait örneklerin sayısını artırarak, sınıflar arasındaki dengesizliği gidermeyi amaçlar (Shelke vd., 2017).

Rastgele aşırı örnekleme yönteminde, mevcut azınlık sınıfına ait örnekler basitçe kopyalanarak veya çoğaltılarak azınlık sınıfının boyutu artırılır. Bu yöntem, veri setindeki sınıf dengesizliğini hızlı bir şekilde düzeltebilir. Ancak, doğrudan kopyalama yapmak bazı durumlarda veri setindeki desenleri ve örüntüleri bozabilir, ayrıca aşırı örnekleme yapma işlemi nedeniyle modelin ezberleme eğilimi artabilir.

Sentetik aşırı örnekleme tekniğinde ise, azınlık sınıfına ait örnekler üzerinde yapay örnekler oluşturulur. Bu örnekler, mevcut azınlık örneklerinin özelliklerini temel alarak, gerçek veriye benzer yapılarda üretilir. Bu sayede, azınlık sınıfının boyutu artarken, veri setindeki desenlerin ve örüntülerin korunması hedeflenir. Sentetik örnekler, modelin azınlık sınıfını daha iyi anlamasını sağlar ve azınlık sınıfındaki örneklerin yanlış sınıflandırılmasını önler.

Her iki yöntemin de avantajları ve dezavantajları bulunmaktadır. Rastgele aşırı örnekleme hızlı ve basit bir çözüm sunarken, sentetik aşırı örnekleme daha karmaşık bir işlem olup, veri setindeki desenleri daha iyi koruyabilir. Hangi yöntemin kullanılacağı, veri setinin özelliklerine ve problemin gereksinimlerine bağlı olarak belirlenmelidir.

4.9.2.1. Hibrit Yöntemler

Hibrit yöntemler, aşırı örnekleme ve azaltma yöntemlerini bir araya getiren ve bunların avantajlarını birleştiren yaklaşımlardır. Hibrit yöntemler, genellikle sınıf dengesizliği problemlerini daha etkili bir şekilde çözmek için kullanılır.

Bir hibrit yöntem örneği, önce azınlık sınıfındaki örneklerin sayısını artırmak için sentetik aşırı örnekleme kullanılabilir. Daha sonra, çoğunluk sınıfındaki örnekler rastgele azaltılarak veya bilgi temelli azaltma yöntemleriyle düzeltilir. Bu yaklaşım, azınlık sınıfının boyutunu artırarak veri setindeki sınıf dengesizliğini azaltırken, çoğunluk sınıfındaki örneklerin sayısını da azaltarak modelin ezberleme eğilimini azaltabilir.

Hibrit yöntemlerin avantajları arasında, farklı veri setleri ve problemler için daha esnek ve özelleştirilmiş bir yaklaşım sunması yer alır. Ayrıca, farklı örneklemleme tekniklerinin birleştirilmesiyle daha dengeli ve güvenilir sonuçlar elde edilebilir.

Ancak, hibrit yöntemlerin karmaşıklığı ve uygulanması gereken ek hesaplama maliyetleri gibi dezavantajları da vardır. Bu nedenle, hangi hibrit yöntemin kullanılacağı, veri setinin özellikleri ve problem gereksinimlerine bağlı olarak dikkatlice değerlendirilmelidir.

4.9.3. Karar Eşiği (Threshold)

Karar eşiklerini taşıma, sınıf dengesizliğiyle başa çıkmak için kullanılan bir tekniktir. Geleneksel sınıflandırma modellerinde, çoğunlukla sınıf tahminini belirleyen bir karar eşiği bulunur. Bu eşik, genellikle %0.5 olarak varsayılır ve çoğunlukla sınıfı değiştirme kararı verir. Ancak, dengesiz veri kümelerinde, bu varsayılan eşik dengesiz sınıf dağılımına uygun olmayabilir (Collell vd., 2018).

Karar eşiklerini taşıma yöntemi, bu eşiği değiştirerek sınıf dengesizliğiyle başa çıkmayı amaçlar. Örneğin, azınlık sınıfının daha hassas bir şekilde tespit edilmesini sağlamak için eşik azaltılabilir, böylece azınlık sınıfına ait daha fazla örnek doğru bir şekilde tanımlanabilir. Benzer şekilde, çoğunluk sınıfının daha spesifik bir şekilde belirlenmesi gerekiyorsa eşik yükseltilebilir.

Bu yöntem, modelin çıktılarını değiştirmek yerine mevcut modelin çıktılarını daha dengeli bir şekilde yorumlamak için kullanılır. Bu nedenle, eğitilmiş bir modelin performansını iyileştirmek için mevcut veri seti üzerinde yapılan bir değişiklik değildir.

4.9.4. Tek Sınıf Sınıflandırması

Tek-sınıf sınıflandırması (TSS), nadiren bulunan veya nadiren görülen bir durumu tanımlamak için kullanılan bir tekniktir. Bu teknikte, sınıflandırıcı yalnızca tek bir sınıfı, yani "hedef sınıfı" öğrenir ve diğer sınıfı (negatif sınıfı) tanımaz veya göz ardı eder. Örneğin, bir güvenlik uygulamasında, normal kullanıcı davranışlarını modellemek için TSS kullanılabilir ve bu sayede anormal veya potansiyel tehditler daha iyi tanımlanabilir (Tsai ve Lin, 2021).

TSS'nin eğitim aşamasında, hedef sınıfa ait verilerle model oluşturulur. Bu veri, hedef sınıfın niteliklerini ve özelliklerini tanımlayan genellikle nadir olan örneklerden oluşur. Diğer sınıf ise ya eğitim verisi olarak kullanılmaz ya da çok az örnek içerir.

Sınıflandırma aşamasında, TSS yeni nesneleri sınıflandırmak için kullanılır. Bu nesneler ya hedef sınıfın yeni örnekleri ya da oluşturulan karar sınırının dışındaki bilinmeyen örneklerdir. TSS, her yeni nesne için bir "anormallik puanı" atar. Bu puan, nesnenin hedef sınıfa ne kadar uygun olduğunu gösterir ve belirli bir eşik değerine göre değerlendirilerek normal veya anormal olarak sınıflandırılır.

Geleneksel sınıflandırma problemlerinden farklı olarak, TSS yalnızca tek bir sınıfı tanır ve diğer sınıfı göz ardı eder. Bu, nadir veya anormal durumları tanımlamak için ideal bir yöntemdir. Ancak, bu yöntem, nadir sınıf verilerinin yetersizliği veya temsil edilmemesi gibi zorluklarla karşılaşabilir.

4.9.5. Maliyet Duyarlı Sınıflandırma

Geleneksel makine öğrenimi ve veri madenciliği uygulamalarında, sınıflandırıcılar genellikle yeni verilerle başa çıkarken yapacakları hataları en aza indirmek üzerine odaklanır. Ancak, gerçek dünya uygulamalarında, farklı hataların maliyetleri genellikle eşit değildir. Örneğin, tıbbi teşhis gibi bir alanda yanlış bir teşhisin maliyeti, yanlışlıkla sağlıklı bir kişiyi hasta olarak tanımlamanın maliyetinden çok daha büyük olabilir, çünkü bu tür bir hata yaşamsal sonuçlara neden olabilir (Zhou ve Liu, 2006).

Bu nedenle, maliyet duyarlı öğrenme kavramı önem kazanmıştır. Maliyet duyarlı öğrenme, farklı hataların farklı maliyetlerle ilişkilendirildiği durumları ele alır. Öğrenme algoritmaları, bu maliyetleri dikkate alarak model oluşturur ve hataların maliyetine göre optimize edilir.

Son yıllarda, maliyet duyarlı öğrenme yöntemlerinin geliştirilmesine rağmen, bazı yöntemlerin, özellikle karar ağaçları gibi, maliyet duyarlı hale getirilmesi konusunda daha fazla araştırma yapılmıştır. Ancak, maliyet duyarlı sinir ağlarının incelenmesi ve uygulanması konusunda daha az çalışma yapılmıştır. Bu, özellikle sinir ağlarının, örneğin ağırlıklı örneklerin kabul edilmesini gerektiren yöntemlerin doğrudan uygulanmasının zor olmasıyla ilgilidir.

Ayrıca, sınıf dengesizliği problemi de önemli bir konudur. Bu problem, bir sınıftaki örneklerin diğerine göre çok daha fazla olduğu veya sınıf dağılımının dengesiz olduğu durumları ifade eder. Sınıf dengesizliği, birçok uygulamada karşılaşılan bir sorundur ve özellikle öğrenme algoritmalarının performansını olumsuz yönde etkileyebilir. Maliyet duyarlı öğrenme, sınıf dengesizliği problemine bir çözüm olarak kabul edilir, çünkü bu yaklaşım, farklı maliyetlerle ilişkilendirilmiş hataları dikkate alarak model oluşturur ve optimize eder.

4.9.6. Veri ve Algoritma Düzeyindeki Yaklaşımın Birleşimi

Veri ve algoritma düzeyindeki yaklaşımların birleşimi, dengesiz sınıf dağılımı gibi karmaşık problemleri ele almak için kullanılan etkili bir stratejidir.

Veri Düzeyinde Yaklaşım: Veri düzeyindeki yaklaşımlar, veri setinin kendisinde değişiklikler yaparak dengesizlik problemini çözmeyi amaçlar. Bu yaklaşımlar genellikle azınlık sınıfını artırmak veya çoğunluk sınıfını azaltmak için kullanılır. Örneğin, azınlık sınıfın örneklerinin kopyalanması veya sentetik örneklerin oluşturulması gibi teknikler veri düzeyindeki yaklaşımlara örnektir.

Algoritma Düzeyinde Yaklaşım: Algoritma düzeyindeki yaklaşımlar, model eğitimi sırasında veya sonrasında kullanılan tekniklerdir. Örneğin, karar ağaçları gibi bazı algoritmaların maliyet duyarlı hale getirilmesi veya sınıflandırma eşiğinin ayarlanması gibi

teknikler algoritma düzeyindeki yaklaşımlara örnektir.

Bu iki yaklaşımın birleştirilmesi, daha güçlü ve daha esnek bir çözüm sağlar. Örneğin, veri düzeyindeki yaklaşımlarla azınlık sınıfını artırarak veya çoğunluk sınıfını azaltarak dengesizliği azaltabilirsiniz. Ardından, algoritma düzeyindeki yaklaşımları kullanarak, maliyet duyarlı bir sınıflandırıcı oluşturabilir veya karar eşiğini ayarlayarak daha iyi sonuçlar elde edebilirsiniz.

Birleşik yaklaşımlar, özellikle sınıf dengesizliği gibi karmaşık problemleri ele alırken, genellikle daha etkili sonuçlar sağlar. Bu yöntemler, makine öğrenimi ve veri madenciliği alanında yaygın olarak kullanılmaktadır ve gerçek dünya veri setlerindeki karmaşıklığı ele almak için önemli bir araç haline gelmiştir.

4.10. Anomali (Aykırı Değer)

Anomali, diğer veri örneklerinden önemli ölçüde farklı olan güçlü bir aykırı değer olarak tanımlanırken, zayıf bir aykırı değer verinin gürültüsü olarak kabul edilir. Aykırı değerler genellikle kötü niyetli faaliyetler, ortamdaki değişiklikler, sistem davranışı, hileli davranışlar, insan hatası, cihaz hatası, kurulum hatası, örnekleme hataları, veri girişi hatası veya basitçe popülasyonlardaki doğal sapmalar nedeniyle ortaya çıkar. Bu nedenle, her gerçek veri setinde aykırı değerlerin varlığını anlamak önemlidir (Kuızınıenė vd., 2022).

Farklı araştırmacılar ve yazılım kaynakları, aykırı değerleri tanımlamak için çeşitli terimler kullanır. Bu terimler arasında "aykırı değer tespiti", "yenilik tespiti", "anomali tespiti", "gürültü tespiti", "sapma tespiti" veya "istisna madenciliği" gibi farklı isimlendirmeler bulunur. Ancak, bu terimlerin hepsi, bir veri setindeki beklenmedik veya normal dağılımdan önemli ölçüde farklı olan noktaları ifade etmek için kullanılır.

Örneğin, "aykırı değer tespiti" terimi, bir veri setindeki nadir veya beklenmedik örnekleri belirlemek için kullanılırken, "anomali tespiti" terimi genellikle bir sistemin normal davranışından sapmaları tanımlamak için kullanılır. "Yenilik tespiti" ise, daha önce görülmemiş veya bilinmeyen desenleri veya örüntüleri tanımlamak için kullanılır. "Gürültü tespiti" terimi ise, veri setindeki rastgele veya anlamsız değişkenlikleri veya sapmaları

belirtir.

Bunların hepsi, aslında aynı temel kavrama işaret eder: bir veri setindeki belirgin ve beklenmedik örneklerin tanımlanması. Bu terimlerin kullanımı, genellikle disiplinler arası farklılıklar veya belirli uygulama alanlarındaki özel gereksinimlerden kaynaklanır. Ancak, temelde, hepsi veri madenciliği veya makine öğrenimi alanlarında aykırı değer tespiti problemine yönelik bir yaklaşımı ifade eder.

4.10.1. Anomali Tespit Yöntemleri

Anomalilerinin tespiti, yatırımcılar ve finans uzmanları için potansiyel kâr fırsatlarını ortaya çıkarabilir ve piyasaların etkinliğini değerlendirmede yardımcı olur. Anomalileri tespit etmek için kullanılan bazı başlıca yöntemler:

4.10.1.1. İstatistiksel Tabanlı Yaklaşım

İstatistiksel anormallik tespit sistemlerinin potansiyeli, bilinmeyen saldırıları tespit etme yeteneğidir. Anormallik tespit sistemleri, bir ağın veya sistemin normal davranışının bir modelini oluşturur ve bu normal profilden sapmaları tespit eder. Bu hem bilinen hem de bilinmeyen kötü amaçlı faaliyetleri tespit etmelerini sağlar (Marhas vd., 2012).

İstatistiksel tabanlı anomali tespiti, özellikle ağ güvenliği, siber güvenlik ve sistem izleme gibi alanlarda yaygın olarak kullanılır. Bu teknik, normalden sapmaları belirleyerek bilinen ve bilinmeyen tehditleri tespit etme yeteneği sağlar. Ancak, bu yöntemin başarılı olabilmesi için normal davranışın doğru bir şekilde modellemesi ve anormalliklerin doğru bir şekilde tespit edilmesi gerekmektedir. Bu nedenle, veri analizi ve istatistik bilgisinin yanı sıra, alan uzmanlığı ve deneyim de bu tekniklerin etkin bir şekilde uygulanmasında önemlidir.

4.10.1.2. Uzaklık, Yoğunluk Tabanlı Yaklaşım

Uzaklık ve yoğunluk tabanlı anomali tespit yaklaşımı, k-NN (k-en yakın komşu) tekniği üzerine kurulmuş ve sıklıkla kullanılan bir yöntemdir. Bu yaklaşım, normal veri noktalarının genellikle daha yoğun veya yakın komşuluklarda bulunduğu, ancak anomalilerin bu yoğun

gruplardan daha uzakta olduğu ve kendi yerel yoğun gruplarını oluşturabileceği prensibine dayanır (Kuızınıenė vd., 2022).

Bu yöntemi daha ayrıntılı bir şekilde açıklamak gerekirse, öncelikle her bir veri noktası için belirli bir uzaklık ölçütü kullanılarak komşulukları belirlenir. K-en yakın komşu (k-NN) tekniği, her bir veri noktasının k-en yakın komşusunu belirleyerek bir komşuluk yapısı oluşturur. Daha sonra, her bir veri noktasının yoğunluğu hesaplanır. Yoğunluk, bir veri noktasının k-en yakın komşusu içinde bulunan diğer veri noktalarının sayısına dayanarak belirlenir.

Normal veri noktaları genellikle yoğun bölgelerde bulunurken, anomaliler bu yoğun bölgelerden uzakta, daha seyrek bölgelerde bulunabilir. Bu nedenle, uzaklık ve yoğunluk tabanlı yaklaşım, normal ve anormal davranış arasındaki bu farklılıkları kullanarak anomalileri tespit etmeyi amaçlar.

Bu yaklaşım, özellikle veri setlerindeki anomali tespiti için etkili bir yöntem olarak kabul edilir. Ancak, doğru sonuçlar elde etmek için k-en değeri gibi parametrelerin dikkatli bir şekilde ayarlanması ve uzaklık veya yoğunluk ölçütünün uygun bir şekilde belirlenmesi önemlidir.

Bu yöntemin başlıca avantajları şunlardır:

- Denetimsiz öğrenme sırasında kullanılabilir,
- Çok boyutlu bir ortamda kolayca ölçeklenebilir,
- Verimli hesaplama özelliklerine sahiptir,
- Veri dağılımı hakkında bir ön varsayıma sahip değildir.

4.10.1.3. Sınıflandırma Tabanlı Yaklaşım

Sınıflandırma tabanlı tekniklerde, öncelikle bir model eğitimi yapılırken, veri seti içindeki örnekler normal ve anormal olarak etiketlenir. Bu etiketlenmiş veri seti, sınıflandırma algoritmasına beslenir ve model, normal ve anormal veriler arasındaki farkları öğrenir.

Model eğitildikten sonra, yeni veri örnekleri model üzerinden geçirilir ve anomali olup olmadığı belirlenir (Örnek vd., 2018).

Bu yöntemin avantajlarından biri, sınıflandırma algoritmalarının genellikle iyi tanımlandığı ve geniş bir yelpazedeki veri türlerinde kullanılabilmesidir. Ayrıca, bu yaklaşım denetimli öğrenme yöntemlerine dayandığı için, eğitim veri setindeki etiketli verilere dayanarak doğru sonuçlar elde etme potansiyeli yüksektir.

Ancak, bu yöntemin bazı dezavantajları da vardır. Özellikle, eğitim veri setindeki dengesizlikler ve nadir görülen anomalilerin sınıflandırılması zor olabilir. Ayrıca, bu yaklaşım, anomali tespitinde tam olarak bilinmeyen veya beklenmeyen desenleri tanımlamakta zorlanabilir.

Genel olarak, sınıflandırma tabanlı teknikler, veri madenciliği ve makine öğrenimi alanında yaygın olarak kullanılan ve etkili sonuçlar veren bir yaklaşımdır. Bu yöntem, doğru şekilde uygulandığında, farklı endüstrilerdeki çeşitli anomali tespiti problemlerine başarıyla uygulanabilir.

4.10.1.4. Kümeleme Tabanlı Yaklaşım

Çok değişkenli zaman serileri, birçok gerçek dünya uygulamasında karşılaşılan karmaşık veri yapılarıdır. Bu tür verilerdeki anormallikleri tespit etmek, genellikle zorlu bir görevdir çünkü zaman serilerinin çok boyutlu yapısı nedeniyle geleneksel yöntemlerin kullanımı zorlaşır. Ancak, kümeleme tabanlı yöntemler bu tür veri yapılarını analiz etmek ve anormallikleri tespit etmek için potansiyel olarak etkili bir yaklaşım sunar (Li vd., 2021).

Kümeleme tabanlı yöntemlerin genel bir tanımı, öncelikle zaman serisi verisinin alt dizilere ayrılmasıdır. Bu alt diziler, genellikle KMeans veya Fuzzy C-Means (FCM) gibi kümeleme algoritmaları kullanılarak oluşturulur. Daha sonra, her bir alt dizi, farklı kümelere atanır ve bu kümelerin merkezleri hesaplanır.

İkinci aşamada, her alt dizi, kümelerle ne kadar uyumlu olduğuna bağlı olarak bir anormallik puanı alır. Bu puanlar, genellikle alt dizilerin küme merkezlerine olan uzaklıklarına

dayanarak hesaplanır. Alt diziler, kümelerle daha az uyumluysa, yani merkezlere daha uzaksa, daha yüksek bir anormallik puanına sahip olabilirler.

Bu yaklaşımda, herhangi bir zaman serisi veri setindeki yapısal özellikler veya desenler önceden bilinmesi gerekmez. Bu nedenle, bu yöntemler genellikle veriye bağlı olmayan ve genel uygulamalarda kullanılabilen esnek bir yaklaşım sunarlar.

Ancak, kümeleme tabanlı yöntemlerin bazı dezavantajları da vardır. Özellikle, bu yöntemlerin performansı, oluşturulan kümelerin kalitesine ve veri setindeki yapıya bağlı olabilir. Ayrıca, çok büyük veya yüksek boyutlu veri setlerinde zaman ve mekansal karmaşıklık gibi sorunlar ortaya çıkabilir.

Sonuç olarak, kümeleme tabanlı yöntemler, çok değişkenli zaman serilerindeki anormallikleri tespit etmek için güçlü bir araç olabilir. Ancak, uygulanacak veri setine ve problem bağlamına bağlı olarak, bu yöntemlerin avantajları ve dezavantajları dikkate alınmalıdır.

4.10.1.5. Bilgi Teorisi Temelli Yaklaşım

Bilgi teorisi temelli yaklaşım, anormallikleri tespit etmek için veri setindeki düzensizlikleri ve bilgi içeriğini analiz etmeye dayanır. Bu yaklaşımın ana varsayımı, anormal durumların normalden farklı bilgi içeriği veya düzensizlikler oluşturduğudur (Kuızınıenė vd., 2022).

Bilgi içeriği, temel olarak bir olayın beklenmeme olasılığını veya belirsizliği ölçer. Bu kavram, Shannon'ın entropi ölçüsüyle temsil edilir. Entropi, bir olayın olasılık dağılımının ne kadar tahmin edilemez olduğunu belirtir. Yüksek entropi değerleri, bir veri setindeki belirsizliğin veya düzensizliğin arttığını gösterebilir ve potansiyel olarak anormalliklerin varlığını işaret edebilir.

Bilgi teorisi temelli yaklaşımda, göreli entropi, koşullu entropi, bilgi kazancı ve bilgi maliyeti gibi farklı bilgi teorisi ölçümleri kullanılarak veri setinin yapısı incelenir. Göreli entropi, iki olasılık dağılımı arasındaki farkı ölçer ve belirsizlik azaldıkça birbirine yaklaşır. Koşullu entropi ise bir olayın diğer bir olay hakkında bilgi sağlayıp sağlamadığını ölçer.

Bilgi kazancı, bir özniteliğin bilgi içeriğindeki azalmayı ölçer ve anormal durumların belirlenmesinde önemli bir rol oynar. Bilgi maliyeti ise bir özelliğin bilgi kazancına göre değerini ölçer ve bilgi kazançlarının dengelenmesine yardımcı olur.

Bu ölçümler, veri setindeki düzensizlikleri ve anormal durumları tespit etmek için kullanılır. Daha yüksek bilgi içeriği veya daha yüksek belirsizlik, potansiyel olarak anormal durumların varlığına işaret edebilir ve bu nedenle bu ölçümler, anormallik tespitinde önemli bir rol oynar.

4.10.1.6. Topluluk Tabanlı Yaklaşım

Tabii, kombinasyon tabanlı yaklaşım veya topluluk tabanlı yöntem, birkaç farklı makine öğrenimi algoritmasının sonuçlarını birleştirerek anormallikleri tespit etmeye çalışır. Bu yöntem, genellikle ağırlıklı oylama veya çoğunluk oylama gibi tekniklerle algoritmaların çıktılarını birleştirir.

Aggarwal (2017) tarafından belirtilene göre, topluluk tabanlı aykırı değer tespit algoritmaları genellikle sıralı ve bağımsız topluluklar olmak üzere iki kategoriye ayrılabilir. Sıralı topluluklar, oluşturulan sıralı algoritmaların verilere bağlı olarak kullanılmasını içerirken, bağımsız topluluklar farklı algoritmaların çıktılarını oylayan bir kombinasyon kullanır.

Bu yöntemin temel avantajları arasında, performansın daha etkili olması, tahmin edilen sonuçların daha istikrarlı olması ve yüksek boyutlu veriler ve veri akışları için uygulanabilir olması bulunur. Kombinasyon tabanlı yaklaşım, farklı algoritmaların güçlü yanlarını bir araya getirerek anormallik tespitinde daha güvenilir sonuçlar elde etmeyi amaçlar.

4.11. Performans Ölçütleri

Günümüzde yapay zekâ (YZ) teknolojileri, birçok alanda hayatımıza derinlemesine nüfuz etmiştir. YZ modelleri, görüntü tanıma, doğal dil işleme, oyun stratejileri ve daha pek çok alanda insan benzeri veya hatta insanı aşan performans sergilemektedir. Ancak, bir YZ modelinin başarısını değerlendirmek ve karşılaştırmak için doğru metrikleri kullanmak önemlidir.

YZ performansını ölçmek için kullanılan metrikler, genellikle modelin türüne ve çözümlediği problem türüne bağlıdır. Sınıflandırma problemleri için hassasiyet, geri çağırma ve doğruluk gibi metrikler yaygın olarak kullanılırken, regresyon problemleri için ortalama mutlak hata (MAE) ve ortalama kare hata (MSE) önemlidir. Ayrıca, sınıflandırma modellerinin performansını değerlendirmek için ROC eğrisi ve alan altındaki alan (AUC) gibi metrikler de sıkça kullanılmaktadır.

Ancak, performans ölçütlerinin seçimi, çözülmek istenen spesifik problem ve kullanım senaryosuna bağlıdır. Örneğin, bir tıbbi teşhis sistemi için hassasiyet ve geri çağırma çok önemli olabilirken, bir görüntü tanıma uygulamasında doğruluk daha kritik olabilir.

Ayrıca, YZ modellerinin performansını değerlendirirken genellikle hesaplama zamanı, karmaşıklık ve genelleme yeteneği gibi faktörler de göz önünde bulundurulmalıdır. Modelin eğitim ve tahmin süreleri, gerçek dünya uygulamalarında kullanılabilirliği etkileyebilirken, modelin yeni ve görülmemiş verilerle ne kadar iyi performans gösterdiği de önemlidir.

Sonuç olarak, yapay zekâ performansını ölçmek karmaşık bir süreçtir ve doğru metrikleri seçmek önemlidir. Doğru metriklerin seçilmesi, modelin gerçek dünya problemlerine daha etkin bir şekilde uyarlanmasına ve karar verme süreçlerine değer katmasına yardımcı olabilir. Bu nedenle, YZ performansını değerlendirirken dikkatli bir yaklaşım benimsemek gerekmektedir.

4.11.1. Doğruluk (Accuracy)

Doğruluk, tahmin edilen Remaining Useful Life (RUL - Kalan Kullanılabilir Ömür) değerinin, gerçek RUL değerine ne kadar yakın olduğunu ölçer. Ancak, bu metrik genellikle tüm hataları eşit olarak değerlendirir. Bu durumda, sistem ömrünün sonuna doğru yapılan tahminlerin, bakım planlaması ve lojistik için daha kritik olabileceği düşünülür. Bu nedenle, bu tahminlerin doğruluğuna daha fazla ağırlık vermek amacıyla bazı metriklerde değişiklikler yapılabilir (Ochella ve Shafiee, 2020).

Bu bağlamda, zamanında doğru tahminler yapmanın önemi ortaya çıkar. Erken ve doğru

tahminler, bakım planlaması ve lojistik süreçlerinin daha verimli olmasına yardımcı olabilir. Bu nedenle, tahminlerin zamanlaması da bir performans ölçütü olarak göz önünde bulundurulmalıdır.

Sonuç olarak, tahmin doğruluğunu artırmak ve zamanında tahmin yapabilme yeteneğini geliştirmek, bakım planlama ve lojistik süreçlerini optimize etmek için kritik öneme sahiptir. Bu nedenle, bu iki faktörü dikkate alan ölçütler ve değerlendirme yöntemleri geliştirilmelidir.

$$Doğruluk = \frac{Doğru\ Tahminler}{Toplam\ Veri\ Sayısı} \quad \text{(Denklem 8)}$$

4.11.2. Hassasiyet (Precision)

Hassasiyet, alınan örnekler arasında ne kadarının ilgili olduğunu gösterir ve doğru bir şekilde sınıflandırılan örneklerin, o sınıfa atanmış tüm örneklerin oranı olarak hesaplanır. Hassasiyet, [0, 1] aralığında sınırlıdır, burada 1 tüm örneklerin sınıfa doğru şekilde tahmin edildiğini, 0 ise sınıftaki hiçbir doğru tahmini temsil eder (Hicks vd., 2022).

Hassasiyet, bir sınıflandırma modelinin pozitif olarak tahmin ettiği örneklerin gerçekte ne kadarının pozitif olduğunu ölçer. Yani, hassasiyet ne kadar "keskin" veya "doğru" tahminler yapıldığını gösterir. Hassasiyet, aşağıdaki formülle hesaplanır:

$$Keskinlik = \frac{Gerçek\ Pozitif}{Gerçek\ Pozitif + Yanlış\ Pozitif} \quad \text{(Denklem 9)}$$

Bu formülde, "gerçek pozitif" sınıflandırılan örneklerin gerçekte pozitif olanların sayısını, "yanlış pozitif" ise pozitif olarak tahmin edilen ancak gerçekte negatif olan örneklerin sayısını temsil eder.

Hassasiyet, bir sınıflandırma modelinin ne kadar "keskin" veya "hassas" olduğunu gösterir. Yüksek hassasiyet değeri, modelin pozitif olarak tahmin ettiği örneklerin gerçekten pozitif olduğu anlamına gelir.

Kesinlik (Precision) ve doğruluk (Accuracy) metrikleri, model performansını

değerlendirirken kullanılan önemli ölçütlerdir. Doğruluk, tüm tahminlerin doğruluğunu yansıtırken, kesinlik yalnızca pozitif tahminlerin doğruluğunu ölçer. Bu nedenle, model performansını değerlendirirken hangi metriğin kullanılacağı, problem bağlamına ve gereksinimlere bağlı olarak değişir.

4.11.3. Duyarlılık (Recall)

Duyarlılık (recall), sınıflandırma modellerinde kullanılan önemli bir performans metriğidir. Ayrıca "Hatırlatma" veya "Doğru Pozitif Oranı" (TPR) olarak da adlandırılır. Duyarlılık, doğru olarak sınıflandırılan pozitif örneklerin oranını belirler. Bu, modelin pozitif sınıfı ne kadar başarıyla tahmin ettiğini ölçer (Hicks vd., 2022).

Duyarlılık, aşağıdaki formülle hesaplanır:

$$Dıyarlılık = \frac{Gerçek\ Pozitif}{Gerçek\ Pozitif + Yanlış\ Negatif} \quad \text{(Denklem 10)}$$

Burada, "gerçek pozitif", gerçekten pozitif olan örneklerin doğru bir şekilde tahmin edilen sayısını, "yanlış negatif" ise gerçekte pozitif olan ancak model tarafından yanlışlıkla negatif olarak tahmin edilen örneklerin sayısını temsil eder.

Duyarlılık (Recall), sınıflandırma modellerinde kullanılan bir performans metriğidir ve gerçek pozitiflerin doğru bir şekilde tahmin edilme oranını ölçer. Doğruluk (Accuracy) ise tüm tahminlerin doğruluğunu değerlendirirken, duyarlılık yalnızca gerçek pozitifleri değerlendirir. Bu nedenle, nadir pozitif sınıfların doğru bir şekilde tanınması gereken durumlarda, duyarlılık metriği daha önemlidir.

4.11.4. F1 Skor

Bu metrik, hem kesinlik (precision) hem de duyarlılık (recall) ölçümlerini dikkate alarak bir algoritmanın performansını değerlendirmek için kullanılır. Özellikle, dengesiz sınıflandırma problemlerinde ve sınıflar arasında belirli bir dengenin sağlanması gereken durumlarda yaygın olarak kullanılır.

Matematiksel olarak, f-skor (f-score veya f-measure), kesinlik ve duyarlılık arasındaki harmonik ortalama ile hesaplanır. Bu, hem kesinliğin hem de duyarlılığın göreceli olarak yüksek olduğu durumları teşvik eder.

F-skor, aşağıdaki formülle hesaplanır:

$$F1\ Skor = \frac{2x(Keskinlik*Duyarlılık)}{Keskinlik+Duyarlılık} \quad \text{(Denklem 11)}$$

F1-Skoru, sınıflandırma modellerinin performansını değerlendirmek için kullanılan bir metriktir. Bu metrik, kesinlik (Precision) ve duyarlılık (Recall) arasındaki dengeyi ölçer. Kesinlik, pozitif olarak tahmin edilen örneklerin gerçekte ne kadarının pozitif olduğunu gösterirken, duyarlılık, gerçek pozitiflerin ne kadarının doğru bir şekilde tahmin edildiğini ölçer. F1-Skoru, bu iki metrik arasındaki dengeyi temsil eder ve modelin pozitif tahminlerinin doğruluğu ile eksikliği arasındaki dengeyi değerlendirir. Bu nedenle, sınıflandırma problemlerinde yaygın olarak kullanılan önemli bir performans metriği olarak kabul edilir.

4.11.5. Karışıklık Matrisi (Confusion Matrix)

Karışıklık matrisi, bir sınıflandırma modelinin performansını değerlendirmek için kullanılan bir tablodur. Bu matris, gerçek değerlerin bilindiği bir dizi test verisi üzerinde modelin tahminlerini karşılaştırarak oluşturulur (Aşçı, ve diğerleri, 2022). Genellikle 2x2 veya daha büyük boyutlarda olabilir. 2x2 Karışıklık Matrisi, iki sınıflı (binary) sınıflandırma problemleri için kullanılır. Matrisin satırları gerçek sınıfları, sütunları ise modelin tahmin ettiği sınıfları temsil eder.

Bir 2x2 Karışıklık Matrisinde kullanılan metrikler şunlardır:

- **Doğru Pozitif (TP - True Positive):** Gerçek pozitif sınıf örneklerinin doğru bir şekilde pozitif olarak tahmin edildiği durumların sayısı.
- **Yanlış Negatif (FN - False Negative):** Gerçek pozitif sınıf örneklerinin yanlışlıkla negatif olarak tahmin edildiği durumların sayısı.

- **Yanlış Pozitif (FP - False Positive):** Gerçek negatif sınıf örneklerinin yanlışlıkla pozitif olarak tahmin edildiği durumların sayısı.
- **Doğru Negatif (TN - True Negative):** Gerçek negatif sınıf örneklerinin doğru bir şekilde negatif olarak tahmin edildiği durumların sayısı.

Bu metrikler genellikle modelin performansını değerlendirmek için kullanılır. Özellikle kesinlik (precision), duyarlılık (recall) ve F1 skoru gibi diğer performans metriklerinin hesaplanmasında temel olarak kullanılırlar. Karışıklık matrisi, modelin hangi sınıflar için daha iyi veya daha kötü performans gösterdiğini anlamak için önemli bir araçtır.

4.11.6. ROC Eğrisi ve AUC

ROC eğrisi (Receiver Operating Characteristic Curve), sınıflandırma modellerinin performansını değerlendirmek için kullanılan önemli bir araçtır, özellikle dengesiz veri kümeleri için idealdir. ROC eğrisi, doğru pozitif oranının (TPR) yanlış pozitif oranına (FPR) karşı çizildiği bir grafiktir (Keskenler vd., 2021).
Doğru pozitif oranı (TPR), gerçek pozitiflerin doğru bir şekilde tanımlandığı oranı ifade eder. Yani, TPR, gerçek pozitiflerin toplam sayısına oranla doğru şekilde pozitif olarak tahmin edilenlerin oranını ifade eder.

Yanlış pozitif oranı (FPR), gerçekte negatif olan ancak yanlışlıkla pozitif olarak tahmin edilen örneklerin oranını ifade eder. Bu, negatif sınıfa ait örneklerin toplam sayısına oranla yanlış şekilde pozitif olarak tahmin edilenlerin oranını temsil eder.

ROC eğrisi, farklı sınırlayıcı (threshold) değerlerine göre TPR ve FPR'nin değişimini gösterir. Bu eğri, bir sınıflandırma modelinin hassasiyetinin ve özgüllüğünün (1 - FPR) bir dizi farklı noktada nasıl değiştiğini görselleştirir.

ROC eğrisinin altındaki alan (AUC - Area Under the Curve), sınıflandırma modelinin performansını tek bir sayısal değerle ölçer. AUC değeri 0 ile 1 arasında değişir ve 1'e ne kadar yakınsa, modelin performansı o kadar iyidir. Bir modelin AUC değeri 0.5'e yakınsa, rastgele sınıflandırma ile aynı performansı sergiler. AUC değeri 0'dan küçükse, modelin performansı rastgele sınıflandırmadan daha kötüdür. Bu nedenle, AUC değeri, bir

sınıflandırma modelinin performansını değerlendirmek için yaygın olarak kullanılan önemli bir ölçüttür.

4.11.7. Logaritmik Kayıp (Log Loss)

Log kaybı metriği, bir olasılık aralığında doğru veya yanlış tahmin etmenin gerektiği durumlarda kullanılır (Manzali vd., 2017). Bu metrik, kesinlikle doğru (1) olma olasılığından kesinlikle yanlış (0) olma olasılığına kadar olan olasılıklarla çalışır. Hatanın logaritması alınarak kullanılması hem emin olup yanlış olmanın hem de yanlış olmanın aşırı cezalandırılmasını sağlar. Bu metrik, özellikle kesin olarak doğru tahminler yapmaya çalışırken yapılacak yanlış tahminlerin ağırlığını vurgular.

Logaritmik kayıp, bir modelin olasılıklı tahminlerini değerlendirir. Özellikle, bir sınıfın doğru olma olasılığını ölçmek için kullanılır. Model, bir örneğin belirli bir sınıfa ait olma olasılığını çıkarırken, logaritmik kayıp bu tahminin doğruluğunu ölçer.

Formül olarak, logaritmik kayıp şu şekilde hesaplanır:

$$LogLoss = -\frac{1}{n}\sum_{i=0}^{n}[y_i \log(y\hat{}_i) + (1 - y_i)\log(1 - y\hat{}_i)] \quad \text{(Denklem 12)}$$

Burada:

- n veri noktalarının toplam sayısını,
- y_i gerçek sınıf etiketlerini (0 veya 1),
- $y\hat{}_i$ tahmin edilen olasılıkları (0 ile 1 arasında),
- log doğal logaritmayı temsil eder.

Bu formül, her bir veri noktası için gerçek sınıf etiketlerini ve modelin tahmin ettiği olasılıkları alarak logaritmik hataları hesaplar ve bu hataların ortalamasını alır.

Logaritmik kayıp, modelin tahminlerinin doğruluğunu ölçerken, aynı zamanda aşırı

tahminlerin de cezalandırılmasını sağlar. Özellikle, kesin olarak doğru veya kesin olarak yanlış tahminler, büyük hatalara neden olur ve logaritmik kayıp bu durumları vurgular. Bu nedenle, logaritmik kayıp, sınıflandırma modellerinin performansını objektif bir şekilde değerlendirmek için yaygın olarak kullanılan bir ölçüdür.

4.11.8. Ortalama Mutlak Hata (Mean Absolute Error)

Ortalama mutlak hata (Mean Absolute Error - MAE), sürekli değişkenler arasındaki farkın ölçüsüdür ve regresyon ve zaman serisi problemlerinde yaygın olarak kullanılır (Kocabaş vd., 2022). Bu metrik, bir tahmin modelinin doğruluğunu değerlendirmek için önemli bir araçtır.

Ortalama mutlak hata, gerçek değerler ile tahmin edilen değerler arasındaki mutlak farkların ortalamasıdır. Bu, her bir veri noktası için gerçek değer ile modelin tahmin ettiği değer arasındaki farkın mutlak değerinin alınması ve bu farkların ortalamasının alınması demektir. Sonuç olarak, ortalama mutlak hata her bir veri noktası arasındaki farkın büyüklüğünü ölçer ve bu farkların ortalamasını sağlar.

Ortalama mutlak hatanın bir avantajı, kolay yorumlanabilir olmasıdır. Her bir veri noktasının gerçek değer ile tahmin edilen değer arasındaki farkın mutlak değerinin ortalaması olduğu için, ortalama mutlak hatanın birimleri orijinal verilerle aynıdır ve bu nedenle kolayca anlaşılabilir.

Ayrıca, ortalama mutlak hatanın, her veri noktasının en iyi uyum sağlayan çizgiye olan dikey mesafesinin ortalama değeri olduğu da söylenebilir. Bu nedenle, Ortalama Mutlak Hata, her bir veri noktasının tahmin edilen değer ile en iyi uyum sağlayan çizgiye ne kadar yakın veya uzak olduğunu ölçer.

Sonuç olarak, ortalama mutlak hata, sürekli değişkenler arasındaki farkın ölçüsü olduğu için regresyon ve zaman serisi problemlerinde yaygın olarak kullanılan bir performans metriğidir. Daha düşük bir ortalama mutlak hata değeri, modelin daha iyi bir tahmin yapma yeteneğine sahip olduğunu gösterirken, daha yüksek bir ortalama mutlak hata değeri, modelin daha kötü tahminler yaptığını gösterir.

4.11.9. R-Kare (R-Squared)

R-kare değeri, bir regresyon modelinin uyumunu ölçmek için kullanılan istatistiksel bir ölçüdür. Bir bağımlı değişkenin, bir veya daha fazla bağımsız değişken tarafından açıklanabilirliğini ifade eder (Özgür vd., 2023). R-kare değeri, bağımsız değişkenlerin bağımlı değişken üzerindeki toplam varyansın ne kadarını açıkladığını belirler.

R-kare değeri 0 ile 1 arasında değişir. Bir modelin R-kare değeri 1'e ne kadar yakınsa, o kadar iyi uyum sağladığı ve bağımsız değişkenlerin bağımlı değişken üzerindeki varyansın büyük bir kısmını açıkladığı anlamına gelir. Öte yandan, R-kare değeri 0'a ne kadar yakınsa, modelin uyumu o kadar zayıftır ve bağımsız değişkenlerin bağımlı değişken üzerindeki varyansın açıkladığı oran düşüktür.

4.11.10. Doğruluk Çapraz Doğrulama (Accuracy Cross-Validation)

Çapraz doğrulama (cross-validation) teknikleri, bir makine öğrenimi modelinin performansını değerlendirmek için yaygın olarak kullanılan ve güvenilir sonuçlar elde etmeye yardımcı olan önemli bir yöntemdir. Bu tekniklerden biri olan k-kat çapraz doğrulama (k-fold cross-validation), veri setini k kadar eşit parçaya böler ve her bir parçayı sırayla test verisi olarak kullanırken diğer parçaları eğitim verisi olarak kullanır. Bu işlem, her bir parçanın test verisi olarak kullanılmasıyla k kez tekrarlanır ve modelin doğruluğunun ortalaması alınır (Ateş ve Şenol, 2021).

Örneğin, 5-kat çapraz doğrulama kullanıyorsak, veri seti beş eşit parçaya bölünür. Her bir iterasyonda, bir parça test verisi olarak seçilirken, geri kalan dört parça eğitim verisi olarak kullanılır. Model eğitildikten sonra, test verisi üzerinde performans ölçülür ve bu işlem k kez tekrarlanır. Sonuç olarak, her bir iterasyonda elde edilen performans ölçümlerinin ortalaması alınarak modelin genel doğruluğu belirlenir.

Bir diğer çapraz doğrulama yöntemi olan "leave-one-out cross-validation" ise her bir veri noktasını sırayla test verisi olarak seçerken, geriye kalan verileri eğitim verisi olarak kullanır. Bu işlem, her bir veri noktası için tekrarlanır ve her bir veri noktasının test verisi olarak seçilmesiyle elde edilen performans ölçümlerinin ortalaması alınarak modelin genel

doğruluğu belirlenir.

Bu çapraz doğrulama teknikleri, bir modelin aşırı uyuma (overfitting) veya veriye özel (data-specific) olma gibi problemleri tespit etmek için oldukça faydalıdır. Ayrıca, modelin performansını daha güvenilir bir şekilde değerlendirmek için kullanılırlar.

5. UYGULAMA

Çalışmada BİST Ulusal Pazardan 14, Yakın İzleme Pazarından 14 şirket olmak üzere toplam 28 şirketin 2013/2022 dönemleri arası 43 adet finansal oranları kullanılmıştır.

5.1. Veri ve Yazılım

Çalışmanın veri seti aşağıda yer verilen finansal tabloları kullanılan ulusal pazardaki ve yakın izleme pazarındaki firmaların finansal verileri kullanılarak elde edilen finansal oranlardan oluşmaktadır;

Tablo 5.1: Çalışmada test edilen şirketler

	Ulusal Pazar'daki Firmalar		Yakın İzleme Pazarındaki Firmalar
1	Aksa Akrilik Kimya Sanayi A.Ş.	**1**	Atlantis Yatırım Holding A.Ş.
2	Aselsan Elektronik Sanayi ve Ticaret A.Ş.	**2**	Birko Birleşik Koyunlulular Mensucat Ticaret ve Sanayi A.Ş.
3	Anadolu Isuzu Otomotiv Sanayi ve Ticaret A.Ş.	**3**	Birlik Mensucat Ticaret ve Sanayi İşletmesi A.Ş.
4	Ereğli Demir ve Çelik Fabrikaları Ticaret A.Ş.	**4**	Casa Emtia Petrol Kimyevi ve Türevleri Sanayi Ticaret A.Ş.
5	Ford Otomotiv Sanayi A.Ş.	**5**	Cosmos Yatırım Holding A.Ş.
6	Gübre Fabrikaları Ticaret A.Ş.	**6**	Diriteks Diriliş Tekstil Sanayi ve Ticaret A.Ş.
7	Otokar Otomotiv ve Savunma Sanayi A.Ş.	**7**	Ekiz Kimya Sanayi ve Ticaret A.Ş.
8	Sasa Polyester Sanayi A.Ş.	**8**	Eminiş Ambalaj Sanayi ve Ticaret A.Ş.
9	Türkiye Şişe ve Cam Fabrikaları A.Ş.	**9**	Kervansaray Yatırım Holding A.Ş.
10	Tofaş Türk Otomobil Fabrikası A.Ş.	**10**	Kuvva Gıda Ticaret ve Sanayi Yatırımları A.Ş.
11	Türk Traktör ve Ziraat Makineleri A.Ş.	**11**	Mmc Sanayi ve Ticari Yatırımlar A.Ş.
12	Tüpraş-Türkiye Petrol Rafineleri A.Ş.	**12**	Otto Holding A.Ş.
13	Ülker Bisküvi Sanayi A.Ş.	**13**	Senkron Siber Güvenlik ve Yazılım ve

			Bilişim Çözümleri A.Ş.
14	Vestel Elektronik Sanayi ve Ticaret A.Ş.	**14**	Zedur Enerji Elektrik Üretim A.Ş.

2013 1.çeyrek ile 2022 4.çeyrek dönemleri arasındaki toplam 40 dönem için hazırlanan finansal oranlar Tablo 4'te sunulmuştur.

Tablo 5.2: Çalışmada kullanılan finansal oranlar

1	Cari Oran	23	Özsermaye Büyümesi
2	Dönen Varlıklar / Aktif	24	ROIC
3	Likit Oran	25	Borç Kaynak Oranı
4	Nakit Oran	26	Net Satışlar / Kısa Vadeli Borçlar
5	Aktif Karlılık	27	Stoklar / Dönen Varlıklar
6	Brüt Esas Faaliyet Kar Marjı	28	Net Borç / Aktifler
7	Esas Faaliyet Kar Marjı	29	FAVÖK / Kısa Vadeli Borçlar
8	FAVÖK Marjı	30	FAVÖK / Mali Borçlar
9	Net Kar Marjı	31	Kısa Vadeli Borç / Aktif
10	Özsermaye Karlılığı	32	Net Satışlar / Kısa Vadeli Borç
11	VAFÖK Marjı	33	Özsermaye / Aktifler
12	Finansman Gideri / Net Satış	34	Toplam Borç / Özsermaye
13	Firma Değeri / Defter Değeri	35	Aktif Devir Hızı
14	Firma Değeri / FAVÖK	36	Alacak Devir Hızı
15	Firma Değeri / Net Satış	37	Alacak Tahsil Süresi
16	FK (Fiyat Kazanç)	38	Dönen Varlıklar Devir Hızı
17	Piyasa Değeri / FAVÖK	39	Duran Varlıklar Devir Hızı Sektörel
18	Piyasa Değeri / Aktifler	40	Nakit Döndürme Süresi
19	Piyasa Değeri / Defter Değeri	41	Özsermaye Devir Hızı
20	PD / Net Satış	42	Stok Devir Hızı
21	Aktif Büyüme	43	Ticari Borç Devir Hızı
22	Net Satışlar Büyüme		

5.2. Yöntem

Çalışmada finansal başarısızlığın tahmini için yapay zekâ yöntemlerinden makine

öğrenmesinin aşağıda yer alan alt başlıklarda açıklamaları bulunan farklı algoritmalar ve eğitim şekilleri ile gerçekleştirerek performans sonuçlarının karşılaştırılması yapılmıştır.

Analiz yapmak için Weka programı kullanılmış olup, veri setini eğitirken Split %66 Modeli ve 10-fold modeli kullanılmıştır;

Split %66 yöntemde veri setinin %66'sı modelin eğitilmesi için kullanılırken, kalan %34'lük kısmı modelin doğruluğunu veya performansını değerlendirmek amacıyla test verisi olarak ayrılır.

Bu ayrımın mantığı şu şekildedir:

- **Eğitim Verisi (%66)**: Model, veri setinin bu kısmını kullanarak örüntüleri öğrenir. Yani, modelin öğrenme süreci bu veriler üzerinden yapılır.
- **Test Verisi (%34)**: Model eğitildikten sonra, modelin gerçek dünyadaki performansını görmek için bu ayrılmış veri seti üzerinde test edilir. Test verileri, modelin eğitimi sırasında hiç görmediği verilerden oluşur ve modelin genel doğruluğunu değerlendirmek için kullanılır.

10 fold yönteminde ise makine öğrenimi modellerinin performansını değerlendirmek için kullanılan popüler bir doğrulama yöntemidir. Bu yöntem, veri setini 10 eşit parçaya (fold) ayırarak, her bir parçayı bir test kümesi olarak kullanır ve kalan 9 parçayı eğitim kümesi olarak değerlendirir. Bu işlem 10 kez tekrarlanır ve her seferinde farklı bir bölüm test kümesi olarak seçilir.

5.2.1. Sınıflandırma Algoritmaları

Başlıca sınıflandırma algoritmaları aşağıdaki gibidir;

5.2.1.1. Random Forest

Random Forest, birden fazla karar ağacının bir araya gelmesiyle oluşturulan bir makine öğrenme algoritmasıdır. Aşağıda, Random Forest modelinin temel matematiksel formülü ve

işleyişi açıklanmıştır.

5.2.2. Karar Ağaçları

Karar ağacı algoritması, veri kümesini farklı koşullara göre bölerek tahminler yapar. Bir karar ağacı aşağıdaki gibi çalışır:

- Her düğümde veri kümesi belirli bir özellik ve eşik değeri kullanılarak bölünür.
- Bu işlem, yaprak düğümlerine ulaşılana kadar tekrarlanır.
- Yaprak düğümleri, sınıflandırma problemi için sınıf etiketlerini veya regresyon problemi için sürekli değerleri temsil eder.

5.2.3. Rastgele Alt Küme Seçimi (Bagging)

Random Forest, her bir karar ağacını eğitmek için eğitim veri kümesinin rastgele alt kümelerini kullanır. Bu teknik, Bagging (Bootstrap Aggregating) olarak bilinir:

- Verilerin bootstrap örnekleri alınır (tekrarlarla rastgele örnekler seçilir).
- Her bir örnek, belirli sayıda veri noktası içerir ve bu örnekler ile ayrı karar ağaçları eğitilir.

5.2.4. Rastgele Özellik Seçimi

Her düğümde veri kümesinin tüm özelliklerinden değil, rastgele seçilen bir alt kümesinden seçim yapılır. Bu özellik, ağaçlar arasındaki bağımlılığı azaltır ve modelin genelleme yeteneğini artırır.

5.2.5. Birleştirme (Aggregation)

Random Forest modeli, her bir karar ağacının tahminlerini birleştirir:

- Sınıflandırma problemi için: Ağaçların oy çokluğuna dayalı sınıflandırmaları birleştirilir.

$$\hat{y} = mode(\hat{y}_1, \hat{y}_2, \dots, \hat{y}_N)$$ (Denklem 19)

Burada $\hat{y}_i, i$-inci karar ağacının tahminini temsil eder ve $\hat{y}$ ise Random Forest modelinin nihai tahminidir.

- Regresyon problemi için: Ağaçların tahminlerinin ortalaması alınır.

$$\hat{y} = \frac{1}{N}\sum_{i-1}^{N} \hat{y}_i$$ (Denklem 20)

Burada $\hat{y}_i, i$-inci karar ağacının tahminini temsil eder ve $\hat{y}$ ise Random Forest modelinin nihai tahminidir.

Genel Formül

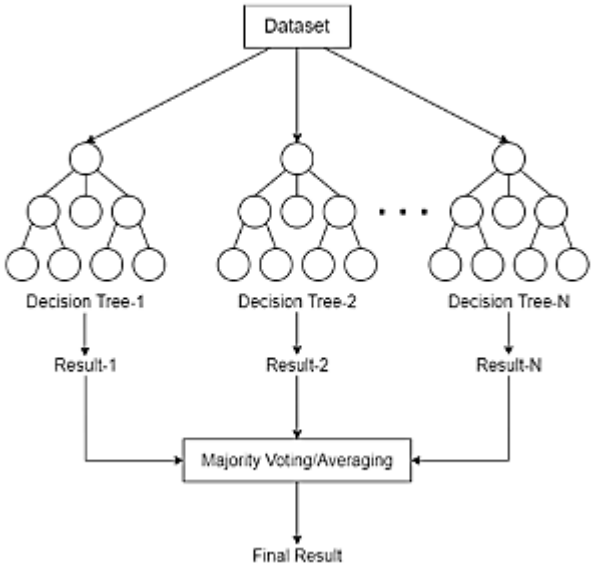
Genel olarak, Random Forest modeli aşağıdaki adımları izler:

1. Eğitim veri kümesinden N adet bootstrap örneği oluştur.
2. Her bir bootstrap örneği ile bağımsız olarak N adet karar ağacı eğit.
3. Her bir ağaçta her düğümde, belirli sayıda rastgele seçilen özelliklerden en iyi bölmeyi seç.
4. Sınıflandırma problemleri için, tüm ağaçların sınıf tahminlerinin modunu alarak son tahmini yap.
5. Regresyon problemleri için, tüm ağaçların tahminlerinin ortalamasını alarak son tahmini yap.

Random Forest modeli hem sınıflandırma hem de regresyon problemlerinde yüksek performans gösterir ve aşırı uyum riskini azaltmak için çeşitli teknikler kullanır.

5.2.6. Algoritma

Random Forest algoritması aşağıdaki gibidir.



Şekil 5.1: Random forest algoritması (Hermawan vd., 2021)

5.2.1.2.Random Tree

Rastgele Ağaç (Random Tree) algoritması, karar ağaçları (decision trees) ile çalışırken bazı rastgelelik unsurları ekleyerek oluşturulur. Bu algoritmanın temel yapısı ve formülleri, karar ağacının genel yapısından türetilir ve rastgelelik eklenerek çeşitlendirilir. Rastgele Ağaç algoritmasının oluşturulma sürecindeki adımlar ve formüller şu şekildedir:

5.2.7. Rastgele Özellik Seçimi (Random Feature Selection)

Her düğümde, belirli sayıda rastgele seçilmiş özellik arasından en iyi bölmeyi sağlayan özellik seçilir. Bu özellikler arasından en iyi olanı, bilgi kazancı (information gain) veya Gini indeksi (Gini index) gibi ölçütlerle belirlenir.

5.2.8. Entropi (Entropy) ve Bilgi Kazancı (Information Gain)

Entropi, veri kümesindeki düzensizliği ölçer ve şu formülle hesaplanır:

$$H(S) = -\sum_{i=1}^{c} p_i - \log_2 p_i \quad \text{(Denklem 21)}$$

Burada:

- $H(S)$: Veri kümesi S için entropi.
- c: Sınıf sayısı.
- p_i: i sınıfının olasılığı.

Bir özellik A kullanılarak veri kümesi S bölündüğünde bilgi kazancı şu şekilde hesaplanır:

$$IG(S, A) = H(S) - \sum_{v \epsilon Values(A)} \frac{|S_v|}{|S|} H(S_v) \quad \text{(Denklem 22)}$$

Burada:

- $IG(S, A)$: Özellik A kullanılarak elde edilen bilgi kazancı.
- $H(S)$: Başlangıç veri kümesi S için entropi.
- $Values(A)$: Özellik A'nın aldığı olası değerler kümesi.
- S_v: A özelliğinin v değerini aldığı alt küme.
- $\frac{|S_v|}{|S|}$: Alt küme S_v'nın S içindeki oranı.

5.2.9. Gini İndeksi

Gini indeksi, bir veri kümesindeki düzensizliği ve yanlış sınıflandırma olasılığını ölçer:

$$Gini(S) = 1 - \sum_{i=1}^{c} p_i^2 \quad \text{(Denklem 23)}$$

Burada:

- $Gini(S)$: Veri kümesi S için Gini indeksi.
- c: Sınıf sayısı.
- p_i: i sınıfının olasılığı.

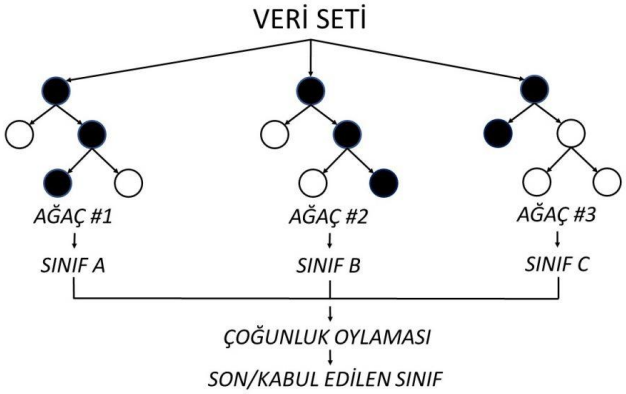
5.2.10. Rastgele Ağaç Yapısı

Rastgele ağaç algoritmasının adımları şu şekildedir:

- **Veri Kümesinin Rastgele Alt Kümesi**: Her düğümde, tüm veri kümesinden rastgele bir alt küme seçilir. Bu alt küme, orijinal veri kümesinin bir parçasıdır.
- **Özelliklerin Rastgele Alt Kümesi**: Her düğümde, tüm özelliklerden rastgele bir alt küme seçilir.
- **Bölme Kararı**: Rastgele seçilen özelliklerin arasından en iyi bölmeyi sağlayan özellik, bilgi kazancı veya Gini indeksi kullanılarak seçilir.
- **Düğüm Bölme**: Seçilen özelliğe göre veri kümesi iki alt kümeye bölünür.
- **Yinelenme (Recursion)**: Bu adımlar, belirli bir durma kriterine (örneğin, belirli bir derinliğe ulaşma, yaprak düğümdeki minimum örnek sayısı, vb.) kadar tekrarlanır.

5.2.11. Algoritma

Random Tree algoritması aşağıdaki gibidir.



Şekil 5.2: Rastgele ağaç algoritması (Öztürk M. , 2022)

Bayes Ağı

Bayes Ağları, değişkenler arasındaki olasılıksal bağımlılık ilişkilerini modelleyen ve bu ilişkileri yönlü oklar aracılığıyla gösteren yönlü, çevrimsiz olasılıksal ağlardır (Babacan ve Karaduman, 2018). Bu ağlar, iki ana bileşenden oluşur:

1. Grafiksel Yapı

Bayes Ağları, düğümler (nodes) ve yönlü oklar (edges) içeren bir grafik yapısından oluşur:

- **Düğüm (Node)**: Her düğüm bir rastgele değişkeni temsil eder. Örneğin, bir hastanın hastalığı olup olmadığını belirleyen bir düğüm.
- **Yönlü Ok (Directed Edge)**: Değişkenler arasındaki olasılıksal bağımlılık ilişkilerini gösterir. Yönlü oklar, bir değişkenin başka bir değişkeni nasıl etkilediğini ifade eder. Örneğin, bir hastalığın semptomları üzerindeki etkisi.

2. Koşullu Olasılık Tabloları (Conditional Probability Tables, CPTs)

Her düğüm için, bu düğümün ebeveynlerine bağlı olasılık dağılımlarını içeren koşullu olasılık tabloları bulunur. Bu tablolar, belirli durumlar altında değişkenlerin alabileceği değerlerin olasılıklarını gösterir.

5.2.12. Bayes Ağları Nasıl Çalışır

Bayes Ağları, Bayes teoremi ile birlikte çalışır. Bayes teoremi, olayların olasılıklarını hesaplamak için kullanılır ve aşağıdaki gibi ifade edilir:

$$P(A|B) = P(B|A) + \frac{P(A)}{P(B)} \quad \text{(Denklem 24)}$$

Burada:

- $P(A \mid B)P(A|B)P(A \mid B)$: B verildiğinde A'nın olasılığıdır (posterior olasılık).

- $P(B \mid A)P(B|A)P(B \mid A)$: A verildiğinde B'nin olasılığıdır (likelihood).
- $P(A)P(A)P(A)$: A'nın öncül olasılığıdır (prior probability).
- $P(B)P(B)P(B)$: B'nin öncül olasılığıdır.

5.2.13. Bayes Ağları Kullanarak Sınıflandırma

Bayes Ağları, belirli bir sınıf etiketini tahmin etmek için kullanılabilir. Bu süreçte şu adımlar izlenir:

Modeli Öğrenme (Learning the Model):

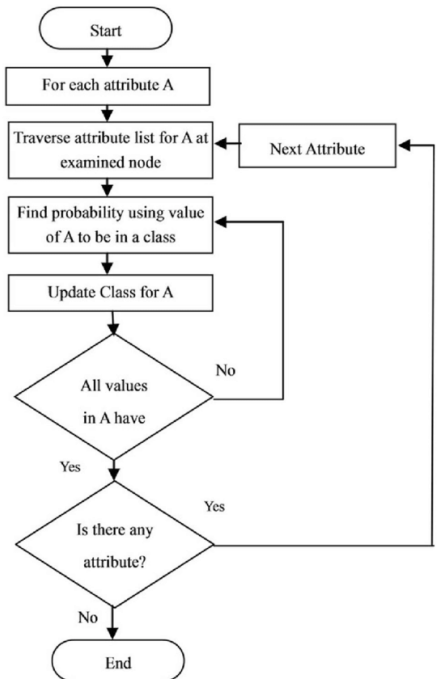
- **Yapı Öğrenme (Structure Learning):** Değişkenler arasındaki bağımlılık ilişkilerinin belirlenmesi.
- **Parametre Öğrenme (Parameter Learning):** Koşullu olasılık tablolarının (CPT) tahmin edilmesi.

Çıkarım (Inference):

- Eğitim verileri kullanılarak ağın yapılandırılması ve parametrelerinin öğrenilmesi.
- Yeni bir örnek geldiğinde, bu örneğin belirli bir sınıfa ait olma olasılığının hesaplanması.

5.2.14. Algoritma

Bayes ağı algoritması aşağıdaki gibidir.



Şekil 5.3: Bayes ağları algoritması (Sneha ve Gangil, 2019)

5.2.1.3.Model Ağaç Oluşturma (M5Prime)

M5P algoritması, M5 olarak bilinen karar ağacı yapısını doğrusal regresyon fonksiyonları saklayabilecek şekilde değiştirmiştir. Bu algoritmanın oluşturulmasında tümevarım ağaç algoritması kullanılmaktadır. İlk olarak, tümevarım ağacı algoritmasıyla bir ağaç oluşturulur. Bu ağaç, veri setindeki öznitelikleri temsil eden düğümleri ve kararları içerir (Aydemir vd., 2020).

Daha sonra, ağaçtaki düğümler bölme kıstasına göre çocuklara bölünür. Bu bölme işlemi, veri setindeki öznitelikleri kullanarak düğümleri alt ağaçlara ayırır. Eğer düğümlerdeki sınıf değerleri çok farklılık göstermez veya düğüm sayısı belirli bir eşiği aşmazsa, bölme işlemi durdurulur. Bu şekilde ağaç oluşturulur ve veri setindeki özniteliklerin ilişkileri belirlenir.

Sonraki aşamada, ağacın her bir yaprağı kontrol edilir ve budama işlemi yapılır. Budama işlemi, ağacın gereksiz dallarını kaldırarak modelin basitleştirilmesini sağlar. Budanmış düğümler, regresyon düzeyine dönüştürülür, yani her bir yaprakta bulunan veri noktalarının ortalaması veya diğer regresyon yöntemleri kullanılarak bir regresyon fonksiyonu belirlenir.

Son olarak, alt ağaçlar arasındaki süreksizlikleri önlemek için bir yumuşatma prosedürü uygulanır. Bu prosedür, modelin aşırı uydurma (overfitting) problemlerini önlemek ve daha genelleştirilebilir hale getirmek için kullanılır.

M5P algoritması, karar ağaçlarının regresyon problemleri için genişletilmiş bir versiyonudur ve doğrusal regresyon fonksiyonları saklayabilme yeteneğiyle bilinir. Bu özelliği sayesinde hem sınıflandırma hem de regresyon problemleri için kullanılabilir ve geniş bir uygulama yelpazesine sahiptir.

M5P sınıflandırma, karar ağaçları ve doğrusal regresyon fonksiyonlarını birleştirerek sınıflandırma problemlerini çözmek için kullanılan bir yöntemdir. Bu yöntem, veri setindeki öznitelikleri kullanarak tümevarım ağaçları oluşturur. Her bir düğüm, belirli bir özniteliği temsil eder ve bu özniteliklere göre veri setini böler. Ağaç oluşturulduktan sonra, her bir yaprak düğümünde veri noktalarının sınıf etiketlerine göre bir regresyon modeli oluşturulur. Test verileri, bu ağaç üzerinde gezinerek sınıflandırılır. Modeldeki gereksiz dallar budanır ve süreksizlikleri önlemek için bir yumuşatma prosedürü uygulanır. Sonuç olarak, M5P sınıflandırması doğrusal regresyon modellerini kullanarak sınıflandırma performansını artırırken modelin yorumlanabilir halini korur.

5.2.15. Ağaç Oluşturma Aşamaları

Ağacın kökünden başlayarak, her bir düğümde verilen veri kaydının belirli bir özniteliğini, o düğümdeki bölünmüş değerle karşılaştıran bir matematiksel mantık kullanılarak yeni bir veri kaydı bir yaprağa ulaşana kadar hareket edilir. Bu süreç, veri kaydının ağaç boyunca nasıl ilerleyeceğini belirler ve sonuç olarak bilginin ağaç yapısından elde edilmesine olanak tanır (Ecemiş, 2018).

Ağaç Oluşturma: Hedef, standart sapma azaltmasını sağlayan adı verilen bir ölçünün en

büyük değerine ulaşmaktır. Bu, karar ağacının bölme kriterlerinin belirlenmesinde kullanılan ana ölçüttür.

Budama: Bu aşamada hedef aşırı öğrenmeyi engellemektir.

Düzgünleştirme: Ortaya çıkabilecek süreksiz durumlar telafi edilir. Bu, modelin daha düzgün ve genelleştirilebilir hale getirilmesini sağlar.

Düzgünleştirme işlemi;

$$p' = \frac{np+kq}{n+k} \qquad \text{(Denklem 25)}$$

Burada;

p': Üst düğüme aktarılan tahmini,
p: Alt düğümden aktarılan tahmini,
q: Düğümde tahmin edilen değeri
n: Alt düğüme erişmek için kullanılan eğitim örneklerinin sayısını,
k: Düzgünleştirme sabitini

ifade eder

Ağaç oluşturma aşağıdaki formülle ifade edilir;

$$SDR = SD(T) - \sum_i \frac{|T_i|}{T} sd(T_i) \qquad \text{(Denklem 26)}$$

Budama sürecinde, her düğüm için test verilerinin beklenen hatası tahmin edilir. Düğümdeki eğitim verilerinde sınıf için tahmin edilen değer ile gerçek değer arasındaki mutlak farkların ortalaması hesaplanır. Bu ortalamaya, görünmeyen durumlar için beklenen hatayı daha küçük değerde tahmin edebilmek için telafi faktörü uygulanır. Telafi faktörü, düğümdeki eğitim örneklerinin sayısı (n) ile düğümün sınıf değeri veren doğrusal modelin parametre sayısı (v) arasındaki farkı ifade eder. Bu şekilde, modelin tahmin hatası en aza indirilmeye çalışılır.

5.2.1.4. Random Comminittee

Random Committee (Rastgele Komite) bir sınıflandırma algoritması olup, zayıf sınıflandırıcıların bir topluluğunu (ensemble) kullanarak nihai tahminleri üretir. Bu yöntemde, her zayıf sınıflandırıcı, farklı bir rastgele sayı tohumuyla eğitilir ve nihai karar, bu sınıflandırıcıların tahminlerinin ortalaması alınarak belirlenir. Sınıflandırma durumunda, tahminler oy kullanma yoluyla değil, olasılık tahminlerinin ortalaması alınarak üretilir (Lira, ve diğerleri, 2007).

Bir test örneği x için, her bir temel sınıflandırıcının tahmini $h_i(x)$ olarak ifade edilir. Nihai tahmin $H(x)$ şu şekilde hesaplanır:

$$H(x) = \frac{1}{\mathrm{N}} \sum_{i=1}^{N} h_i(x) \qquad \text{(Denklem 27)}$$

Bu formül, N adet temel sınıflandırıcının tahminlerinin ortalamasını alarak nihai sonucu verir.

Random Committee algoritması, özellikle çeşitli ve bağımsız tahminlerin birleştirilmesi yoluyla tahmin doğruluğunu artırmayı hedefler. Bu yöntem, tek bir modelin yetersiz kalabileceği durumlarda topluluk yaklaşımları kullanarak daha iyi genelleme performansı sağlayabilir.

5.2.1.5. Locally Weighted Learning

Yerel Ağırlıklı Öğrenme (LWL), karmaşık görevleri eğitmek ve sınıflandırmak için istatistiksel öğrenme teknikleri kullanan bir tembel sınıflandırıcıdır (K, 2015). LWL, büyük veri setleriyle hızlı bir şekilde başa çıkabilir, eğitim süresini kısaltır ve yüksek boyutlu giriş verileriyle başarılı bir şekilde çalışabilir. İki stratejisi vardır: Bellek tabanlı LWL, tüm eğitim verilerini bellekte saklar ve yeni bir girdi için verimli arama ve enterpolasyon teknikleri kullanır. Bellek tabanlı olmayan LWL ise veri depolamaktan kaçınmak için yinelemeli sistem tanımlama tekniklerini kullanır ve eğitim verilerini kompakt temsillerde önbelleğe alır, böylece yeni veriler için arama süreleri daha hızlı hale gelir.

$$\hat{y}(x) = \frac{\sum_{i=1}^{N} w(x,x_i) * y_i}{\sum_{i=1}^{N} (x,x_i)}$$ (Denklem 28)

Burada:

- $\hat{y}(x)$: Yeni veri noktası x için yapılan tahmin.
- N: Eğitim veri setindeki toplam veri noktası sayısı.
- $w(x, x_i)$: Veri noktası x_i ile yeni veri noktası x arasındaki ağırlık. Bu ağırlık genellikle bir mesafe fonksiyonu kullanılarak hesaplanır.
- y_i: Eğitim veri setindeki x_i veri noktasının hedef değeri (etiketi).

5.2.1.6.Lazy Instance-Based k (Lazy IBk)

IBK, aynı mesafe metriğini kullanan bir k-en yakın komşu (k-NN) sınıflandırıcısıdır. En yakın komşuların sayısı, nesne düzenleyicisinde açıkça belirtilebilir veya bırak-bir-çık çapraz doğrulama kullanılarak otomatik olarak belirlenebilir; bu durumda belirli bir üst sınır dikkate alınır (Mohan, 2013). Farklı arama algoritmaları, en yakın komşuları bulma görevini hızlandırmak için kullanılabilir. Varsayılan olarak doğrusal arama kullanılır, ancak KD-ağaçları, top ağaçları ve "kaplama ağaçları" gibi diğer seçenekler de mevcuttur.

Mesafe fonksiyonu, arama yönteminin bir parametresidir ve varsayılan olarak Öklidyen mesafe kullanılır; diğer seçenekler arasında Manhattan ve Minkowski mesafeleri bulunur. Birden fazla komşudan gelen tahminler, test örneğine olan uzaklıklarına göre ağırlıklandırılabilir ve mesafeyi ağırlığa dönüştürmek için iki farklı formül uygulanır.

Öklidyen Uzaklığı

$$\sqrt{\sum_{i=1}^{k} (x_i - y_i)^2}$$ (Denklem 98)

Manhattan uzaklığı

$$\sum_{i=1}^{k} |x_i - y_i|$$ (Denklem 30)

Minkowski Uzaklığı

$$\left(\sum_{i=1}^{k}(|x_i - y_i|)^q\right)^{1/q} \quad \text{(Denklem 31)}$$

Sınıflandırıcı tarafından tutulan eğitim örneklerinin sayısı, pencere boyutu seçeneği ayarlanarak sınırlandırılabilir. Yeni eğitim örnekleri eklendikçe, en eski olanlar çıkarılarak eğitim örneklerinin sayısının bu boyutta kalması sağlanır.

$$F(x_q) = \arg\max \sum_{i=1}^{k} \delta(v, f(x_i)) \quad \text{(Denklem 32)}$$

5.2.1.7.M5Rules

M5Rules, regresyon tabanlı bir sınıflandırma algoritmasıdır ve M5P (M5 Prime) algoritmasının bir türevidir. M5Rules, M5P ağacının yapraklarını bir dizi kurala dönüştürerek sınıflandırma yapar. Bu kural tabanlı yaklaşım, sınıflandırma kararlarını açıklanabilir bir formda sunar.

M5 ağaçları, parçalı doğrusal fonksiyonlar gibi çalışır. Bunun nedeni, regresyon ağaçlarının yaprak düğümlerindeki değerlerin benzer şekilde olduğu, M5 algoritmasında ise ağaçların çok değişkenli doğrusal modeller kullanılarak oluşturulmasıdır. M5'in kullanılmasının olumlu bir yanı, öğrenme ve 100'lere kadar özniteliği olan görevlerle başa çıkma konusundaki etkinliğidir. Bu model ağaçları genellikle regresyon ağaçlarından daha küçüktür ve tahmin gibi görevlerde daha doğru olduğu araştırılmıştır. Model ağaçları, sabit bir öznitelik setine sahip eğitim örneklerinden oluşan bir koleksiyon üzerine inşa edilir. Bir ağaç oluşturmak için, eğitim örneklerinin hedef değerleri bazı diğer öznitelik değerleri ile ilişkilendirilebilir. Tahmin doğruluğu, görülmeyen vakaların hedef değerlerini tahmin etme yeteneği ile değerlendirilebilir (Kolli Sai Poorna Chand vd., 2019).

Bu ağaçların oluşturulmasında böl ve yönet metodolojisi kullanılır. Küme T, test sonuçlarına eşleşen alt kümeler halinde tekrar tekrar bölünür. Bazen bölme nedeniyle aşırı derecede karmaşık yapılar elde edilir, bu durumda budama yapılabilir. Bir sonraki adım, T'deki vakaların hedef değerlerinin standart sapmasının hesaplanmasıdır. T, bir testin sonuçlarına

göre bölünür. Testin sonucu, hata azalmasının beklenen miktarını verir ve m5, hata azalmasının en yüksek olduğu yerleri seçer. Yalnızca testler tarafından referans alınan öznitelikler için standart regresyon teknikleri kullanılarak, her ağaç düğümü için çok değişkenli doğrusal model oluşturulur. En alt düğümden başlayarak, model ağacının her bir yaprağı, bu düğüm için son modeli seçmek için incelenir. Son model, basitleştirilmiş bir doğrusal model veya en az tahmin edilen hataya sahip olan model alt ağaç olabilir. Eğer doğrusal model en az hataya sahipse, budama yapılır.

M5Rules algoritması, doğrudan matematiksel bir formülle ifade edilmez. Bunun yerine, M5P ağacının yapraklarının kural tabanlı bir formata dönüştürülmesi ve bu kuralların kullanılmasıyla sınıflandırma yapılır. Her bir kural, bir sınıf etiketini tahmin etmek için bir dizi koşul ifadesinden oluşur.

Diyelim ki, bir M5Rules modeli aşağıdaki gibi bir kural seti oluşturmuş olsun:

1. "Eğer A > 5 ve B < 10 ise Sınıf = X"
2. "Eğer A <= 5 ve C = 'evet' ise Sınıf = Y"
3. "Eğer B > 15 ise Sınıf = Z"

Bu kural seti, test örneklerini sınıflandırmak için kullanılır. Örneğin, test örneği xxx için A= 7, B = 8 ve C = 'evet' olsun. Bu örneğin sınıfını belirlemek için kurallar sırayla kontrol edilir. İlk kurala uyan xxx sınıflandırılır ve Sınıf = X olarak tahmin edilir.

M5Rules algoritması, açıklanabilirlik ve yorumlanabilirlik açısından avantaj sağlar çünkü sınıflandırma kararlarını kurallar şeklinde sunar. Bu, özellikle karar süreçlerinin anlaşılması ve doğrulanması gereken durumlarda faydalıdır.

5.3. BULGULAR

Tablo 3'de yer alan veriler, çalışmada kullanılan makine öğrenimi algoritmalarının performans metriklerini göstermektedir. Bu metrikler arasında Korelasyon Katsayısı, Ortalama Mutlak Hata, Kök Ortalama Kare Hatası ve Göreceli Mutlak Hata bulunmaktadır. Bu metrikler, modellerin doğruluğunu ve hata oranlarını değerlendirmek için kullanılır. Her model, 66% oranında veri bölünmesi (split %66) ve 10-kat çapraz doğrulama (10-fold cross-validation) ile test edilmiştir.

5.3.1. Metriklerin Anlamı

- **Correlation coefficient (Korelasyon katsayısı)**: Modelin tahminleri ile gerçek değerler arasındaki doğrusal ilişkinin gücünü gösterir. 1'e yakın değerler, anlamlı bir ilişki olduğunu gösterir.
- **Mean absolute error (Ortalama mutlak hata)**: Tahmin edilen ve gerçek değerler arasındaki ortalama mutlak farktır. Daha düşük değerler, modelin daha anlamlı olduğunu gösterir.
- **Root mean squared error (Ortalama karekök hatası)**: Hataların karesinin ortalamasının kareköküdür. Daha düşük değerler, modelin daha iyi olduğunu gösterir.
- **Relative absolute error (Göreceli mutlak hata)**: Modelin mutlak hatasının, tahmin edilmeyen bir modelin hatasına göre ne kadar iyi olduğunu gösterir. Daha düşük değerler, modelin daha iyi performans gösterdiğini ifade eder.

5.4. Analiz Sonuçları

Çalışmanın veri seti ile yapılan analiz sonuçları aşağıdaki gibidir;

5.4.1. M5Prime

M5Prime yöntemi %66 eğitim modeline göre başarı oranı %99.98, 10-fold eğitim modeline

göre başarı oranı %99.99 olmuştur. Buna göre M5Prime yönteminde 10-fold modeli daha başarılı bir sınıflandırma gerçekleştirmiştir.

Tablo 5.3: M5Prime algoritması test sonuçları

Test Mode	*split 66%*	*10-fold*
Correlation Coefficient	*0.9998*	*0.9999*
Mean Absolute Error	*0.0084*	*0.006*
Root Mean Squared Error	*0.01*	*0.0072*
Relative Absolute Error	*1,6782%*	*1,2049%*
Root Relative Squared Error	*1,9915%*	*1,4458%*
Total Number of Instances	*380*	*1119*

5.4.2. Random Forest

Random Forest yöntemi %66 eğitim modeline göre başarı oranı %99.56, 10-fold eğitim modeline göre başarı oranı %99.56 olmuştur. Buna göre Random Forest yönteminde her iki model de aynı sınıflandırmayı yapmıştır.

Tablo 5.4: Random forest algoritması test sonuçları

Test mode	*split 66%*	*10-fold*
Correlation Coefficient	*0.9956*	*0.9956*
Mean Absolute Error	*0.0276*	*0.0218*
Root Mean Squared Error	*0.0525*	*0.0503*
Relative Absolute Error	*5,5093%*	*4,3452%*
Root Relative Squared Error	*10,4719%*	*10,0568%*
Total Number of Instances	*380*	*1119*

5.4.3. Random Tree

Random Tree yöntemi %66 eğitim modeline göre başarı oranı %98.94, 10-fold eğitim modeline göre başarı oranı %97.93 olmuştur. Buna göre Random Tree yönteminde %66

eğitim modeli daha başarılı bir sınıflandırma gerçekleştirmiştir.

Tablo 5.5: Random tree algoritması test sonuçları

Test mode	*split 66%*	*10-fold*
Correlation Coefficient	*0.9894*	*0.9793*
Mean Absolute Error	*0.0054*	*0.0106*
Root Mean Squared Error	*0.0726*	*0.1021*
Relative Absolute Error	*1,0764%*	*2,1166%*
Root Relative Squared Error	*14,4808%*	*20,3926%*
Total Number of Instances	*380*	*1119*

5.4.4. M5Rules

M5Rules yöntemi %66 eğitim modeline göre başarı oranı %99.99, 10-fold eğitim moeline göre başarı oranı %100 olmuştur. Buna göre M5Rules yönteminde 10-fold modeli daha başarılı bir sınıflandırma gerçekleştirmiştir.

Tablo 5.6: M5Rules algoritması test sonuçları

Test mode	*split 66%*	*10-fold*
Correlation Coefficient	*0.9999*	*1*
Mean Absolute Error	*0.0037*	*0,0028*
Root Mean Squared Error	*0.0064*	*0,0049*
Relative Absolute Error	*0,7382%*	*0,5607%*
Root Relative Squared Error	*1,2771%*	*0,9844%*
Total Number of Instances	*380*	*1119*

5.4.5. Random Committee

Random Committee yöntemi %66 eğitim modeline göre başarı oranı %99.24, 10-fold eğitim modeline göre başarı oranı %99.35 olmuştur. Buna göre M5Prime yönteminde 10-fold modeli daha başarılı bir sınıflandırma gerçekleştirmiştir.

Tablo 5.7: Random committee algoritması test sonuçları

Test mode	*split66%*	*10-fold*
Correlation Coefficient	*0,9924*	*0,9935*
Mean Absolute Error	*0,0257*	*0,0153*
Root Mean Squared Error	*0,0645*	*0,058*
Relative Absolute Error	*5,1244%*	*3,0643%*
Root Relative Squared Error	*12,8674%*	*11,5789%*
Total Number of Instances	*380*	*1119*

5.4.6. Lazy Instance-Based k (Lazy IBk)

Lazy Instance-Based k yöntemi %66 eğitim modeline göre başarı oranı %92.89, 10-fold eğitim modeline göre başarı oranı %94.77 olmuştur. Buna göre Lazy Instance-Based k yönteminde 10-fold modeli daha başarılı bir sınıflandırma gerçekleştirmiştir.

Tablo 5.8: Lazy instance-based k algoritması test sonuçları

Test mode	*split 66%*	*10-fold*
Correlation Coefficient	*0,9289*	*0,9477*
Mean Absolute Error	*0,0368*	*0,0268*
Root Mean Squared Error	*0,1919*	*0,1637*
Relative Absolute Error	*7,3552%*	*5,3558%*
Root Relative Squared Error	*38,3011%*	*35,7075%*
Total Number of Instances	*380*	*1119*

5.4.7. Locally Weighted Learning

Locally Weighted Learning yöntemi %66 eğitim modeline göre başarı oranı %100, 10-fold eğitim modeline göre başarı oranı %100 olmuştur. Buna göre Locally Weighted Learning yönteminde her iki model de aynı sınıflandırmayı yapmıştır.

Tablo 5.9: Locally weighted learning algoritması test sonuçları

Test mode	*split 66%*	*10-fold*
Correlation Coefficient	*1*	*1*
Mean Absolute Error	*0*	*0*
Root Mean Squared Error	*0*	*0*
Relative Absolute Error	*0%*	*0%*
Root Relative Squared Error	*0%*	*0%*
Total Number of Instances	*380*	*1119*

Tablo 5.10: Sonuç tablosu

	M5P		RandomForest		RandomTree		M5Rules		RandomCommittee		IBk		LWL	
Test Modu	split 66%	10-fold	split 66%	10-fold	split 66%	10-fold	split 66%	10-fold	split66%	10-fold	split 66%	10-fold	split 66%	10-fold
Korelasyon Katsayısı	0.9998	0.9999	0.9956	0.9956	0.9894	0.9793	0.9999	1	0,9924	0,9935	0,9289	0,9477	1	1
Ortalama Mutlak Hata	0.0084	0.006	0.0276	0.0218	0.0054	0.0106	0.0037	0,0028	0,0257	0,0153	0,0368	0,0268	0	0
Ortalama Karekök Hatası	0.01	0.0072	0.0525	0.0503	0.0726	0.1021	0.0064	0,0049	0,0645	0,058	0,1919	0,1637	0	0
Göreceli Mutlak Hata	1,6782%	1,2049%	5,5093%	4,3452%	1,0764%	2,1166%	0,7382%	0,5607%	5,1244%	3,0643%	7,3552%	5,3558%	0%	0%
Göreceli Karekök Hatası	1,9915%	1,4458%	10,4719%	10,0568%	14,4808%	20,3926%	1,2771%	0,9844%	12,8674%	11,5789%	38,3011%	35,7075%	0%	0%
Toplam Örnek Sayısı	380	1119	380	1119	380	1119	380	1119	380	1119	380	1119	380	1119

5.5. Modellerin Performansları

En iyi performans gösteren modeller;

M5Rules ve LWL modelleri, tüm metriklerde en iyi performansı göstermektedir. Korelasyon katsayısı 1, yani tahmin edilen ve gerçek değerler arasında mükemmel bir uyum vardır. Ortalama mutlak hata ve kök ortalama kare hatası da sıfır veya çok düşük seviyelerde, bu da modellerin tahminlerinin son derece doğru olduğunu gösterir.

Yüksek performanslı modeller;

M5Prime, oldukça yüksek bir korelasyon katsayısı ve düşük hata metrikleri ile iyi performans sergilemektedir. Bu model de finansal başarısızlık tahmininde güvenilir bir seçenektir.

İyi performans gösteren modeller

RandomForest ve RandomCommittee modelleri, nispeten iyi performans göstermektedir. Korelasyon katsayıları 0.99 civarında ve hata metrikleri makul seviyelerdedir. Bu modeller, M5Rules ve M5Prime kadar olmasa da yine de güçlü tahmin yeteneklerine sahiptir.

Daha düşük performans gösteren modeller

RandomTree ve IBk modelleri, diğer modellere kıyasla daha düşük performans göstermektedir. Özellikle IBk modeli, en yüksek hata metriklerine ve en düşük korelasyon katsayısına sahiptir.

6. SONUÇ

Dünya genelinde finans ve bilişim alanları son yıllarda hızla gelişim göstermiştir. Finans sektöründe dijitalleşme, bankacılık hizmetlerinin dijital platformlara taşınmasıyla hız kazanmış, internet bankacılığı, mobil bankacılık ve dijital ödeme sistemleri yaygınlaşmıştır. Fintech (finansal teknoloji) şirketleri, bankacılık, sigortacılık ve yatırım hizmetlerinde yenilikçi çözümler sunarak sektörü dönüştürmekte, blok zincir (blockchain) ve kripto paralar gibi yeni teknolojiler de bu dönüşümde önemli rol oynamaktadır. Dünya genelinde finansal regülasyonlar, sektörün güvenliğini ve şeffaflığını artırmak amacıyla sürekli güncellenmektedir. Bilişim sektöründe ise bulut bilişim, veri depolama ve işleme maliyetlerini düşürerek şirketlerin daha esnek ve verimli çalışmalarını sağlamaktadır. Yapay zekâ ve makine öğrenimi teknolojileri, veri analitiği, otomasyon ve kişiselleştirilmiş hizmetlerde büyük ilerlemeler sağlamaktadır.

Finansal teknoloji (fintech) şirketleri, finans sektöründe yenilikçi çözümler sunarak önemli bir rol oynamaktadır. Fintech, dijital cüzdanlar, çevrimiçi ödeme sistemleri, kişisel finans yönetimi uygulamaları ve kredi değerlendirme sistemleri gibi alanlarda büyük ilerlemeler kaydetmiştir. Özellikle blok zincir (blockchain) teknolojisi, finansal işlemlerin güvenliğini ve şeffaflığını artırarak, bankalar ve diğer finansal kuruluşlar arasında veri paylaşımını ve işlem takibini kolaylaştırmıştır. Kripto paralar ise, geleneksel para birimlerine alternatif olarak ortaya çıkmış ve finansal piyasalarda yeni yatırım ve ödeme yöntemleri sunmuştur.

Yapay zeka (AI) ve makine öğrenimi (ML) teknolojileri, finansal veri analitiği, risk yönetimi, dolandırıcılık tespiti ve müşteri hizmetleri gibi alanlarda kullanılmaktadır. AI tabanlı sistemler, büyük veri setlerini analiz ederek daha doğru ve hızlı kararlar alınmasını sağlamaktadır. Örneğin, makine öğrenimi algoritmaları, kredi başvurularının değerlendirilmesinde ve yatırım stratejilerinin belirlenmesinde kullanılmaktadır.

Genel olarak, bilişim teknolojilerinin finans sektöründe uygulanması, sektördeki verimliliği ve güvenliği artırmış, müşteri deneyimini iyileştirmiş ve finansal hizmetlere erişimi kolaylaştırmıştır. Bu gelişmeler, finans sektörünün dijital dönüşümünde kritik bir rol oynamaktadır.

Çalışmada, işletmelerin finansal başarı durumlarını yapay zekâ teknikleriyle tahmin etmeye

yönelik kapsamlı bir analiz gerçekleştirilmiştir. Bu amaç doğrultusunda, Borsa İstanbul'da ulusal pazardan ve yakın izleme pazarından toplamda 28 şirket seçilmiştir; yakın izleme pazarında ilgili dönem için 14 şirket bulunduğundan, her iki pazardan 14'er şirket alınarak dengeli bir örneklem oluşturulmuştur. Seçilen bu şirketlerin mali tablolarından elde edilen verilerle hesaplanan ve Tablo 4'te detaylı bir şekilde sunulan 43 adet finansal oran yapılan analizlerde kullanılmıştır. Bu oranlar, işletmelerin finansal sağlık durumlarını yansıtmakta ve finansal başarısızlık riskinin tahmin edilmesinde kritik rol oynamaktadır.

Çalışmada gerçekleştirilen analizler, makine öğrenimi algoritmalarını ve veri madenciliği tekniklerini kullanmak için yaygın olarak tercih edilen Weka yazılımı ile gerçekleştirilmiştir. Uygulanan regresyon modelleri arasında M5Prime, Random Forest, Random Tree, M5Rules, Random Committee, Lazy Instance-Based K (IBk) ve Locally Weighted Learning (LWL) algoritmaları bulunmaktadır. Burada kullanılan çeşitli algoritmalar ile, finansal başarı durumu tahminlerinde sonuçların doğruluğunu ve güvenilirliğini artırmayı amaçlamıştır.

Elde edilen bulgulara göre, M5Rules ve LWL algoritmaları en başarılı tahmin sonuçlarını vermiştir. Bu iki algoritma, finansal başarısızlık riskinin tahmin edilmesinde yüksek doğruluk oranları sergilediği için işletmelerin gelecekteki mali durumlarını öngörmede etkin bir şekilde kullanılmıştır. M5Rules algoritması, kurallar tabanlı yaklaşımı sayesinde finansal verilerdeki desenleri ve ilişkileri etkili bir şekilde ortaya koymuş, LWL ise yerel ağırlıklandırma yöntemiyle tahminlerin doğruluğunu artırmıştır.

M5Prime algoritması da benzer şekilde oldukça başarılı bulunmuş ve çok yakın tahminlerde bulunmuştur. Bu algoritma, tahmin edilen ve gerçek değerler arasındaki farkı minimize ederek yüksek performans göstermiştir. Random Forest ve Random Committee modelleri ise 0.99 gibi yüksek korelasyon katsayılarıyla nispeten iyi performans sergilemiştir. Bu modeller, çoklu karar ağaçları kullanarak verilerin çeşitliliğini ve karmaşıklığını dikkate almış, böylece güçlü tahmin yetenekleri sunmuştur. Random Forest, özellikle aşırı uyum sorununu minimize eden yapısıyla dikkat çekmiştir.

Ancak, Random Tree ve IBk modelleri diğer modellere kıyasla daha düşük performans göstermiştir. Random Tree modeli, tek bir karar ağacı kullanarak sınırlı bir tahmin kapasitesi

sunmuş, IBk modeli ise en yakın komşu algoritmasını kullanarak yüksek hata metriklerine ve en düşük korelasyon katsayısına sahip olmuştur. Bu durum, IBk modelinin finansal başarısızlık tahmininde yetersiz kaldığını göstermektedir.

Yapılan analiz sonuçlarından yola çıkarak, M5Rules, LWL ve M5Prime algoritmaları, finansal başarısızlık tahmini konusunda en iyi seçenekler olarak öne çıkmaktadır. Bu algoritmalar, yüksek doğruluk oranları ve genelleme kabiliyetleri ile dikkat çekmektedir. Random Forest ve Random Committee algoritmaları da bu üç modele alternatif olarak kullanılabilecek sağlam seçeneklerdir. Dolayısıyla, işletmelerin finansal başarısızlık riskini tahmin etmek için bu modellerin kullanımı, daha yüksek doğruluk oranları ve genelleme kabiliyetleri sunarak, gelecekte yapılacak olan çalışmalar için yol gösterici olacaktır.

Çalışmada belirli bir zaman aralığındaki verilerin kullanılması, ulusal pazardaki ilgili firmaların rastgele belirlenmesi, literatürde yer alan tüm modellerin sınanmaması gibi birtakım kısıtlar bulunmaktadır. Gelecekte yapılacak araştırmalarda bu kısıtların azaltılması ile elde edilecek sonuçlar daha başarılı olabilir.

KAYNAKLAR

Akbulut, S. ve Adem, K. (2023). Derin Öğrenme ve Makine Öğrenmesi Yöntemleri Kullanılarak Gelişmekte Olan Ülkelerin Finansal Enstrümanlarının Etkileşimi ile Bist 100 Tahmini. *Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi, 12*(1), 52-63. doi:https://doi.org/10.28948/ngumuh.1131191

Akgüç, Ö. (1985). *Finansal Yönetim.* İstanbul: Seçkin Yayıncılık.

Akgün, A. (2013). Firmalarda Finansal Başarısızlığın Tahmini ve İstanbul Menkul Kıymetler Borsası'nda Bir Uygulama. *Doktora Tezi, Selçuk Üniversitesi Sosyal Bilimler Enstitüsü*.

Akkaya, G. C., Demireli, E. ve Yakut, H. (2009). İşletmelerde Finansal Başarısızlık Tahminlemesi: Yapay Sinir Ağları Modeli İle İmkb Üzerine Bir Uygulama. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi, 10*(2), 187-216.

Akkoç, S. (2007). Finansal Başarısızlığın Öngörülmesinde Sinirsel Bulanık Ağ Modelinin Kullanimi Ve Ampirik Bir Çalışma. *Doktora Tezi, Dumlupınar Üniversitesi Sosyal Bilimler Enstitüsü*.

Akpınar, O. ve Akpınar, G. (2017). Finansal Başarısızlık Riskinin Belirleyicileri: Borsa İstanbul'da Bir Uygulama. *İşletme Araştırmaları Dergisi, 9*(4), 932-951. doi:10.20491/isarder.2017.366

Aktaş, R. (1991). Endüstri İstletmeleri İçin Mali Başarısızlık Tahmini - Çok Boyutlu Model Uygulaması. *Doktora Tezi, Ankara Üniversitesi Sosyal Bilimler Entitüsü*.

Aktaş, R., Doğanay, M. ve Yıldız, B. (2003). Mali Başarısızlığın Öngörülmesi: İstatistiksel Yöntemler ve Yapay Sinir Ağı Karşılaştırması. *Ankara Üniversitesi SBF Dergisi, 58*(4), 1-24.

Aktümsek, E. ve Göker, İ. E. (2018). Mali Başarısızlık Tahminlemesinde Sektör Bazlı Bir Karşılaştırma. *İşletme Araştırmaları Dergisi, 10*(4), 401-421. doi:10.20491/isarder.2018.529

Altınöz, U. (2013). BankalarIn Finansal Başarısızlıklarının Yapay Sinir Ağları Modeli Çerçevesinde Tahmin Edilebilirliği. *Dokuz Eylül Üniversitesi İktisadi İdari Bilimler Fakültesi Dergisi, 28*(2), 189-217.

Altman, E. I. (1968). Financial Ratios, Diskriminant Analysis and the Prediction. *The Journal of Finance, 23*(4), 589-609.

Anandarajan, M., Lee, P. ve Anandarajan, A. (2001). Bankruptcy Prediction of Financially Stressed Firms: an Examination of the Predictive Accuracy of Artificial Neural

Networks. *Intelligent Systems in Accounting, Finance and Management, 10*(2), 69-81. doi:doi.org/10.1002/isaf.199

Arda, E. ve Küçükkocaoğlu, G. (2021). Yapay Zeka Yöntemleri ile Hisse Senedi Fiyat Öngörüleri. *Ekonomi, Politika & Finans Araştırmaları Dergisi, 6*(2), 565-586. doi:10.30784/epfad.878664

Aşçı, E., Kılıç, M., Çelik, Ö., Bayrakdar, İ. Ş., Bilgir, E., Aslan, A. F., . . . Orhan, K. (2022). Derin Öğrenme Yöntemi Kullanılarak Geliştirilen Yapay Zekâ Yöntemi ile Panoramik Radyografilerde Dental Restorasyonların Otomatik Tespiti ve Sınıflandırılması: Metodolojik Çalışmalar. *Turkiye Klinikleri Journal of Dental Sciences, 28*(2), 329-337. doi:10.5336/dentalsci.2021-84835

Ataseven, B. (2013). Yapay Sinir Ağları ile Öngörü Modellemesi. *Öneri Dergisi, 10*(39), 101-115.

Ateş, F. ve Şenol, R. (2021). Hava Araçlarında Buzlanma Risk Derecesinin Yapay Zekâ ile Tahmin Edilmesi. *International Journal of 3D Printing Technologies and Digital Industry, 5*(3), 457-468. doi:10.46519/ij3dptdi.957478

Aydemir, E., Kaysi, F. ve Yavuz, M. (2020). İlaç Satış Verileri Kullanılarak Ağaç Algoritmaları İle Elde Edilen Gelirin Tahmin Edilmesi. *Bilgisayar Bilimleri, 5*(1), 14-21.

Aydın, A. D. ve Çavdar, Ş. Ç. (2015). Prediction of Financial Crisis with Artificial Neural Network: An Empirical Analysis on Turkey. *International Journal of Financial Research, 6*(4), 36-45. doi:10.5430/ijfr.v6n4p36

Aydoğdu, A. S., Hatipoğlu, P. U., Özparlak, L. ve Yüksel, S. E. (2015). LWIR and MWIR Images Dimension Reduction and Anomaly Detection With Locally Linear Embedding. *2015 23nd Signal Processing and Communications Applications Conference (SIU)* (s. 819-822). Malatya, Türkiye: Institute of Electrical and Electronics Engineers. doi:10.1109/SIU.2015.7129954

Ayesha, S., Hanif, M. K. ve Talib, R. (2020). Overview and Comparative Study of Dimensionality Reduction Techniques for High Dimensional Data. *Information Fusion, 59*, 44-58. doi:doi.org/10.1016/j.inffus.2020.01.005

Babacan, E. K. ve Karaduman, M. Ö. (2018). Bayes Ağları-K2 Algoritması Üzerine Bir Çalışma. *Karadeniz Fen Bilimleri Dergisi, 8*(2), 24-38. doi:https://doi.org/10.31466/kfbd.418862

Bağcı, H. ve Sağlam, Ş. (2020). Sağlık ve Spor Kuruluşlarında Finansal Başarısızlık Tahmini: Altman, Sprıngate Ve Fulmer Modeli Uygulaması. *Hacettepe Sağlık İdaresi Dergisi, 23*(1), 149-164.

Bakhshıyev, İ. (2009). Bankalarda Mali Başarısızlık Tahmini ve Örnek Bir Uygulama. *Yüksek Lisans Tezi, Dokuz Eylül Üniverstesi Sosyal Bilimler Enstitüsü*.

Baş, M. (2010). İşletmelerde Finansal Başarısızlığın Öngörülmesinde Gri İlişkisel Analiz Tekniği: Tekstil ve Deri Sektöründe Bir Uygulama. *Doktora Tezi, Dumlupınar Üniversitesi Sosyal Bilimler Enstitüsü*.

Baykal, A. (2006). Veri Madenciliği Uygulama Alanları. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*(7), 95-107.

Bilgin, G., Ertürk, S. ve Yıldırım, T. (2008). Nonlinear Dimension Reduction Methods and Segmentation of Hyperspectral Images. *2008 IEEE 16th Signal Processing, Communication and Applications Conference* (s. 1-4). Aydın, Türkiye: Institute of Electrical and Electronics Engineers. doi:10.1109/SIU.2008.4632577

Bilir, H. (2015). Finansal Sıkıntının Tanımı ve Piyasa Odaklı Çözümleri: Borç Yapılandırma, Varlık Satışı ve Yeni Sermaye Enjeksiyonu. *Sosyoekonomi, 23*(9), 9-24. doi:https://doi.org/10.17233/se.72179

Blum, M. (1974). Failing Company Discriminant Analysis. *Journal of Accounting Research, 12*(1), 1-25. doi:https://doi.org/10.2307/2490525

Büyükarıkan, U. ve Büyükarıkan, B. (2014). Bilişim Sektöründe Faaliyet Gösteren Firmaların Finansal Başarısızlık Tahmin Modelleriyle İncelenmesi. *Akademik Bakış Uluslararası Hakemli Sosyal Bilimler Dergisi*(46), 160-172.

Ceyhan, İ. F. (2023). Finans Alanında Makine ve Derin Öğrenmenin Kullanılması: Lisansüstü Tezlerde Sistematik Literatür Taraması. *İnsan ve Toplum Bilimleri Araştırmaları Dergisi, 12*(3), 2187-2209. doi:https://doi.org/10.15869/itobiad.1329889

Civan, M. ve Dayı, F. (2014). Altman Z Skoru ve Yapay Sinir Ağı Modeli İle Sağlık İşletmelerinde Finansal Başarısızlık Tahmini. *Akademik Bakış Uluslararası Hakemli Sosyal Bilimler Dergisi*(41).

Collell, G., Prelec, D. ve Patil, K. R. (2018). A Simple Plug-in Bagging Ensemble Based on Threshold-Moving for Classifying Binary and Multiclass Imbalanced Data. *Neurocomputing, 275*, 330-340. doi:10.1016/j.neucom.2017.08.035

Coşkun, E. ve Sayılgan, G. (2008). Finansal Sıkıntının Dolaylı Maliyetleri: İMKB'de İşlem Gören Şirketlerde Bir Uygulama. *Gazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 10*(3), 45-66.

Çelik, M. K. (2010). Bankaların Finansal Başarısızlıklarının Geleneksel ve Yeni Yöntemlerle Öngörüsü. *Yönetim ve Ekonomi Dergisi, 17*(2), 129-143.

Çelik, M. ve Dursun, G. D. (2021). Z Skor Modeli ile Finansal Risk Tahmini ve BIST100 Tekstil, Giyim Eşyası ve Deri İmalat Sanayi Sektöründe Faaliyet Gösteren Firmalar Üzerine Bir Uygulama. *İktisadi İdari ve Siyasal Araştırmalar Dergisi, 6*(14), 19-31. doi:https://doi.org/10.25204/iktisad.786034

Deniz, Ö. (2005). Poisson Regresyon Analizi. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, 4*(7), 59-72.

Dizgil, E. (2018). Bist Ticaret Endeksinde Yer Alan Şirketlerin Springate Finansal Başarısızlık Modeli İle İncelenmesi. *Bilecik Şeyh Edebali Üniversitesi Sosyal Bilimler Dergisi, 3*(2), 248-267. doi:https://doi.org/10.33905/bseusbed.462059

Doğan, F. ve Türkoğlu, İ. (2019). Derin Öğrenme Modelleri ve Uygulama Alanlarına İlişkin Bir Derleme. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi, 10*(2), 409-445. doi:https://doi.org/10.24012/dumf.411130

Ecemiş, O. (2018). Model Ağaç Yöntemiyle Satış Tahmini: Paslanmaz Çelik Sektöründe Bir Uygulama. *Journal of Academic Social Science*(84), 336-350. doi:10.16992/ASOS.14538

Elarm, R. (1972). The Effect of Lease Data on the Predictive Ability of Financial Ratios. *The Accounting Review, 50*(1), 25-43.

Ertan, A. S. ve Ersan, Ö. (2019). Finansal Başarısızlığı Belirleyen Etkenler: Türkiye İmalat Sektörü Örneği. *Marmara Üniversitesi İktisadi ve İdari Bilimler Dergisi, 40*(2), 181-207.

Espezua, S., Vıllanueva, E., Macıel, C. D. ve Carvalho, A. (2015). A Projection Pursuit Framework gor Supervised Dimension Reduction of High Dimensional Small Sample Datasets. *Neurocomputing, 149, Part B*, 767-776. doi:doi.org/10.1016/j.neucom.2014.07.057

Gacar, A. (2019). Yapay Zekâ ve Yapay Zekânın Muhasebe Mesleğine Olan Etkileri: Türkiye'ye Yönelik Fırsat ve Tehditler. *BALKAN SOSYAL BILIMLER DERGISI, 8*(EUREFE'19), 389-394.

Gökalp, Ö. M. (2022). Makine Öğrenmesi. *Gazi Üniversitesi, Gazi Bilişim Enstitüsü, Adli Bilişim Bölümü*, 1-14.

Gül, M. ve YILMAZ, T. (2023). Altman Z-Skor Modeline Göre Finansal Başarı ve Belirleyicileri. *İşletme Ekonomi ve Yönetim Araştırmaları Dergisi, 6*(2), 202-218. doi:https://doi.org/10.33416/baybem.1297200

Gümüştaş, E. ve Pehlivanlı, A. Ç. (2023). Dengesiz Sınıf Dağılımında Kayıp Gözlem Sorunu için Topluluk Öğrenmesi Sonuçlarının İstatistiksel Değerlendirmesi. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 27*(2), 181-190. doi:10.19113/sdufenbed.1090596

Gündüz, G. ve Cedimoğlu, İ. H. (2019). Derin Öğrenme Algoritmalarını Kullanarak Görüntüden Cinsiyet Tahmini. *Sakarya University Journal of Computer and Information Sciences, 2*(1), 9-17. doi:10.35377/saucis.02.01.517930

Hermawan, D. R., Fatihah, M. F., Kurniawatı, L. ve Helen, A. (2021). Comparative Study of J48 Decision Tree Classification Algorithm, Random Tree, and Random Forest on In-Vehicle CouponRecommendation Data. *2021 International Conference on Artificial Intelligence and Big Data Analytics*, (s. 1-6). Bandung, Endolezya. doi:10.1109/ICAIBDA53487.2021.9689701

Hicks, S. A., Strümke, I., Thambawita, V., Hammou, M., Riegler, M. A., Halvorsen, P. ve Parasa, S. (2022). On Evaluation Metrics for Medical Applications of Artificial Intelligence. *Scientific Reports, 12*(1), 5979. doi:10.1038/s41598-022-09954-8

İloğlu, H. S. (2020). Altman Z Skor Yöntemi ile Havayolu Şirketlerinde Finansal Başarısızlık Tahmini. *Yüksek Lisans Tezi, Başkent Üniversitesi Sosyal Bilimler Enstitüsü*.

Jabeur, S. B., Gharıb, C., Mefteh-Mali, S. ve Arfi, W. B. (2021). Catboost Model and Artificial İntelligence Techniques for Corporate Failure Prediction. *Technological Forecasting and Social Change, 166*, 120658. doi:10.1016/j.techfore.2021.120658

K, R. K. (2015). Detection and Classification of Mammogram mages Using Principle Component Analysis and Lazy Classifiers. *nternational Journal of Innovative Research and Creative Technology, 1*(2), 173-182.

Karacan, S. ve Savcı, M. (2011). Kriz Dönemlerinde İşletmelerin Mali Başarısızlık Nedenleri. *Kocaeli Üniversitesi Sosyal Bilimler Dergisi*(21), 39-54.

Karadeniz, E. ve Öcek, C. (2019). Borsa İstanbul Konaklama İşletmelerinde Finansal Başarısızlık Riskinin Beaver Modeli Kullanılarak Analizi. *Journal of Tourism Theory and Research, 5*(2), 99-111. doi:https://doi.org/10.24288/jttr.524794

Karadeniz, E. ve Öcek, C. (2019). Finansal Başarısızlık Riski Taşıyan ile Taşımayan İşletmelerin Finansal Oranlarının Karşılaştırmalı Analizi: Borsa İstanbul Turizm İşletmelerinde Bir Araştırma. *Seyahat ve Otel İşletmeciliği Dergisi, 16*(2), 191-206. doi:10.24010/soid.600207

Karadeniz, E. ve Öcek, C. (2020). Finansal Başarısızlık Tahmin Modellerinin Geçerliliğinin Analizi: Thomas Cook Örneği. *Seyahat ve Otel İşletmeciliği Dergisi, 17*(3), 394-406. doi:10.24010/soid.719656

Kasztelnik, K. (2020). Innovative Empirical Model for Predicting National Banks' Financial Failure with Artificial Intelligence Subset Data Analysis in the United States. *Open Economics, 3*(1), 98-111. doi:10.1515/openec-2020-0106

Kayaalp, K. ve Süzen, A. A. (2018). *Derin Öğrenme ve Türkiye'deki Uygulamaları.* IKSAD Yayınevi.

Kayalı, N. ve Doğan, İ. (2018). Kurumsal Yönetim Derecelendirme Notu İle İşletmelerin Finansal Başarıları Arasındaki İlişkinin İncelenmesi. *Muhasebe ve Finansman Dergisi*(78), 111-124. doi:10.25095/mufad.412537

Keskenler, M. F., Dal, D. ve Aydın, T. (2021). Yapay Zeka Destekli ÇOKS Yöntemi ile Kredi Kartı Sahtekarlığının Tespiti. *El-Cezeri, 8*(2), 1007-1023. doi:10.31202/ecjse.908260

Kılıç, S. (2013). Doğrusal Regresyon Analizi. *Journal of Mood Disorders, 3*(2), 90-92. doi:https://doi.org/10.5455/jmood.20130624120840

Kilimci, Z. H. (2020). Borsa Tahmini için Derin Topluluk Modellleri (DTM) ile Finansal Duygu Analizi. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 35*(2), 635-650. doi:10.17341/gazimmfd.501551

Kılınç, Y. ve Seyrek, İ. H. (2012). Finansal Başarısızlık Tahmininde Yapay Sinir Ağlarının Kullanılması: İmalat Sektöründe Bir Uygulama. *1. Uluslararası Muhasebe ve Finans Sempozyumu.*

Kim, K. P., Yusof, F. ve Daud, Z. b. (2014). Multi-Dimensional Reduction Using Self-Organizing Map. *AIP Conference Proceedings, 1065*(1), 932-937. doi:10.1063/1.4887715

Kocabaş Akay, M., Canik, F., Yeşilyurt, C. ve Şahin Günkut, M. (2022). Yapay Zeka ile Kripto Para Değeri Tahmini. *Ekonomi Bilimleri Dergisi, 14*(1), 72-101. doi:10.55827/ebd.1060983

Koç, S. ve Ulucan, S. (2016). Finansal Başarısızlıkların Tespitinde Kullanılan Altman Z Yönteminin Bulanık Mantık (Anfis) Yöntemi İle Test Edilmesi: Teknoloji ve Tekstil Sektöründe Bir Uygulama. *Maliye ve Finans Yazıları*(106), 147-167. doi:10.33203/mfy.341768

Kolli Sai Poorna Chand, P. N., S, R., Rao, D. V. ve Vemparala, S. (2019). Assessment Analysis and Performance Prediction using M5 Rules. *International Journal of Recent Technology and Engineering, 8*(4), 12945-12951. doi:10.35940/ijrte.D9583.118419

Korkmaz, G. (2020). Yapay Zekâ Yöntemleriyle Siniflandirma ve Finans Sektöründe Kurumsal Müşterilere Dönük Bir Uygulama. *Akademik Yaklaşımlar Dergisi, 11*(2), 91-109.

Kozan, E. (2023). Finansal Teknolojide Yapay Zekanın Rolü.

Kuızınıenė, D., Krılavıčıus, T., Damaševičius, R. ve Maskelıūnas, R. (2022). Systematic Review of Financial Distress Identification using Artificial Intelligence Methods. *Applied Artificial Intelligence, 16*(1), 3527-3571. doi:doi.org/10.1080/08839514.2022.2138124

Kul, Ö. (2012). İşletmelerde Mali Başarısızlık Tahminlemesi İmkb'de Faaliyette Bulunan Tekstil İşletmelerine Uygulama. *Yüksek Lisans Tezi, Cumhuriyet Üniversitesi Sosyal Bililmler Enstitüsü*.

Kulalı, İ. (2016). Altman Z-Skor Modelinin Bist Şirketlerinin Finansal Başarısızlık Riskinin Tahmin Edilmesinde Uygulanması. *Uluslararası Yönetim İktisat ve İşletme Dergisi, 12*(27), 283-292. doi:10.17130/10.17130/ijmeb.2016.12.27.1076

Kurt, R., Karayılmazlar, S. ve İmren, E. (2017). Yapay Sinir Ağları İle Öngörü Modellemesi: Türkiye Kağıt-Karton Sanayi Örneği. *Bartın Orman Fakültesi Dergisi, 19*(2), 99-106. doi:10.24011/barofd.334773

Küçük, D. ve Arıcı, N. (2018). Doğal Dil İşlemede Derin Öğrenme Uygulamaları Üzerine Bir Literatür Çalışması. *Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi, 2*(2), 76-86.

Landauer, T. K., Foltz, P. W. ve Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*(2-3), 259-284. doi:10.1080/01638539809545028

Lee, H. H. ve Viviani, J.-L. (2018). Predicting Bank Failure: An İmprovement By İmplementing a Machine-Learning Approach to Classical Financial Ratios. *Research in International Business and Finance*(44), 16-25. doi:10.1016/j.ribaf.2017.07.104

Li, J., Izakian, H. ve Jamal, I. (2021). Clustering-Based Anomaly Detection in Multivariate Time Series Data. *Applied Soft Computing, 100*, 106919. doi:10.1016/j.asoc.2020.106919

Lira, M. M., Aquino, R. R., Ferreira, A. A., Carvalho, M. A., Neto, O. N. ve Santos, G. S. (2007). Combining Multiple Artificial Neural Networks Using Random Committee to Decide upon Electrical Disturbance Classification. *International Joint Conference on Neural Networks*, (s. 2863-2868). Orlando. doi:10.1109/IJCNN.2007.4371414

Liu, X.-Y., Wu, J. ve Zhou, Z.-H. (2009). Exploratory Undersampling for Class-Imbalance Learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 39*(5), 539-550. doi:10.1109/TSMCB.2008.2007853

Manzali, Y., Chahhou, M. ve Mohajir, M. E. (2017). Impure Decision Trees for Auc and Log Loss Optimization. *2017 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS).* Fez. doi:10.1109/WITS.2017.7934675

Marhas, M. K., Bhange, A. ve Ajankar, P. (2012). Anomaly Detection in Network Traffic: A Statistical Approach. *International Journal of IT, Engineering and Applied Sciences Research (IJIEASR), 1*(3), 16-20.

Medetoğlu, B. ve Tutar, S. (2023). Springate S Ve Fulmer H Skor Modelleri İle Finansal Başarısızlık Tespiti: Borsa İstanbul Tekstil, Giyim Eşyası Ve Deri Sektörü Üzerine Bir Uygulama. *Doğuş Üniversitesi Dergisi, 24*(1), 307-319. doi:10.31671/doujournal.1186265

Medsker, L. ve Jain, L. (2001). *Recurrent Neural Networks Desing and Applications.* CRC Press.

Mohan, V. (2013). Comparative Analysis of Bayes and Lazy Classification Algorithms. *International Journal of Advanced Research in Computer and Communication Engineering, 2*(8), 3118-3124.

Nanga, S., Bawah, A. T., Acquaye, B. A., Bılla, M.-I., Baeta, F. D., Odıı, N. A., . . . Nsiah, A. D. (2021). Review of Dimension Reduction Methods. *Journal of Data Analysis and Information Processing, 9*, 189-231. doi:10.4236/jdaip.2021.93013

Ochella, S. ve Shafiee, M. (2020). Performance Metrics for Artificial Intelligence (AI) Algorithms Adopted in Prognostics and Health Management (PHM) of Mechanical Systems. *2020 International Symposium on Automation, Information and Computing (ISAIC 2020).* Beijing, China. doi:10.1088/1742-6596/1828/1/012005

Öcek, C. (2018). Konaklama İşletmelerinde Finansal Başarısızlık Riskinin Ölçümü: Türkiye Ve Avrupa'daki Halka Açık Konaklama İşletmelerinin Karşılaştırmalı Analizi. *Yüksek Lisans Tezi, Mersin Üniversitesi Sosyal Bilimler Enstitüsü*.

Öner, A. (2018). Finansal Bilgi Düzeyinin Finansal Başarısızlık Üzerindeki Etkinliğinin Analizi: Erzincan İli Kobi Yöneticileri Üzerinde Bir Uygulama. *Yüksek Lisans Tezi, Erzincan Binali Yıldırım Üniversitesi Sosyal Bilimler Enstitüsü*.

Örnek, Ö., Vatan, S., Sarıoğlu, S. ve Yazıcı, A. (2018). Trafik Ağlarında Anomali Tespiti. *Eskişehir Osmangazi Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi, 26*(3), 132-138. doi:10.31796/ogummf.440285

Özekes, S. (2003). Veri Madenciliği Modelleri ve Uygulama Alanları. *İstanbul Ticaret Üniversitesi Dergisi, 2*(3), 65-82.

Özgür, A. S., Tarhan, Ç., Komesli, M. ve Tecim, V. (2023). Yapay Zeka Teknikleri Kullanılarak Proje Üretim Sistemlerinin Tasarımı ve Geliştirilmesi. *Journal of Information Systems and Management Research, 5*(1), 1-12. doi:10.59940/jismar.1214440

Öztemel, E. (2012). *YAPAY SİNİR AĞLARI.* İstanbul: PAPATYA YAYINCILIK.

Öztürk, K. ve Şahin, M. E. (2018). Yapay Sinir Ağları ve Yapay Zekâ'ya Genel Bir Bakış. *Takvim-i Vekayi, 6*(2), 2-25.

Öztürk, M. (2022, 04 13). *python-ile-siniflandirma-analizleri-rastgele-orman-random-forest-algoritmasi*. miracozturk.com: https://miracozturk.com/ adresinden alındı

Permai, S. D. ve Tanty, H. (2018). Linear regression model using bayesian approach for energy performance of residential building. *Procedia Computer Science*(135), 671-677. doi:10.1016/j.procs.2018.08.219

Sakarya, Ş. ve Yılmaz, Ü. (2019). Derin Öğrenme Mimarisi Kullanarak BİST30 İndeksinin Tahmini. *European Journal of Educational and Social Sciences, 4*(2), 106-121.

Sakız, B. ve Ünkaya, G. (2018). Hava Yolu Taşımacılığı Sektöründe İflas Riski - Yapay Sinir Ağları İle Airscore Tahmini. *Öneri Dergisi, 13*(50), 159-172. doi:10.14783/maruoneri.v13i38778.408996

Salakhutdınov, R. ve Hinton, G. (2009). Deep Boltzmann Machines. *Proceedings of the Twelth International Conference on Artificial Intelligence and Statistics, 5*, 448-455.

Salur, M. N. (2021). Finansal Başarısızlık Tahmininde Yapay Sinir Ağları Modelinin Kullanımı: Borsa İstanbul'da Bir Uygulama. *Journal of Economics Finance and Accounting, 8*(1), 17-30. doi:10.17261/Pressacademia.2021.1375

Saxena, A., Gupta, A. ve Mukerjee, A. (2004). Non-linear Dimensionality Reduction by Locally Linear Isomaps. *Neural Information Processing* (s. 1038-1043). Berlin, Heidelberg: Springer, Berlin, Heidelberg. doi:10.1007/978-3-540-30499-9_161

Sayım, F. (2022). *İşletmelerde ve Yönetimde Yapay Zeka.* Ankara: Seçkin Yayıncılık.

Selçik, S. (2019). Öngörü Teknikleri ile Finansal Başarısızlık Tahmini: BIST Üzerine Bir Uygulama. *Yüksek Lisans Tezi, İstanbul Üniversitesi Sosyal Bilimler Enstitüsü*.

Selimoğlu, S. ve Orhan, A. (2015). Finansal Başarısızlığın Oran Analizi ve Diskriminant Analizi Kullanılarak Ölçümlenmesi: BİST'de İşlem Gören Dokuma, Giyim Eşyası ve Deri İşletmeleri Üzerine Bir Araştırma. *Muhasebe ve Finansman Dergisi*(66), 21-40. doi:10.25095/mufad.396529

Seyrek, İ. H. ve Ata, H. A. (2010). Veri Zarflama Analizi ve Veri Madenciliği ile Mevduat Bankalarında Etkinlik Ölçümü. *BDDK Bankacılık ve Finansal Piyasalar Dergisi, 4*(2), 67-84.

Shelke, M., Deshmukh, D. P. ve Shandılya, V. (2017). A Review on Imbalanced Data Handling Using Undersampling and Oversampling Technique. *International Journal of Recent Trends in Engineering ve Research (IJRTER), 3*(4), 444-449. doi:10.23883/IJRTER.2017.3168.0UWXM

Sneha, S. N. ve Gangil, T. (2019). Analysis of Diabetes Mellitus for Early Prediction Using Optimal Features Selection. *Journal of Big Data, 6*(1), 1-19. doi:10.1186/s40537-019-0175-6

Söylemez, Y. ve Türkmen, S. Y. (2017). Yapay Sinir Ağları Modeli İle Finansal Başarısızlık Tahmini. *Finans Ekonomi ve Sosyal Araştırmalar Dergisi, 2*(4), 270-284. doi:10.29106/fesa.364323

Şahin, T. ve Yangil, F. M. (2023). Finansal Başarısızlık Konusunda Yapılan Lisansüstü Tez Çalışmalarına Yönelik Bir İçerik Analizi. *Kastamonu Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 25*(2), 509-5027.

Şanlı, E. (2018). Yapay Sinir Ağı Kontrollü Otonom RC Araç Uygulaması. *Yapay Sinir Ağı Kontrollü Otonom RC Araç Uygulaması*.

Şeker, A., Diril, B. ve Balık, H. H. (2017). Derin Öğrenme Yöntemleri ve Uygulamaları Hakkında Bir İnceleme. *Gazi Mühendislik Bilimleri Dergisi, 3*(3), 47-64.

Şişmanoğlu, G., Koçer, F., Önde, M. A. ve Şahingöz, Ö. K. (2020). Derin Öğrenme Yöntemleri ile Borsada Fiyat Tahmini. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi, 9*(1), 434-445. doi:10.17798/bitlisfen.571386

Taffler, R. (1982). Forecasting Company Failure in the UK Using Discriminant Analysis and Financial Ratio Data. *Journal of the Royal Statistical Society. Series A (General), 15*(3), 342-358. doi:10.2307/2981867

Tas, O. ve Mert, H. (2019). Denetimde Yapay Zeka Uygulaması. *PressAcademia Procedia, 9*, 65-68. doi:10.17261/Pressacademia.2019.1067

Tekin, B. ve Gör, Y. (2022). Finansal Başarısızlık Tahmin Modelleri Ve Bankacılık Sektörü Mali Tablolari Üzerinden Bir Uygulama: Altman Ve Springate Modelleri. *Adıyaman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*(40), 373-404. doi:10.14520/adyusbd.992296

Terzi, S. (2011). Finansal Rasyolar Yardımıyla Finansal Başarısızlık Tahmini: Gıda Sektöründe Ampirik Bir Araştırma. *Çukurova Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 15*(1), 1-18.

Tsai, C.-F. ve Lin, W.-C. (2021). Feature Selection and Ensemble Learning Techniques in One-Class Classifiers: An Empirical Study of Two-Class Imbalanced Datasets. *IEEE Access, 9*, 13717-13726. doi:10.1109/ACCESS.2021.3051969

Türk, Ö. (2022). MR Görüntülerinden Alzheimer Tespitinde Boyut Azaltma ve Derin Öğrenme Yaklaşımlarının Karşılaştırılması. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi, 13*(3), 485-491. doi:10.24012/dumf.1141233

Ural, K. (2014). İşletmelerde Finansal Başarısızlık Tahminlemesi ve Borsa İstanbul'da Faaliyet Gösteren Gıda Şirketlerinde Uygulama. *Yüksek Lisans Tezi, Gediz Üniversitesi Sosyal Bilimler Enstitüsü,*.

Ural, K., Gürada, Ş. ve Önemli, M. B. (2015). Lojistik Regresyon Modeli İle Finansal Başarısızlık Tahminlemesi: Borsa İstanbul'da Faaliyet Gösteren Gıda, İçki Ve

Tütün Borsa İstanbul'da Faaliyet Gösteren Gıda, İçki Ve Tütün Borsa İstanbul'da Faaliyet Gösteren Gıda, İçki Ve Tütün Şirketlerinde Uygulam. *Muhasebe ve Finansman Dergisi*(67), 85-100. doi:10.25095/mufad.396578

Uzun, E. (2005). İşletmelerde Finansal Başarısızlığın Teorik Olarak İrdelenmesi. *Muhasebe ve Finansman Dergisi*(27), 158-168.

V.S, S. ve Surendran, S. (2015). A Review of Various Linear and Non Linear Dimensionality Reduction Techniques. *International Journal of Computer Science and Information Technologies, 6*(3), 2354-2360.

Wang, S., Lu, J., Gu, X., Du, H. ve Yang, J. (2016). Semi-Supervised Linear Discriminant Analysis for Dimension Reduction and Classification. *Pattern Recognition, 57*, 179-189. doi:10.1016/j.patcog.2016.02.019

Xu, M., Watanachaturaporn, P., Varshney, P. K. ve Arora, M. K. (2005). Decision Tree Regression for Soft Classification of Remote Sensing Data. *Remote Sensing of Environment, 3*(97), 322-336. doi:10.1016/j.rse.2005.05.008

Xu, X. ve Wang, Y. (2009). Financial Failure Prediction Using Efficiency As A Predictor. *Expert Systems with Applications, 1*, 366-373.

Yazıcı, A. C., Öğüş, E., Ankaralı, S., Canan, S., Ankaralı, H. ve Akkuş, Z. (2007). Yapay Sinir Ağlarına Genel Bakış. *Türkiye Klinikleri Tıp Bilimleri Dergisi, 27*(1), 65-71.

Yıldız, A. (2022). Finans Alaninda Yapay Zekâ Teknolojisinin Kullanimi: Sistematik Literatür İncelemesi. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*(52), 47-66. doi:10.30794/pausbed.1089134

Yıldız, E. ve Sevim, Y. (2016). 016 National Conference on Electrical, Electronics and Biomedical Engineering (ELECO). *Comparison of Linear Dimensionality Reduction Methods on Classification Methods*, (s. 161-164). Bursa, Türkiye.

Yıldız, K., Çamurcu, Y. ve Doğan, B. (2010). A Comperative Analize of Principal Component Analysis and NonNegative Matrix Factorization Techniques in Data Mining. *12. Akademik Bilişim Konferansı*, (s. 207-213). Muğla,Türkiye.

Yıldız, Ş. ve Gürkan, S. (2022). Finansal Başarısızlık Tahmin Modellerinin Karşılaştırılması: Borsa İstanbul Turizm Şirketlerinde Bir Araştırma. *Ekonomi İşletme ve Yönetim Dergisi, 6*(2), 235-262. doi:10.7596/jebm.31122022.005

Yürük, M. F. ve Ekşi, İ. H. (2019). Yapay Zekâ Yöntemleri İle İşletmelerin Finansal Başarisizliğinin Tahmin Edilmesi: Bist İmalat Sektörü Uygulaması. *Mukaddime, 10*(1), 393-422. doi:10.19059/mukaddime.533151

Zahıd, M., Chen, Y., Jama, A. ve Mamadou, C. Z. (2020). Freeway Short-Term Travel Speed Prediction Based on Data Collection Time-Horizons: A Fast Forest Quantile Regression Approach. *Sustainability, 12*(2), 2-19. doi:10.3390/su12020646

Zhang, Z., Wu, C., Qu, S. ve Chen, X. (2022). An Explainable Artificial İntelligence Approach for Financial Distress Prediction . *Information Processing & Management, 59*(4). doi:10.1016/j.ipm.2022.102988

Zhou, Z.-H. ve Liu, X.-Y. (2006). Training Cost-Sensitive Neural Networks with Methods Addressing the Class Imbalance Problem. *IEEE Transactions on Knowledge and Data Engineering, 18*(1), 63-77. doi:10.1109/TKDE.2006.17

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı Soyadı	:	Çağatay Sinoplu
Doğum Yeri ve Tarihi	:	Bartın / 25.06.1996

Eğitim Durumu

Lisans Öğrenimi	:	Anadolu Üniversitesi İşletme Fakültesi İşletme Bölümü
Yüksek Lisans Öğrenimi	:	
Bildiği Yabancı Diller	:	İngilizce
Bilimsel Faaliyet/Yayınlar	:	Yatırımcıların Risk İştahları ve Hisse Senedi Getirileri Arasındaki İlişki: Borsa İstanbul'da Bir Araştırma. Muhasebe ve Finansman Dergisi(98), 155-170. https://doi.org/10.25095/mufad.1211699 Ekonomik Özgürlük Endeksi İle BİST Mali Endeksi Arasında Nedensellik Araştırması: Vektör Hata Düzeltme Modeli İle Analiz. JOEEP: Journal of Emerging Economies and Policy, 8(2), 445-456.
Aldığı Ödüller	:	

İş Deneyimi

Stajlar	:	Bartın Belediyesi İmar ve Şehircilik Müdürlüğü - Eylül 2015 – Ocak 2016
Projeler ve Kurs Belgeleri	:	ISO 9001 : 2015 Kalite Yönetim Sistemi Standardı Eğitimi – Udemy – 01/2024 Power BI İstanbul – Uçtan Uca Power BI Eğitimi – 02/2023 2021-2022 TÜBİTAK 2209-A Üniversite Öğrencileri Araştırma Projeleri Destekleme Programı Proje Kabulü Yönetim ve Organizasyon Eğitimi - İstanbul İşletme

Enstitüsü – 10/2020
E-ticaret Eğitimi – Bemar Kariyer Okulu – 01/2017
Girişimcilik Sertifikası – KOSGEB – 01/2017
Grafik Animasyon Kursu – MEB – 06/2012

Çalıştığı Kurumlar : Reymer Tekstil – İnsan Kaynakları ve Satın Alma Sorumlusu - Nisan 2022 – Halen

İletişim

E-Posta Adresi :

Tarih : 25/10/2024 (Tez Savunma Tarihi)