

Sosyal Sorumlu Yatırım Bağlamında Pay Senedi Getirisinin Makine Öğrenmesi Yöntemleri ile Tahmin Edilmesi: Borsa İstanbul Örneği

(Estimating using Machine Learning Methods of Stock Return in the Context of Socially Responsible Investment: Borsa İstanbul Example)

Barış AKSOY  ^a

^a Sivas Cumhuriyet Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Bankacılık ve Finans Bölümü, Sivas, Türkiye.
baksoy@cumhuriyet.edu.tr

MAKALE BİLGİSİ	ÖZET
<p>Anahtar Kelimeler: Sosyal Sorumlu Yatırım Pay Senedi Getiri Tahmini Makine Öğrenmesi Yöntemleri BIST Sürdürülebilirlik Endeksi Borsa İstanbul</p> <p>Gönderilme Tarihi 15 Ağustos 2020 Revizyon Tarihi 27 Kasım 2020 Kabul Tarihi 11 Aralık 2020</p> <p>Makale Kategorisi: Araştırma Makalesi</p>	<p>Amaç – Bu çalışmada pay senetleri Borsa İstanbul 30/100 ve sürdürülebilirlik endeksindeki imalat sanayi şirketlerinin yılsonu mali tablo verileri ve ekonomik göstergeler kullanılarak örnek kapsamındaki şirketlerin bir yıl sonraki ortalama pay senedi getirilerinin tahmin edilmesi amaçlanmıştır. Çalışmada kullanılan yöntemlerin tahmin performansının, yöntemlerin ayırt edici özellikleri altında karşılaştırılması çalışmanın diğer amacını oluşturmaktadır.</p> <p>Yöntem – Örnek kapsamındaki 13 şirketin 2010-2018 döneminde yılsonu mali tabloları ve yıllık ortalama ekonomik göstergeler alınarak 2011-2019 yıllık ortalama pay senedi getirileri Yapay sinir ağları (ANN), Sınıflandırma ve Regresyon Ağaçları (CART) ve K En Yakın Komşu Algoritması (KNN) yöntemleri ile tahmin edilmiştir. Çalışmada optimal veri dağılımı için 10 katlı çapraz doğrulama yöntemi ve verilen parametre aralıklarında en yüksek tahmin sonucu veren modelin belirlenmesi amacıyla parametre optimizasyonu kullanılmıştır. Yöntemlerin performansı ROC eğrisi ile karşılaştırılmıştır.</p> <p>Bulgular – Analiz sonucunda CART (%94,87), ANN (%94,02) ve KNN (%92,31) genel tahmin ve sınıflandırma doğruluğu elde etmiştir. %20’den %50 ye kadar negatif getiri sağlayan örnekleri ANN (%80,00), CART (%100), KNN (%100) oranında doğru tahmin etmesi ve %20’den %50’ye kadar pozitif getiri tahmininde KNN (%100), ANN (%96,55), CART (96,55) tahmin doğruluğu elde etmesi dikkate değer bulunmuştur. Araştırmada CART karar ağacının oluşturulmasında verileri bölen önemli değişkenler olarak “Özsermaye Kârlılık Oranı” ve “Piyasa Değeri /Defter Değeri” bulunmuştur.</p> <p>Tartışma – Bu çalışmada kullanılan tüm yöntemler %90,00’in üzerinde genel tahmin doğruluğu elde ederek sınıflandırma ve tahmin gerçekleştirmiştir. Gaganis (2009) çalışmasında %75’in üzerinde tespit doğruluğunun sosyal bilimler alanında iyi bir sonuç olduğunu belirtmiştir. Gaganis (2009) sınıflandırmasına göre bu çalışmadaki ANN, CART ve KNN analiz sonuçları çok iyi olarak değerlendirilebilir.</p>
ARTICLE INFO	ABSTRACT
<p>Keywords: Socially Responsible Investment Prediction of Stock Return Machine Learning Methods BIST Sustainability Index Borsa İstanbul</p> <p>Received 15 August 2020 Revised 27 November 2020 Accepted 11 December 2020</p> <p>Article Classification: Research Article</p>	<p>Purpose – In this study, it is aimed to predict the average stock returns of the companies within the scope of the sample in the next year by using the year-end financial statement data and economic indicators of the manufacturing industry companies in BIST 30/100 indexes and the sustainability index. The other aim of the study is to compare the prediction performance of the methods used in the study under the distinctive features of the methods.</p> <p>Design/methodology/approach – 2011-2019 average annual stock returns were predicted using Artificial Neural Networks (ANN), Classification and Regression Trees (CART) and K Nearest Neighbor Algorithm (KNN) after taking the year-end financial statements and annual average economic indicators of 13 companies within the scope of the sample for the period of 2010-2018. In the study, 10-fold cross validation method was used for optimal data distribution and parameter optimization to determine the model with the highest prediction success in the given parameter ranges. The performance of the methods was compared with the ROC curve.</p> <p>Findings – As a result of the analysis, CART (94.87%), ANN (94.02%) and KNN (92.31%) achieved overall prediction and classification accuracy. It was remarkable that ANN, CART, KNN predicted samples with negative returns from 20% to 50% with 80%, 100% and 100% prediction accuracy, respectively and KNN, ANN, CART predicted samples with positive returns from 20% to 50% with 100%, 96.55% and 96.55% prediction accuracy, respectively. “Return On Equity” and “Market</p>

Önerilen Atf/ Suggested Citation

Aksoy, B. (2020). Sosyal Sorumlu Yatırım Bağlamında Pay Senedi Getirisinin Makine Öğrenmesi Yöntemleri ile Tahmin Edilmesi: Borsa İstanbul Örneği, *İşletme Araştırmaları Dergisi*, 12 (4), 3859-3878.

Value/Book Value" were found as important variables dividing the data in the creation of the CART decision tree.

Discussion – All methods used in this study achieved classification and prediction success by obtaining more than 90% overall prediction accuracy. Gaganis (2009) stated that prediction accuracy of over 75% was a good result in the field of social sciences. According to Gaganis (2009) classification, ANN, CART and KNN analysis results in this study can be evaluated as very good.

1. Giriş

Finansal piyasalar vade uzadıkça belirsizlik içermektedir. Piyasanın belirsizliğini azaltmak için farklı tahmin teknikleri kullanılmaktadır. Belirsizlikleri azaltmayı amaçlayan, pazar dinamiklerini yakalayabilen modellerin tasarımı akademik araştırmacıların ve piyasada görev yapan profesyonellerin ilgisini çekmektedir. Finansal analistler ve yatırımcılar, yaptıkları finansal işlemlerden getiri elde etmek amacıyla borsa trendlerini gözlemlemekte, karar vermek için tarihsel fiyat verilerini ve ekonomik göstergeleri analiz etmektedirler (Oliveira vd., 2013: 7596). Finansal piyasalar karmaşık, evrimsel yapıda olan ve doğrusal olmayan bir dinamik sisteme sahiptir. Finansal tahmin çalışma alanında, veri yoğunluğu, gürültü, durağan olmayan, yapılandırılmamış, yüksek derecede belirsizlik ve gizli ilişkiler bulunmaktadır. Siyasi olaylar, genel ekonomik koşullar ve finansal piyasada alım/satım yapanların beklentileri dâhil olmak üzere birçok faktör etkileşime girmektedir. Bu nedenle, finans piyasasında fiyat hareketlerini tahmin etmek oldukça zordur (Huang vd., 2005: 2514).

Borsa İstanbul (BIST) gibi gelişmekte olan piyasalarda pay senedi fiyat dalgalanmalarının sık yaşanması görece doğal karşılanabilir. Yatırımcılar, fiyat dalgalanmaları karşısında risk altında olduklarından pay senedi fiyatlarının önceden tahmin edilmesine önem vermekte ve belirsizliklerin azaltılmasıyla mevcut ve olası risklerin portföyleri üzerindeki olumsuz etkilerini minimize etmeye çalışmaktadırlar. Araştırmacılar tarafından kullanılan birçok yöntem bulunmakla birlikte günümüzde makine öğrenme yöntemlerinin kullanımı zamanla yaygınlaşmaktadır. (Akcan ve Kartal, 2011: 31). Günümüzde bazı yatırımcılar ve fonlar yatırım kararı alırken sürdürülebilirlik raporları ve sürdürülebilirlik endeksi gibi araçlardan yararlanmaktadırlar. Sürdürülebilirlik endeksleri, sürdürülebilirlik performansını değerlendirmelerinde yatırımcıların ihtiyaç duydukları bilgileri sağlayan finansal piyasa endeksleridir (Özdemir ve Pamukçu, 2016: 19). Sosyal sorumlu yatırımın (SRI: Socially Responsible Investment) en yaygın şekli, paydaşlarının tasarruflarını kurumsal sosyal sorumluluk kriterlerine göre seçilen şirketlere yönlendiren sosyal sorumluluk sahibi yatırım veya emeklilik fonlarında hisse satın almaktır. Sosyal sorumlu yatırım sadece yatırımın getirisi ile ilgili değildir; sosyal, çevresel ve kurumsal yönetim kriterlerine uygun olup olmadıklarını da dikkate almaktadır(Charlo vd. 2015: 278).

2. Çalışmanın Amacı

Bu çalışmada pay senetleri Borsa İstanbul 30/100 ve sürdürülebilirlik endeksindeki imalat sanayi şirketlerinin yılsonu mali tablo verileri ve ekonomik göstergeler kullanılarak örnek kapsamındaki şirketlerin bir yıl sonraki ortalama pay senedi getirilerinin tahmin edilmesi amaçlanmıştır. Çalışmada kullanılan yöntemlerin tahmin performansının, yöntemlerin ayırt edici özellikleri altında karşılaştırılması çalışmanın diğer amacını oluşturmaktadır.

Literatürde, yıllık mali oranlar ve ekonomik göstergeler kullanılarak örnek kapsamındaki şirketlerin bir yıl sonraki ortalama getirilerinin (-%20/-%50, 0/-%20, 0/+%20, +%20/+%50 ve +%50 fazla getiri) 5 kategoriye ayrıldıktan sonra CART, ANN ve KNN yöntemlerinin kullanıldığı ve her üç yönteme ait modellerin tahmin sonuçlarının, modellerin ayırt edici özellikleri altında karşılaştırmasını gösteren bir çalışmaya rastlanmamıştır. Bu nedenle çalışmanın, finansal piyasalarda işlem yapan yatırımcılar, analistler ve araştırmacılar için t+1 dönemindeki ortalama pay senedi getirisinin tahmin edilmesinde önemli yararlar elde edebileceklerinden hareketle literatüre katkısı olacağı düşünülmektedir.

Çalışmada optimal veri dağılımı için 10 katlı çapraz doğrulama yöntemi ve verilen parametre aralıklarında en yüksek sınıflandırma ve tahmin başarısı gösteren modelin belirlenmesi amacıyla parametre optimizasyonu gerçekleştirilmiştir. ROC eğrisi ile çalışmada kullanılan yöntemlerin karşılaştırılması verilmiştir. Bu çalışma altı bölümden oluşmaktadır. Birinci bölümde giriş, ikinci bölümde çalışmanın amacı, üçüncü bölümde BIST Sürdürülebilirlik Endeksi, dördüncü bölümde literatür incelemesi, beşinci bölümde yöntem ve veri, altıncı bölümde sonuç yer almaktadır.

3. BİST Sürdürülebilirlik Endeksi

Günümüzde şirketler kâr elde etme gibi yalnızca ekonomik yönüyle değil, sosyal ve çevre konularında da rakip şirketlere göre fark yarattığını ve ilgili konulara duyarlı davrandığını gösterme ihtiyacı duymaktadır. Şirketlerin sosyal sorumluluğu ile ilgili toplumun beklentisi, finansal piyasalarda sürdürülebilirlik endeksleri ile karşılık bulmuştur (Önder, 2017: 938). Sürdürülebilir büyüme, Brundtland Raporu'nda (1987) bulunan ve genel olarak benimsendiği şekliyle "gelecek nesillerin imkânlarını azaltmadan ve kötüleştirmeden mevcut neslin ihtiyaçlarını karşılama yolu" şeklinde tanımlanabilir. Hem firmalar hem de yatırımcılar sürdürülebilirlik kriterlerini dikkate alan stratejilerin uzun vadeli değer yaratma kapasitesine sahip olduğuna inanmaktadır. Bu tür eylemler yatırımcıların ilgisini uyandırmış ve finansal piyasalara bağlı sürdürülebilirlikle ilgili endekslerin ortaya çıkmasına neden olmuştur (Lo'pez, 2007: 296). Çevresel, sosyal ve yönetim (ESG: Environmental, Social, Government) parametrelerinde üstün performansları açısından incelenen şirketler, küresel ve ulusal borsalarda sunulan sürdürülebilirlik endeksini oluşturmaktadır. Yatırımcıların yatırım yapacakları şirketler ile ilgili performans beklentilerine finansal olmayan sürdürülebilirlik performansı da dâhil edilmiş sosyal sorumlu yatırım (SRI: Socially Responsible Investment) olarak ifade edilen yatırım anlayışı ortaya çıkmıştır (Gök ve Özdemir, 2017: 88). Sürdürülebilir kalkınma ve sosyal sorumluluk sahibi yatırımlar giderek artan sayıda yatırımcının yatırım kararlarında sosyal ve çevresel faktörleri göz önünde bulundurması, şirketlerin finansal kaynaklara erişimde avantaj elde etmek için Kurumsal Sosyal Sorumluluk (KSS) uygulamalarını benimsemeye daha yatkın hale getirmiştir (Charlo vd. 2015: 277).

Bazı yatırımcılar "iyi" ESG performansını, bir şirketin uzun vadeli sürdürülebilirlik konusundaki bilincini yansıtan bir yönetim kalitesi olarak algılamakta ve dolayısıyla bu tür "sürdürülebilir" şirketlere yatırım yapmak isteyebilmektedirler. Değer sistemlerine veya etik ilkelerine uygun şirketlere yatırım yapan yeni bir yatırımcı grubu bulunmaktadır. Bazı yatırımcılar, tütünle ilgili ürünler, silahlar ve mühimmat üretimi yapan firmalar gibi doğrudan veya dolaylı olarak toplum üzerinde zararlı etkisi olan şirketlere yatırım yapmamaktadır. Bu durum, yalnızca sürdürülebilirlik parametreleriyle (yani çevre, sosyal ve yönetim performansı) son derece uyumlu şirketlere yapılan yatırıma odaklanan münhasır yatırım fonlarının veya hisse senedi endekslerinin ortaya çıkmasına neden olmuştur. Dolayısıyla, sosyal sorumluluk sahibi yatırım (SRI) tüm paydaşlara, yani hissedarlara, topluma, çalışanlara, müşterilere, hükümete ve ayrıca doğal çevreye karşı "etik" kurumsal davranış sergileyen şirketlere yatırım yapmak anlamına gelmektedir (Sudha, 2015: 1330). Gelecekte sürdürülebilirlik raporu yayınlayan şirketlerin artması sonucunda paydaşların şirketlere güveni artacağından yatırımcıların bu şirketlere yöneleceği beklenmektedir (Altınay vd. 2017: 224). Borsa İstanbul'da sürdürülebilirlik endeksi 2014 yılından itibaren oluşturulmaya başlamıştır. Bu nedenle sürdürülebilirlik endeksi ile ilgili çalışmalar ve Türkiye'de SRI fonlarının gelişimi çok yenidir. BIST Sürdürülebilirlik Endeksi çevre, sosyal ve yönetim alanında araştırmalar ve bağımsız analizler yapan EIRIS (Ethical Investment Research Services Limited) tarafından yapılmaktadır. EIRIS uluslararası sürdürülebilirlik kriterlerine göre hesaplama yapmakta ve değerlemelerde sadece şirketlerin "kamuya açık" bilgilerini kullanmaktadır (Önder, 2017: 941-942).

4. Literatür İncelemesi

Pay senedi getiri tahmini, yatırımcıların ellerindeki fonları gelir getirici araçlarda değerlendirdikleri fonlardan belirli oranda getiri elde etme potansiyelinin varlığı nedeniyle yatırımcılar, analistler ve araştırmacılar tarafından yoğun bir ilgi alanı haline getirilmiştir. Etkin piyasa hipotezine göre, pay senedi fiyatının veya piyasa endeksinin davranışını etkileyebilecek tüm bilgilerin mevcut piyasa fiyatına dâhil edilmiş olması gerektiğinden, bu tür tahmin girişimlerinin sonuçsuz kalacağı belirtilmektedir. Borsaların gelecekteki davranışlarının bir dereceye kadar doğrulukla tahmin edilmesinin mümkün olduğunu göstererek etkin piyasa hipotezini sorgulayan çeşitli çalışmalar yapılmıştır (Nair vd., 2012: 1).

Anbalagan ve Maheswari (2015) Hindistan'daki Bombay Menkul Kıymetler Borsası'nda (BSE) listelenen pay senetleri için Bulanık Metagraf (FM) temelli karar verme, sınıflandırma ve tahmin modeli önermişlerdir. Önerilen FM tabanlı modelden elde edilen sonuçların çok düşük risk hatası ile tatmin edici olduğu bulunmuştur. Analizler sonucunda Random Walk (RW) Modelinin %50,00, Neural Networks (ANN) %69,00, SVM %73,00 ve FM %75,00 tahmin doğruluğuna ulaşmıştır.

Enke ve Thawornwong (2005) çalışmalarında genelleştirilmiş regresyon, olasılıklı ve çok katmanlı ileri beslemeli sinir ağlarını pay senedi tahmini için kullanmışlardır. Mart 1976'dan Aralık 1999'a kadar toplam 286 dönem ve 31 finansal ve ekonomik değişken için aylık veriler toplanmış ve analiz edilmiştir. Analiz sonucunda sinir ağlarının, özellikle sınıflandırma modellerinin, gelecekteki pay senedi getirilerinin bir kısmını doğru bir şekilde tahmin etmede daha yüksek performans gösterdiği belirtilmiştir.

Boyacıoğlu ve Avcı (2010) çalışmalarında uyarlanabilir Ağ Tabanlı Bulanık Çıkarım Sistemi (ANFIS) ile Borsa İstanbul (BIST) pay senedi fiyat endeksinin getirisini tahmin etmişlerdir. Girdi değişkenleri olarak altı makroekonomik değişken ve üç endeks kullanmışlardır. ANFIS modelinin %98,30 doğruluk oranıyla BIST Ulusal 100 Endeksinin aylık getirisini başarılı bir şekilde tahmin ettiği bulunmuştur.

Adebiyi vd. (2012) çalışmalarında pay senedi fiyat tahmininde daha yüksek doğruluk oranı elde etmek için geri ilerleme algoritması ile eğitilmiş ileri beslemeli çok katmanlı bir algılayıcı sinir ağı tahmin modeli oluşturmuşlardır. Elde edilen ampirik sonuçlar, teknik analiz yönteminden daha iyi performans gösteren hibrid yöntem ile günlük pay senedi fiyat tahmini için yüksek düzeyde doğruluk gösterdiği sonucuna ulaşmışlardır.

Yetiş vd. (2014) çalışmalarında NASDAQ'nun pay senedi değerinin, belirli bir pay senedi piyasası girdi parametrelerine sahip ANN genelleştirilmiş ileri besleme ağlarını kullanmışlardır. Çalışmada NASDAQ Borsası endeksinin reel döviz kuru değeri alınmıştır. Beş giriş, giriş katmanında 10 nöron ve çıkış katmanında bir nöron ile eğitim durumunda kaydedilen ağı performansının %99,00 olarak bulunmuştur.

Patel vd. (2015) çalışmalarında Hint borsalarından CNX Nifty, S&P BSE Sensex, Infosys Ltd. ve Reliance Industries'in on yıllık geçmiş verilerine (2003-2012) dayanarak pay senedi ve pay senedi fiyat endeksi hareketini tahmin etmek için ANN, SVM, Random Forest ve Naive-Bayes algoritmalarının tahmin performansını karşılaştırmışlardır. Analizler, Naive Bayes (Gauss işlemi) modelinin %73,30 doğrulukla en az performans ve rastgele ormanın %83,56 doğrulukla en yüksek performans sergilediğini göstermektedir. Modellerin performansı, trend belirleyici veriler kullanıldığında ANN %86,69, SVM %89,33, Random Forest %89,98 ve Naive Bayes %90,19'lük doğruluğa sahip olduğu sonucu elde edilmiştir.

Özçalıcı (2017), çalışmasında Goodyear, Amazon ve Wal-Mart isimli şirketler için 12 teknik gösterge ile ANN ve Aşırı Öğrenme Makineleri (AÖM) yöntemini kullanarak ANN modelinde pay senedi fiyat yönü %56,00 oranında doğru bir şekilde önceden tahmin edilmiştir. Üç pay senedinde AÖM yöntemi ANN yönteminden daha yüksek tahmin oranları sergilerken, endekste ise ANN modelinin daha iyi isabet performansı sergilediği belirlenmiştir. Çalışmada, AÖM yöntemi %59,32 oranında pay senedi fiyatındaki yönü doğru bir şekilde önceden tahmin edebildiği belirtilmiştir.

Yang vd. (2017) çalışmalarında Çin borsa endeksini (Şangay bileşik endeksi ve SZSE bileşen endeksi) modellemek ve tahmin etmek için derin sinir ağı topluluğu kullanmışlardır. Günlük eğilim tahminlerinin doğruluğu, yüksekten düşüğe Şangay endeksi için sırasıyla, %74,15, %74,15, %71,34 ve SZSE bileşen endeksi için sırasıyla %75,95, %73,95, %72,34 olarak bulunmuştur. Çalışmada önerilen modelin Çin borsalarını kısmen modelleyebileceği ve tahmin edebileceği belirtilmiştir. Ayrıca kısa zaman dönemi öncesi tahminlerin tatmin edici olmadığı sonucuna ulaşmışlardır.

Zhong ve Enke (2017) çalışmalarında S&P 500 Endeksi ETF (SPY) getirisinin günlük yönünü 60 finansal ve ekonomik özelliğe göre tahmin etmek için bir veri madenciliği süreci yürütmüşlerdir. Özetle, ANN - PCA modellerini kullanan madencilik süreci, FRPCA (Fuzzy Robust Principal Component Analysis) ve KPCA'yı (Kernel-Based Principal Component Analysis) içeren madencilik sürecine kıyasla ertesi gün SPY yönü için biraz daha yüksek tahmin doğruluğu verdiği sonucuna ulaşılmıştır.

Borsadaki yüksek doğrusal olmama durumu nedeniyle, pay senedi fiyat hareketlerini geleneksel tahmin yöntemleriyle açıklamak zordur. Tahmin zorluğu insan davranışını modellemenin karmaşıklığında yatmaktadır. Son çalışmalar doğrusal olmayan modellerin volatil borsaları iyi simüle edebildiğini ve borsa eğilimi araştırmasında geleneksel lineer modellere göre daha iyi tahmin sonuçları üretebildiğini ortaya koymaktadır (Imandoust ve Bolandraftar, 2014: 106). Sonuç olarak, son birkaç yıl içinde çeşitli araştırmacılar ve finansal analistler tarafından gelecekteki pay senedi getirileri veya endeksleri hakkında daha iyi tahminler üretmek için finansal piyasa hareketlerinde doğrusallığın bulunmaması nedeniyle makine öğrenme yöntemleri ile analizler yaygınlaşmaya başlamıştır (Enke ve Thawornwong, 2005: 928).

5. Yöntem ve Veri

Ampirik finansın temel araştırma konularından pay senedi getirilerinin tahmininde istatistikî yöntemlere alternatif olarak daha az varsayım gerektiren esneklik sağlayan ve gerçek riske dayanan genelleme kabiliyeti daha üstün olan makine öğrenmesi tekniklerinin giderek artan bir yoğunlukta kullanıldığı görülmektedir (Özdemir vd., 2011: 45). Bu çalışmada Gestel vd. (2007) çalışması takip edilerek, veri setinin kısıtlı olması nedeniyle ve mevcut verinin optimal dağılımını sağlamak için ANN, CART ve KNN analizlerinde doğrulama türü olarak 10 katlı çapraz doğrulama yöntemi kullanılmıştır. ANN, CART ve KNN analizlerinde Ambrose ve Seward (1988) çalışması takip edilerek örneklem seçimi olarak tabakalı örneklem seçimi yöntemi kullanılmıştır.

Sınıflandırma sürecinde modelin ne kadar iyi öğrendiği, performans ölçümleri yapılarak anlaşılmaktadır. Temel performans ölçütleri olarak doğruluk (accuracy), duyarlılık (sensitivity), belirleyicilik (specificity), Kappa istatistiği kullanılmaktadır. Bu performans ölçümleri 2x2 boyutunda dört farklı sonucun olabileceği durumlar için kullanılmakta ancak bazı sınıflandırma problemlerinde bu dört durum dışında farklı durumlara da bakılması ya da daha hassas tahminler yapılması gerekmektedir. 2x2'lik kontenjans tablosundan elde edilen doğruluk, belirleyicilik, duyarlılık formülleri yetersiz kalacağından bu formüllerin nxn boyutundaki bir kontenjans matrisi için güncellenmesinin makine öğrenmesi alanında her sınıflandırma problemi için daha hassas sınıflandırma elde edileceği tahmin edilmektedir (Çelik vd., 2017: 232-233). Tablo 1'de nxn boyutundaki kontenjans tablosu verilmiştir.

Tablo 1. NxN Boyutundaki Kontenjans Tablosu

		Gözlemlenen Grup (Gerçek)					
		C1	C2	C3	C4	C5	
Tahmin Edilen Grup	C1	C11	C12	C13	C14	C15	Toplam
	C2	C21	C22	C23	C24	C25	
	C3	C31	C32	C33	C34	C35	
	C4	C41	C42	C43	C44	C45	
	C5	C51	C52	C53	C54	C55	
		Toplam					T

DTP (Desired Case, Truly Predicted, Hedeflenen durumun gerçek değerinin model tarafından doğru bir şekilde tahmin edilmesi): C1 için C11, C2 için C22, C3 için C33, C4 için C44, C5 için C55'dir.

UTP (Undesired Case Truly Predicted, Hedeflenen durum dışında bir durumun gerçek değerinin model tarafından doğru bir şekilde tahmin edilmesi): C1 için C22, C33, C44, C55; C2 için C11, C33, C44, C55; C3 için C11, C22, C44, C55; C4 için C11, C22, C33, C55; C5 için C11, C22, C33, C44'tür.

FT1 (Desired Case, False Type I, Hedeflenen durumun gerçek değerinin, model tarafından hedeflenen durum dışındaki bir durum olarak tahmin edilmesi): C1 için C21+C31+C41+C51; C2 için C12+C32+C42+C52; C3 için C13+C23+ C43+C53; C4 için C14+ C24+C34+C54; C5 için C15+C25+C35+C45'dir.

FT2 (Undesired Case, Predicted as Desired Case, False, Type II, hedeflenen durum dışındaki bir durumun gerçek değerinin model tarafından hedef durum olarak yanlış biçimde tahmin edilmesi): C1 için C12+C13+C14+C15; C2 için C21+C23+C24+C25; C3 için C31+C32+C34+C35; C4 için C41+C42+C43+C45; C5 için C51+C52+C53+C54'tür.

$$\text{Doğruluk (ACC)} = \frac{C_{11}+C_{22}+C_{33}+C_{44}+C_{55}}{\sum_{i=1}^n \sum_{j=1}^n C_{ij}} = \frac{DTP+UTP}{T} \quad (1)$$

$$\text{Duyarlılık (TPR)} = \frac{DTP}{DTP+FT1} \quad (2)$$

$$\text{Belirleyicilik (specificity- TNR, SPC)} = \text{C1 için } \frac{C_{11}}{C_{11}+C_{21}+C_{31}+C_{41}+C_{51}}; \text{ C2 için } \frac{C_{22}}{C_{12}+C_{22}+C_{32}+C_{42}+C_{52}};$$

$$\text{C3 için } \frac{C_{33}}{C_{13}+C_{23}+C_{33}+C_{43}+C_{53}}; \text{ C4 için } \frac{C_{44}}{C_{14}+C_{24}+C_{34}+C_{44}+C_{54}}; \text{ C5 için } \frac{C_{55}}{C_{15}+C_{25}+C_{35}+C_{45}} \quad (3)$$

Formüllerin $n \times n$ boyutuna güncellenmesi sonucunda doğruluk; modelin hedef durumdan bağımsız bir şekilde gerçek durumları tahmin değeri olarak, duyarlılık; hedeflenen durumun yalnızca hedef durumu içeren gerçek değerler içindeki tahmin oranı olarak ifade edilebilir. Belirleyicilik ise hedef durum dışındaki her durum için ayrı ayrı hesaplanmalıdır (Çelik vd., 2017: 235-236).

5.1. Veri Seti ve Araştırmanın Kısıtları

Finansal piyasalarda pay senedi fiyatları politik olaylar, firmaların politikaları, genel ekonomik koşullar, yatırımcıların beklentileri, kurumsal yatırımcıların tercihleri, diğer borsaların hareketleri ve yatırımcıların psikolojisi gibi makroekonomik faktörlerden etkilenmektedir (Kara vd., 2011: 5311). Borsa İstanbul'da işlem gören perakende, finans şirketleri, yatırım şirketleri, leasing şirketleri, bankalar ve sigorta şirketleri finansal oranlarının belirli özellikler taşıması nedeniyle çalışmaya dâhil edilmemiştir. Veri seti 2010/2018 döneminde yıllık mali tablolarına düzenli olarak ulaşılabilen Borsa İstanbul 30/100 ve sürdürülebilirlik endeksindeki 13 imalat sanayi şirketiyle sınırlandırılmıştır. Çalışma örneklemini oluşturan şirketler Tablo 2'de verilmiştir.

Tablo 2. Çalışma Örneklemini Oluşturan Şirketler

Kodu	Unvanı
AEFES	Anadolu Efes Biracılık ve Malt Sanayii A.Ş.
ANACM	Anadolu Cam Sanayii A.Ş.
AYGAZ	Aygaz A.Ş.
BRISA	Brisa Bridgestone Sabancı Lastik Sanayi Ve Ticaret A.Ş.
CIMSA	Çimsa Çimento Sanayi ve Ticaret A.Ş.
FROTO	Ford Otomotiv Sanayi A.Ş.
KERVT	Kereviş Gıda Sanayi ve Ticaret A.Ş.
KORDS	Kordsa Teknik Tekstil A.Ş.
OTKAR	Otokar Otomotiv ve Savunma Sanayi A.Ş.
TATGD	Tat Gıda Sanayi A.Ş.
TTRAK	Türk Traktör ve Ziraat Makineleri A.Ş.
ULKER	Ülker Bisküvi Sanayi A.Ş.
VESTL	Vestel Elektronik Sanayi ve Ticaret A.Ş.

Örnek kapsamındaki şirketlerin 2010-2018 yılları arasındaki mali tablo verileri Kamuyu Aydınlatma Platformu'nun (KAP) internet sitesinden alınmıştır. Makroekonomik veriler <https://tr.investing.com/> adresinden alınmıştır. TÜFE verileri <https://kpmgvergi.com> internet sitesinden alınmıştır. Mali tablolardan elde edilen finansal oranlar yıllık alındığından aylık ilan edilen makroekonomik verilerin 12 aylık ortalaması alınmıştır. Örnek kapsamındaki 13 şirketin 2010-2018 arasında 9 yıllık mali tablolarından elde edilen 11 finansal oran tek tek hesaplanarak programa girilmiştir. 2010-2018 yılları arasındaki 4 makroekonomik değişken her şirketin ilgili yılları için aynı olup sadece yıllar itibariyle değişiklik göstermiştir. Araştırmanın gerek çalışma dönemi, gerekse örnek sayısı yönünden sınırlandırılması amacıyla Borsa İstanbul 30/100 endeksi ve sürdürülebilirlik endeksindeki imalat sanayinde faaliyet gösteren 13 şirket verileri ile analizler yürütülmüştür. Borsa İstanbul imalat sanayi sektöründe faaliyet gösteren şirket sayısı 2020 yılı itibariyle 175'tir. Çalışmanın en önemli kısıtı örnek sayısının sınırlandırılmış olmasıdır. Örneklemini oluşturan şirketler Borsa İstanbul'da işlem gören tüm şirketleri temsil etmediği gibi Borsa İstanbul imalat sanayi sektörü işletmelerini de temsil etmemektedir. Dolayısı ile çalışmada kullanılan modellerin BIST Tüm Endeksi veya BIST İmalat Sanayi Endeksindeki şirketlere genellenmesi tavsiye edilmemektedir.

Makroekonomik değişkenlerden "CDS Primi", "Dow Jones Endeksi" "TCMB Cumhuriyet Altını Satış Fiyatı" değişkenleri yüksek değerler alabildiğinden bu üç değişkenin doğal logaritması alınmıştır. Bazı değişkenlerin sistemin doğruluğu üzerinde olumsuz bir etkisi olabilir. Bu nedenle, değişken seçimi eğilim tahmini için gerekli olan değişkenleri tanımlamak ve tahmin doğruluğunu iyileştirmeye yardımcı olmak için kullanılır. Seçilmeyen değişkenler gereksiz sayılır ve veri setinden çıkarılmaktadır (Nair vd., 2010: 3). Yakut ve Gemici (2017) çalışması takip edilerek bağımsız değişkenler arasında çoklu doğrusal bağlantı varlığını ortaya çıkarmak için korelasyon analizi gerçekleştirilmiştir. Çalışmada kullanılan teknikler parametrik olmayan yöntemler olduğundan yüksek korelasyona sahip değişkenlerin bulunmasının analiz sonuçlarında önemli etki yapmayacağı bilinmektedir. Ancak yüksek korelasyona sahip değişkenlerin veri setinde bulunması analiz

sonuçlarında herhangi bir değişiklik yaratmadığından %80'in üzerinde korelasyona sahip olan değişkenler veri setinden çıkarılmıştır. Korelasyon tablosu Ek 3'te verilmiştir. Buna göre "Hisse Başına Kâr" ve "TCMB Cumhuriyet Altın Fiyatı" değişkeni veri setinden çıkarılmıştır. Bu çalışmada kullanılan değişkenler literatürde pay senedi fiyat/getiri tahmininde en çok kullanılan değişkenler arasından seçilmiştir. Tablo 3'de çalışmada kullanılan değişkenler ve bu çalışmada kullanılan değişkenleri çalışmalarında kullanan araştırmacılar bilgisi verilmiştir.

Tablo 3. Çalışmada Kullanılan Değişkenler ve Bu Çalışmada Kullanılan Değişkenleri Çalışmalarında Kullanan Araştırmacılar

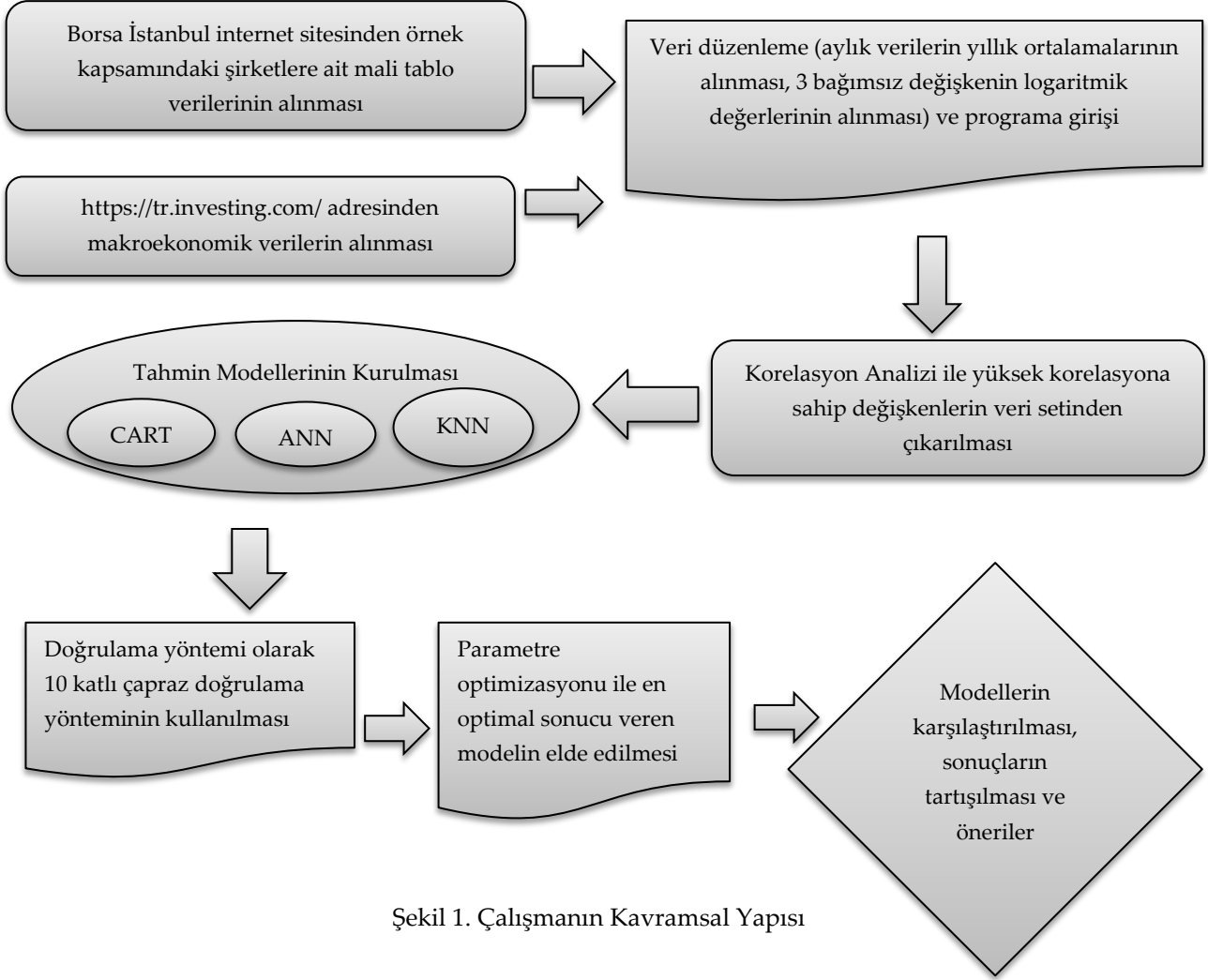
No	Değişken	Hesaplanması	Bu çalışmada kullanılan değişkenleri çalışmalarında kullanan araştırmacılar
X1	Cari Oran	Dönen Varlıklar/KVYK	Yakut ve Gemici (2017); Oliveira, Nobre ve Zarate (2013)
X2	Aktif Devir Hızı	Net Satışlar /Aktif Toplam	Yakut ve Gemici (2017)
X3	Alacak Devir Hızı	Net Satışlar/Kısa Vadeli Ticari Alacaklar	Yakut ve Gemici (2017)
X4	Özsermaye Devir Hızı	Net Satışlar/Özsermaye	Yakut ve Gemici (2017)
X5	Faaliyet Kâr Marjı	Faaliyet Karı /Net Satışlar	Yakut ve Gemici (2017)
X6	Net Kâr Marjı	Net Kar/Net Satışlar	Yakut ve Gemici (2017); Oliveira vd., (2013)
X7	Özsermaye Karlılık Oranı	Net Kar/Özsermaye	Yakut ve Gemici (2017); Oliveira vd., (2013); Akcan ve Kartal (2011)
X8	Kaldıraç Oranı	Toplam Borç/Aktif Toplam	Yakut ve Gemici (2017); Oliveira vd., (2013)
X9	Fiyat/Kazanç	(Hisse Senedi Fiyatı x ödenmiş Sermaye)/(Net Dönem Karıx1000)	Yakut ve Gemici (2017); Oliveira vd., (2013); Akcan ve Kartal (2011)
X10	Piyasa/Defter Değeri	(Hisse Senedi Fiyatı x Ödenmiş Sermaye/1000)/Özsermaye	Yakut ve Gemici (2017); Akcan ve Kartal (2011)
X11	Hisse Başına Kâr Oranı	Net Kar/Ödenmiş Sermaye	Oliveira vd., (2013); Akcan ve Kartal (2011)
X12	TÜFE (Yıllık)	Aylık verilerin yıllık aritmetik ortalaması alınmıştır.	Yakut ve Gemici (2017); Oliveira vd., (2013); Akcan ve Kartal (2011)
X13Ln	CDS Primi Aylık (Yıllık Ortalama)	Aylık verilerin yıllık aritmetik ortalaması ve değerlerin doğal logaritması alınmıştır.	Yazar tarafından eklenmiştir.
X14Ln	Dow jones Endeksi (Yıllık ortalama)	Aylık verilerin yıllık aritmetik ortalaması ve değerlerin doğal logaritması alınmıştır.	Boyacıoğlu ve Avcı (2010)
X15Ln	TCMB Cumhuriyet Altını Satış Fiyatı (Yıllık ortalama)	Aylık verilerin yıllık aritmetik ortalaması ve değerlerin doğal logaritması alınmıştır.	Akcan ve Kartal (2011)
X16	Bağımlı Değişken	(-%20den -%50 ye kadar 0,0'dan -%20 ye kadar 1, 0'dan +%20'ye kadar 2, +%20'den +%50'ye kadar 3, +%50'den fazla 4)	Yazar tarafından oluşturulmuştur.

5.2. Metodoloji

Bu çalışmada Borsa İstanbul 30/100 endeksi ve sürdürülebilirlik endeksinde faaliyet gösteren 13 imalat sanayi şirketine ait 2010-2018 yıllarına ilişkin 11 mikro ekonomik değişken 4 makroekonomik değişken kullanılarak 2011-2019 yılları ortalama yıllık pay senedi getirileri tahmin edilmiştir. Bağımlı değişken, örnek kapsamındaki şirketlerin pay senedi getirileri t dönemi fiyatına göre (-%20'den -%50 ye kadar 0, 0'dan -%20 ye kadar 1, 0'dan +%20'ye kadar 2, +%20'den +%50'ye kadar 3, +%50'den fazla ise 4 olarak kodlanmıştır) 0'dan 4'e kadar kategorik hale getirilmiştir. Çalışma örneğini oluşturan 13 şirkete ait 2010-2018 döneminde 9 yıllık verilerle

oluşturulan toplam 117 örneğin 13'ü t+1 döneminde %50'den fazla pozitif getiri sağlamış, 29 örnek +%20/+%50 arasında pozitif getiri, 41 örnek %0/+%20 arasında pozitif getiri, 24 örnek 0/-%20 arasında negatif getiri, 10 örnek ise -%20/-%50 arasında negatif getiri sağlamıştır.

Veri Madenciliği uygulamalarını gerçekleştirmek için ticari ve açık kaynak olmak üzere birçok program mevcuttur. Bu programlar arasında RapidMiner (YALE), WEKA ve R programları en çok kullanılanlar arasındadır. (Dener vd., 2009). Bu nedenle bu çalışmada ANN, CART ve KNN analizleri için RAPIDMINER 9.7 programı kullanılmıştır. Şekil 1'de çalışmanın kavramsal yapısı verilmiştir.



Şekil 1. Çalışmanın Kavramsal Yapısı

K-katlı çapraz doğrulama yöntemi uygun veri dağılımı için literatürde sıklıkla kullanılan etkili bir yöntemdir. K-katlı çapraz doğrulamada veriler rastgele k sayıda eşit miktarda parçaya ayrılır. Sırasıyla bir parça test için, kalanlar eğitim için kullanılarak analiz yapılır. Sonra başka bir parça test, diğerleri eğitim için kullanılır. Her aşamada veri madenciliği analizi yapılır ve parçaların tümü test edildikten sonra genel performans elde edilir. Yapılan deneysel çalışmalarda, uzman görüşlerine göre k sayısı için en uygun değer 10 bulunmuştur (Çelik vd., 2017: 243). Bu çalışmada kullanılan on katlı çapraz doğrulama yöntemi, aşağıdaki gibi açıklanabilir: Veri örneği rastgele on eşit boyutlu kısma bölünür ve ağ on kez eğitilir. Eğitim geçişlerinin her birinde, on kısma ayrılan eğitim verilerinden bir kat atlanır ve ortaya çıkan model, bir doğrulama seti olarak da bilinen atlanmış kattaki verilerle doğrulanır. Tahminleri güvenilir bir şekilde değerlendirmek için, yalnızca tahmin doğruluğu değil, duyarlılık ve kesinlik de göz önünde bulundurulmuştur. Tabakalı örneklem (stratified sampling) seçiminde her beş sınıftan eşit sayıda ancak rastgele örnekler alınmaktadır. Bağımlı değişkenin kategorik olduğu durumlarda başarılı sonuç vermektedir (Liang vd., 2015: 291). Bu nedenle çalışmamızda örneklem seçiminde tabakalı örneklem seçimi (stratified sampling) kullanılmıştır. Gerçek yön ve tahmini yön değişkenleri aynı değeri aldığı anda doğru tahmin gerçekleştirilmiş demektir.

5.3. Çalışmada Kullanılan Yöntemler

5.3.1. CART Analizi

Karar ağacı (Decision Tree, DT) algoritması, tüm veri öğeleri belirli bir sınıfa ait oluncaya kadar derinlik veya genişlik yöntemini kullanarak bir veri kümesini yinelemeli olarak bölümleyen bir veri madenciliği tekniğidir (Imandoust ve Bolandraftar, 2014: 108). Sınıflandırma ve regresyon ağacı (Classification and Regression Tree, CART) veri kümesinin sınıflandırılmasında kullanılan karar ağacı tekniğidir ve hangi kayıtların belirlenmiş bir sonuca dâhil olacağını tahmin etmek için yeni (sınıflandırılmamış) bir veri kümesine uygulanabilecek bir dizi kural sunar. CART ikili bölmeler oluşturarak veri kümesini sınıflara ayırmaktadır (Dhanalakshmi ve Subramanian, 2014: 330). Hangi düğümün kök düğüm olacağına karar vermenin dışında düğümün hangi noktadan ikiye ayrılması gerektiğini de hesaplar. CART, dallara ayırma kriterini hesaplarken kayıp verileri dikkate almaz. Hesaplanan $\Psi(s/t)$ değerleri içinden en büyük değere sahip nokta, düğüm olarak seçilir ve işlem tüm yapraklara kadar aynı şekilde devam ettirilir (Silahtaroglu, 2016: 83).

$$\Psi(s/t=2P_L P_R \sum_{j=1}^M |P(C_j|t_L) - P(C_j|t_R)| \quad (4)$$

t: Dalların yapılacağı düğüm

c: Kriter

L: Ağacın sol tarafı

R: Ağacın sağ tarafı

P_L, P_R : Öğrenim kümesindeki bir kaydın sağda veya solda olma olasılığı

$|P(C_j|t_L) - P(C_j|t_R)|$: C sınıfındaki bir kaydın sağda veya solda olma olasılığı

Gini ölçütü bir frekans dağılımında değerler arasındaki eşitsizliğin ölçüsüdür. Bu ölçütte öznitelik değerlerinin sol ve sağda olmak üzere iki bölüme ayrılması, her bir bölüm için ayrı ayrı Gini ölçütünün hesaplanması ve elde edilen sonuçların karşılaştırılmasına dayanmaktadır. (Özkan, 2016). Gini Endeksi $g(t)$, pi her bir kategorinin olasılığı ve c kategori sayısı olmak üzere 5 nolu eşitlik kullanılarak hesaplanır (Akpınar, 2014).

$$g(t) = 1 - \sum_{i=1}^c p_i^2 \quad (5)$$

5.3.2. ANN Analizi

Yapay sinir ağı (Artificial Neural Networks, ANN) tekniği, doğrusal olmayan değişkenler arasındaki ilişkiyi öğrenme ve algılama becerisi nedeniyle iş alanında giderek daha fazla kabul gören veri madenciliği tekniklerinden biridir (Adebiyi, 2012: 1). ANN'de işlem elemanı ağırlık değerlerinin belirlenmesinde (ağın eğitilmesi) ağırlıklar rastgele atanmaktadır. ANN'ler kendilerine örnekler gösterildikçe bu ağırlık değerleri değişmektedir. Örnekler ağa defalarca gösterilerek en doğru ağırlık değerleri bulunmaya çalışılır ve doğru ağırlık değerine ulaşıldığında olay hakkında genelleme yapılırsa ağın öğrendiği belirtilebilir. Ağın eğitimden sonra öğrenip öğrenmediğini test etme işlemi eğitim sırasındaki bağlantı ağırlıkları değiştirilmeden önce ağın görmediği örnekler için çıktı üretmesi ile olur. Test çıktı değeri ne kadar iyi olursa eğitimin performansı da o kadar iyi demektir (Öztemel, 2012: 55).

Gizli katmandaki nöron sayısı (n), öğrenme hızı değeri (lr), momentum sabiti (mc) ve yineleme sayısı (ep) etkin bir şekilde belirlenmesi gereken ANN modeli parametreleridir (Kara vd. 2011: 5314). Nöron, önceki aşamada (n) bir ağırlıkla çarpılan ve daha sonra bir eşik b ile toplanan birkaç düğümünden oluşan bir giriş olan x_i 'den oluşur. Transfer fonksiyonu, bir nöron çıkışını belirleyen matematiksel bir fonksiyonla hesaplanır (Gaganis, 2009; 213):

$$in = \sum_{i=1}^n w_i \cdot x_i + b \quad (6)$$

İleri sürümlü ANN'da kullanılan ağırlıklar her seferinde Δw kadar düzeltilerek yenilenir.

$$W_1^{yeni} = w_1^{eski} + \Delta w_1 \quad (7)$$

Algoritmanın en hassas noktası Δw değerlerini bularak en uygun w ağırlıklarını elde etmektir. Bunun için her seferinde oluşan hatayı minimuma indirecek bir yapı kullanılır. Gerçekte var olan değer g ile; w ağırlıklarıyla

elde edilen değer de y ile gösterilirse en küçük kareler yöntemiyle elde edilecek hata fonksiyonu E şu şekilde hesaplanabilir (Silahtaroglu, 2016: 124-125):

$$E_r = 1/2 e^2 = 1/2 (g - y)^2 \quad (8)$$

Her bağlantı, bir aktivasyon fonksiyonu, çoğunlukla bir lojistik fonksiyon veya hiperbolik teğet girişlerinin ağırlıklı toplamı kullanılarak, iki nöron arasındaki ve her bir nöron arasındaki ilişkinin gücünü gösteren bir ağırlık ile temsil edilir. Sinir ağı kullanılarak tasarlanan bir pay senedi getiri tespit modeli, belirli bir firma için, pay senedi getiri ihtimalini temsil eden bir gizli katman, bir çıktı nöronu ve bir girdi katmanından oluşan ağ ile aşağıdaki şekilde ifade edilebilen bir Z skoru hesaplamaktadır (Öztemel, 2012: 55).

$$Z = f(\sum_{i=1}^n w_{ij}x_i + b_j) \cdot (\sum_{j=1}^p w_j) + b \quad (9)$$

Formülde f aktivasyon fonksiyonu, n değişken sayısı, p gizli nöron sayısı, x_i girdi katmanı nöronları, w_{ij} girdi katmanı ve gizli katman arasındaki ilişkileri temsil eden ağırlıkları, w_j gizli katman ve çıktı katmanı grupları arasındaki ağırlıkları, b_j gizli nöronların ağırlıkları ve b çıkış nöronunun ağırlığını göstermektedir (Jardin, 2016: 241).

5.3.3. KNN Analizi

Bellek tabanlı yöntemler arasında sayılan K-En Yakın Komşu Algoritması (KNN, K Nearest Neighbor Algorithm) gözlem değerlerinin arasındaki uzaklıklardan yararlanarak sınıflandırma işlemi yapmaktadır. Bu algoritma denetimli bir öğrenme yöntemi olup verilerin sınıflandırmasında seçilen bir özelliğin kendine en yakın özellikler arasındaki yakınlığı kullanmaktadır. KNN yöntemi hangi sınıfa girdiği bilinmekte olan bir örnek kümesindeki gözlem değerlerinden yararlanarak yeni katılacak olan bir gözlemin sınıflardan hangisine gireceğini belirlemek amacıyla kullanılmaktadır. K-En Yakın Komşu Algoritması örnek kümede bulunan gözlemlerin sonradan belirlenen bir gözlem değerine olan uzaklıkları hesaplandıktan sonra en yakın k sayıda gözlemin seçilmesine dayalı bir yöntemdir. Uzaklıkların hesaplanmasında i ve j noktaları için Öklid uzaklık formülü kullanılmaktadır (Özkan, 2016):

$$d(i, j) = \sqrt{\sum_{k=1}^p (X_{ik} - X_{jk})^2} \quad (10)$$

5.4. Bulgular

5.4.1. CART Analizi Bulguları

Tablo 4'de verilen parametre aralıklarında en yüksek sınıflandırma ve tahmin başarısı gösteren modelin belirlenmesi amacıyla parametre optimizasyonu gerçekleştirilmiştir.

Tablo 4. CART Sınıflandırma Algoritması Analiz Parametreleri

Analiz İçin Belirlenen Parametreler	Açıklama			
Doğrulama Türü	10 Katlı Çapraz Doğrulama			
Değişken Sayısı	13 Bağımsız Değişken			
Örnekleme Seçimi	Tabakalı Örnekleme Seçimi			
Bölünme Kriteri	Gini Index			
Analiz İçin Belirlenen Parametreler	En Düşük	En Yüksek	Adımlar	Ölçek
Bölünme İçin Minimum Boyut (Minimal Size For Split)	1.0	4.0	10	Doğrusal
En Düşük Dal Boyutu (Minimal Leaf Size)	1.0	2.0	10	Doğrusal
Minimum Kazanım (Minimal Gain)	1.0	10.0	10	Doğrusal
Maksimum Derinlik (Maximal Depth)	1.0	10.0	10	Doğrusal
Güven Düzeyi (confidence)	0.0	0.25	-	-
Ön Budama Sayısı	0.0	10	-	-

CART Analizi Karışıklık Matrisi Tablo 5'de verilmiştir. Gözlem grubunda (gerçek durumda) +%50'den fazla getiri elde eden (4 ile kodlanmıştır) 13 örneğin 11'i doğru tahmin edilmiş, bir örnek gerçek durumda +%50'den fazla getiri elde ettiği halde 0/-%20 getiri (2 ile kodlanmıştır) ve bir örnek -%20/-%50 getiri (0 ile kodlanmıştır) olarak yanlış sınıflandırılmıştır. Gözlem grubunda (gerçek durumda) +%20/+%50 getiri elde eden (3 ile kodlanmıştır) 29 örneğin 28'i doğru tahmin edilmiş, bir örnek gerçek durumda +%20/+%50 getiri elde ettiği halde 0/-%20 getiri (1 ile kodlanmıştır) elde edilmiş şeklinde yanlış sınıflandırma ve tahmin gerçekleştirmiştir. Gözlem grubunda 0'dan +%20'ye kadar getiri (2 ile kodlanmıştır) elde eden 41 örneğin 39'u doğru tahmin

edilmiş, bir örnek 0'dan +%20'ye kadar getiri elde ettiği halde 0/-%20 (1 ile kodlanmıştır) getiri olarak ve bir örnek -%20/-%50 (0 ile kodlanmıştır) getiri şeklinde yanlış tahmin edilmiştir. Gözlem grubunda 0'dan -%20 ye kadar (1 ile kodlanmıştır) getiri sağlayan 24 örneğin 23'ü doğru tahmin edilmiş, bir örnek gerçek durumda 0/-%20 getiri sağladığı halde -%20/-%50 (0 ile kodlanmıştır) getiri şeklinde yanlış tahmin edilmiştir. Gözlem grubunda -%20/-%50 getiri sağlayan 10 örneğin tamamı doğru tahmin edilmiştir. Karışıklık matrisinin okunması açısından yukarıda detaylı verilen açıklamaya çalışma kapsamını genişletmemek için ANN ve KNN analiz bulgularında yer verilmemiştir.

Tablo 5. CART Analizi Karışıklık Matrisi

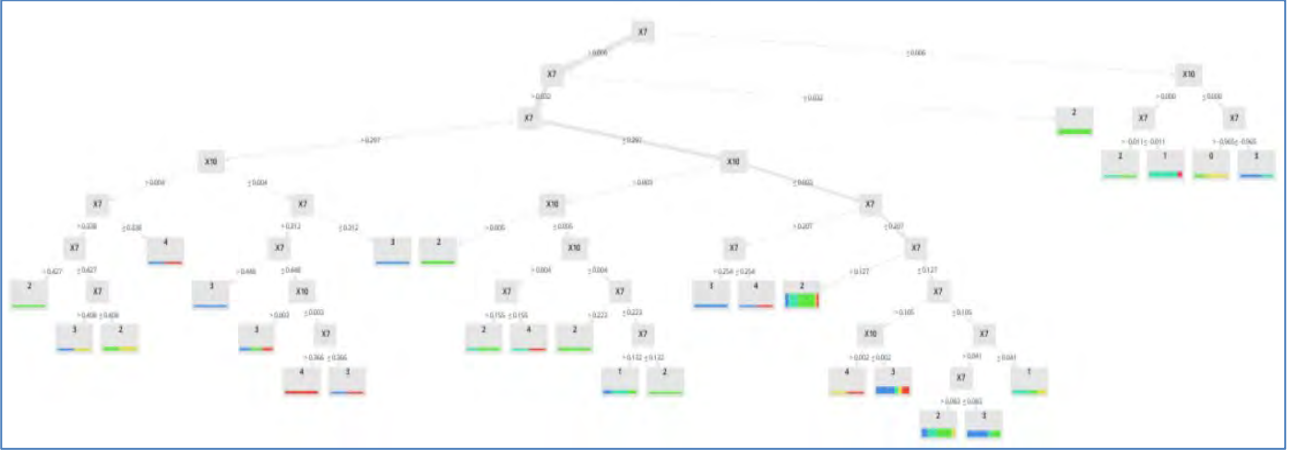
		Gözlemlenen Grup (Gerçek)					Toplam
		4	3	2	1	0	
Tahmin Edilen Grup	4	11	0	0	0	0	11
	3	0	28	0	0	0	28
	2	1	0	39	0	0	40
	1	0	1	1	23	0	25
	0	1	0	1	1	10	13
Toplam		13	29	41	24	10	117

Tablo 6'da parametre optimizasyonu sonucu en iyi performansı gösteren CART analizi performans sonuçları verilmektedir. Tablo 6'da yer alan Kappa istatistiğinden kısaca bahsetmek gerekmektedir. Kappa testi, iki veya daha fazla gözlemci arasındaki uyumun güvenilirliğini ölçen bir istatistik yöntemidir. Uyumun değerlendirildiği değişken kategorik (nominal) olduğu için uygulanan istatistik parametrik olmayan istatistik türüdür. "Cohen'in kappa katsayısı" sadece iki gözlemci arasındaki uyumu ele alırken, uyumun ölçüldüğü gözlemci sayısı ikiden fazla ise "Fleiss'in kappa katsayısı" kullanılır. Fleiss tarafından yapılan sınıflamada, Kappa değerinin 0.75 ve üzeri olması mükemmel, 0.40-0.75 arası orta-iyi, buna karşılık 0.40'ın altında bulunması zayıf uyum olarak değerlendirilmektedir (Kılıç, 2015: 142). Bu çalışmada elde edilen CART modeline ait Kappa değeri 0,932 olduğundan iki gözlemcinin mükemmel uyumlu olduğu görülmektedir.

Tablo 6. Parametre Optimizasyonu Sonucu En İyi Performansı Gösteren CART Sınıflandırma ve Regresyon Ağacı Analizi Performans Sonuçları

Parametreler	Sonuçlar
Doğruluk (accuracy)	%94,87
Sınıflandırma Hatası (classification error)	%5,13
Kappa	0,932
Ağırlıklı Ortalama Kesinlik (weight mean precision)	%93,28
Ağırlıklı Ortalama Duyarlılık (weight mean recall)	%94,42
Bölünme İçin Minimum Boyut (Minimal Size for Split)	2
En Düşük Dal Boyutu (Minimal Leaf Size)	1
En düşük Kazanım (Minimal Gain)	3,699
En Yüksek Derinlik	10

Şekil 2'de CART karar ağacı görüntüsü verilmiştir. "X7, Özsermaye Kârlılık Oranı" ağacın kökünü oluşturan en önemli değişken olarak bulunmuştur. X7, Net Kâr Marjı" 0,006'dan küçük olan örnekler için ağacın karar vermesinde ikinci önemli değişken "X10, Piyasa/Defter Değeri" değişkenidir. Çalışma hacmini sınırlamak için tüm karar kuralları verilmemiş, örnek olması açısından karar kuralının okunması yukarıda kısaca ele alınmıştır. Ek 2'de CART Analizi Karar Kuralı verilmiştir.



Şekil 2. CART Karar Ağacı Görüntüsü

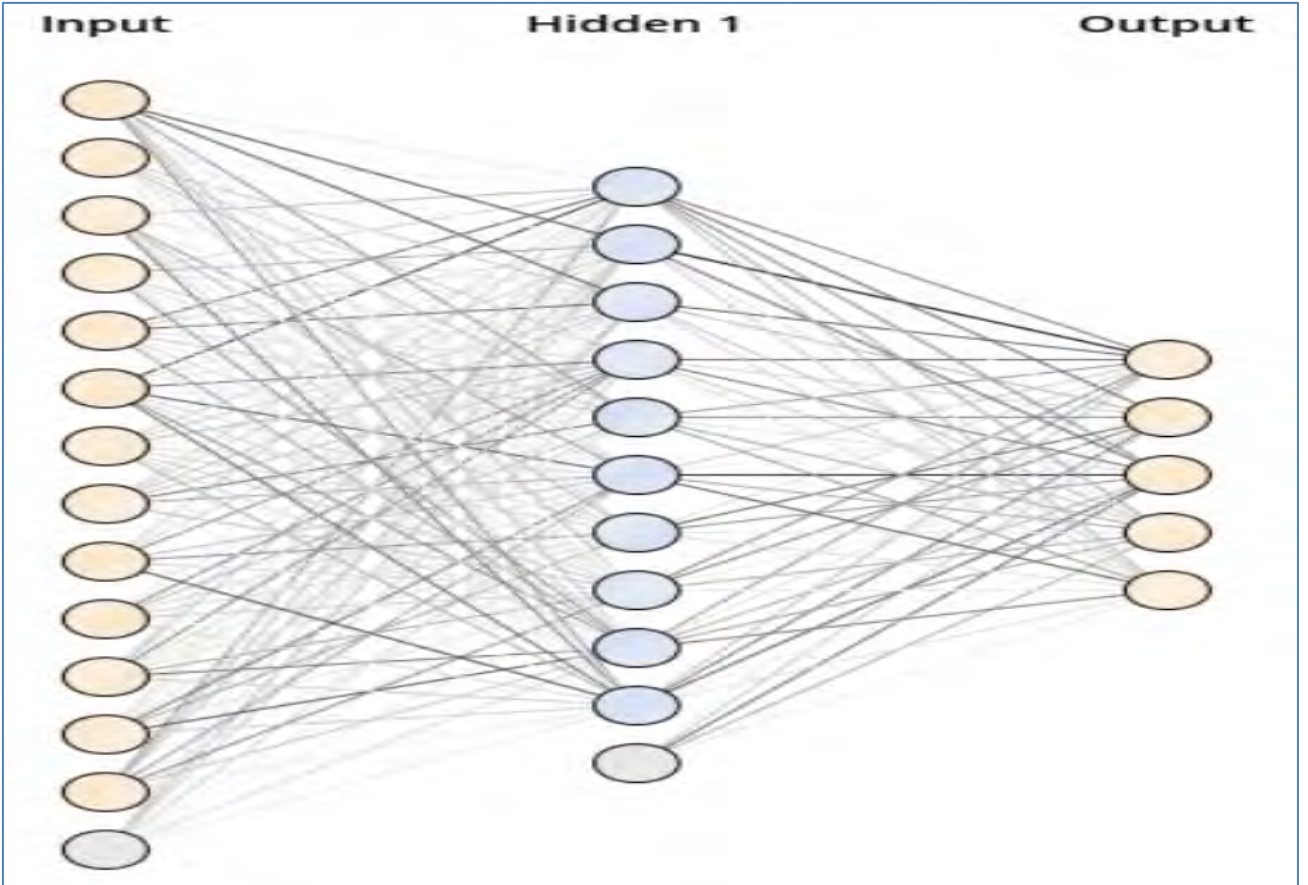
5.4.2. ANN Analizi Bulguları

Çeşitli gizli katman nöronları, öğrenme algoritmaları ve öğrenme oranları ile yapılan birçok deneyden sonra, giriş katmanında 13 nöron, gizli katmanda 10 nöron, 0,27 öğrenme oranı, momentum 0,18, devir sayısı 800 ve esnek bir geri yayılım algoritması kullanan ileri beslemeli sinir ağı, on katlı çapraz doğrulama yöntemi ile en iyi ağ mimarisi olarak bulunmuştur. Şekil 3’de ANN mimarisi verilmiştir.

Girdi Katmanı

Gizli Katman

Çıktı Katmanı



Şekil 3. ANN Mimarisi

Tablo 7’de verilen parametre aralıklarında en yüksek sınıflandırma ve tahmin başarısı gösteren modelin belirlenmesi amacıyla parametre optimizasyonu gerçekleştirilmiştir.

Tablo 7. ANN Analiz Parametreleri

Ağ Türü	Çok Katmanlı Perseptron			
Öğrenme Algoritması	Geri Yayılım			
Öğrenme Kuralı	Momentum			
Girdi Katmanındaki Düğüm Sayısı	13 Bağımsız Değişken			
Gizli Katman Sayısı	1			
Çıktı Katman Düğüm Sayısı	5 (Getiri -%20/-%50 arası "0", 0/-%20 arası "1", 0/+%20 arası "2", +%20/+%50 arası "3", %50'den fazla "4" ile kodlanmıştır.			
Doğrulama Türü	10 Katlı Çapraz Doğrulama			
Örnekleme Seçim Türü	Tabakalı Örnekleme Seçimi (Stratified Sampling)			
Aktivasyon Fonksiyonu	Sigmoid			
Öğrenme Oranı	En Düşük: 0,00	En Yüksek: 0,30	Adımlar: 10	Ölçek: Doğrusal
Momentum	En Düşük: 0,00	En Yüksek: 0,20	Adımlar: 10	Ölçek: Doğrusal
Eğitim Devir Sayısı	En Düşük: 1,00	En Yüksek: 1000	Adımlar: 10	Ölçek: Doğrusal

ANN analizi karışıklık matrisi Tablo 8'de verilmiştir.

Tablo 8. ANN Analizi Karışıklık Matrisi

		Gözlemlenen Grup (Gerçek)					Toplam
		4	3	2	1	0	
Tahmin Edilen Grup	4	12	0	0	0	0	12
	3	0	28	0	0	1	29
	2	0	1	40	2	1	44
	1	1	0	1	22	0	24
	0	0	0	0	0	8	8
Toplam		13	29	41	24	10	117

Ek 1'de ANN ağırlıkları verilmiştir. Tablo 9'da parametre optimizasyonu sonucu en iyi performans gösteren ANN model sonucu verilmektedir.

Tablo 9. Parametre Optimizasyonu Sonucu En İyi Performansı Gösteren Yapay Sinir Ağları Model Sonucu

Parametreler	Sonuçlar
Doğruluk (accuracy)	%94,02
Sınıflandırma Hatası (classification error)	%5,98
Kappa	0,920
Ağırlıklı Ortalama Kesinlik (weight mean precision)	%95,83
Ağırlıklı Ortalama Duyarlılık (weight mean recall)	%91,62
Öğrenme Oranı (learning rate)	0,27
Momentum	0,18
Devir Sayısı (training cycles)	800

5.4.3.KNN Analizi Bulguları

KNN analizinde en yüksek sınıflandırma sonucu k=3 alındığında elde edildiğinden k değeri 3 olarak alınmıştır. Tablo 10'da verilen parametre aralıklarında en yüksek sınıflandırma ve tahmin başarısı gösteren model belirlenmiştir.

Tablo 10. KNN Analiz Parametreleri

Değişken Seçimi	13 Değişkenli Veri Seti
Doğrulama Türü	10 Katlı Çapraz Doğrulama
Örnekleme Seçimi	Tabakalı Örnekleme Seçimi (Stratified sampling)
k sayısı	1-10
Ölçüm Tipi (Measure types)	Karışık ölçü (Mixed Measures)
Karışık Ölçü (Mixed Measure)	Karışık Öklid Mesafesi (Mixed Euclidean Distance)

Tablo 11'de KNN analizi karışıklık matrisi verilmiştir.

Tablo 11. KNN Analizi Karışıklık Matrisi

		Gözlemlenen Grup (Gerçek)					Toplam
		4	3	2	1	0	
Tahmin Edilen Grup	4	10	0	0	0	0	10
	3	2	29	4	1	0	36
	2	1	0	36	0	0	37
	1	0	0	1	23	0	24
	0	0	0	0	0	10	10
Toplam		13	29	41	24	10	117

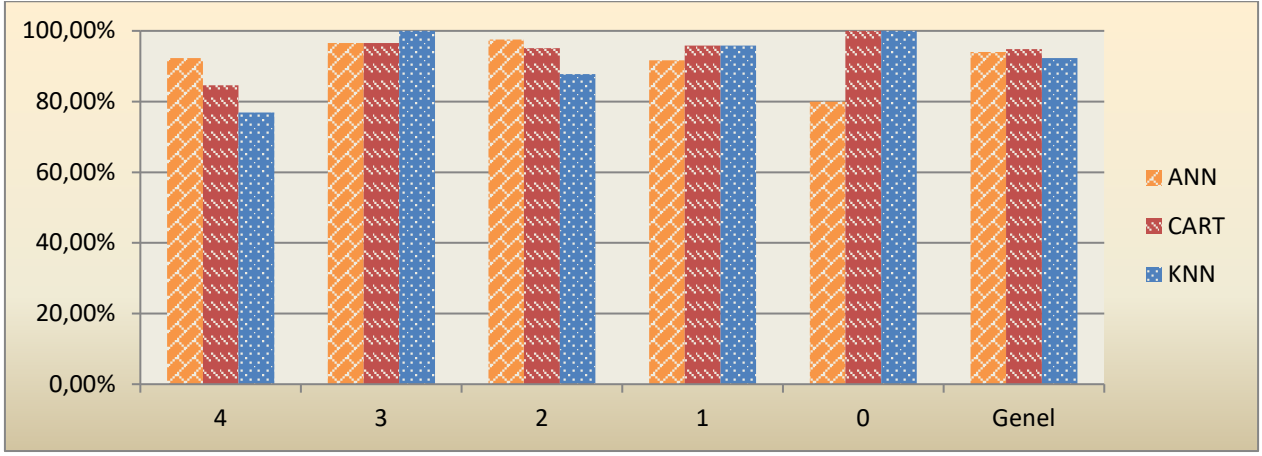
Tablo 12’de KNN analizi model sonucu verilmektedir.

Tablo 12. KNN Analizi Model Sonucu

Parametreler	Sonuçlar
Doğruluk (accuracy)	%92,31
Sınıflandırma Hatası (classification error)	%7,69
Kappa	0,898
Ağırlıklı Ortalama Kesinlik (weight mean precision)	%94,74
Ağırlıklı Ortalama Duyarlılık (weight mean recall)	%92,11
k Değeri	3

5.4.4. Bulguların Değerlendirilmesi

Şekil 4’de çalışmada kullanılan yöntemlerin tahmin performansının karşılaştırılması verilmiştir. Örnek kapsamındaki şirketlerin t+1 döneminde pay senedi t dönemine göre %50’den fazla pozitif getiri elde eden örnekleri (4 ile kodlanmıştır) ANN (%92,31), CART (%84,62), KNN (%76,92) doğruluk oranları ile tahmin etmiştir. %20’den %50’ye kadar pozitif getiri elde eden örnekleri (3 ile kodlanmıştır) ANN (%96,55), CART (%96,55), KNN (%100) oranında doğru tahmin etmiştir. 0’dan %20’ye kadar pozitif getiri elde eden örnekleri (2 ile kodlanmıştır) ANN (%97,56), CART (%95,12), KNN (%87,80) oranında doğru tahmin etmiştir. 0’dan %20 ye kadar negatif getiri sağlayan örnekleri (1 ile kodlanmıştır) ANN (%91,67), CART (%95,83), KNN (%95,83) oranında doğru tahmin etmiştir. %20’den %50 ye kadar negatif getiri sağlayan örnekleri (0 ile kodlanmıştır) ANN (%80,00), CART (%100), KNN (%100) oranında doğru tahmin etmiştir. ANN (%94,02), CART (%94,87) ve KNN (%92,31) oranında genel tahmin doğruluğu elde etmiştir. %50’den fazla pozitif getiri tahmininde ANN (%92,31), %20’den %50’ye kadar pozitif getiri tahmininde KNN (%100) başta olmak üzere ANN (%96,55), CART (%96,55), 0’dan %20’ye kadar pozitif getiri tahmininde ANN (%97,56), 0’dan %20 ye kadar negatif getiri tahmininde CART (%95,83), KNN (%95,83), %20’den %50 ye kadar negatif getiri tahmininde KNN (%100,00), CART (%100,00) sınıflandırma ve tahmin performansları dikkate değer bulunmuştur. Genel tahmin doğruluğunda bu çalışmada kullanılan tüm yöntemlerin %90,00’in üzerinde başarı ile sınıflandırma ve tahmin yaptığı görülmektedir. Gaganis (2009) çalışmasında model doğrulama türü olarak 10 katlı çapraz doğrulamanın tespit doğruluğunu artırmak için en iyi yöntemlerden biri olduğu ve %75,00’in üzerinde tespit doğruluğunun sosyal bilimler alanında iyi bir sonuç olduğunu belirtmiştir. Gaganis (2009) sınıflandırmasına göre bu çalışmadaki ANN, CART ve KNN analiz sonuçları çok iyi olarak değerlendirilebilir.

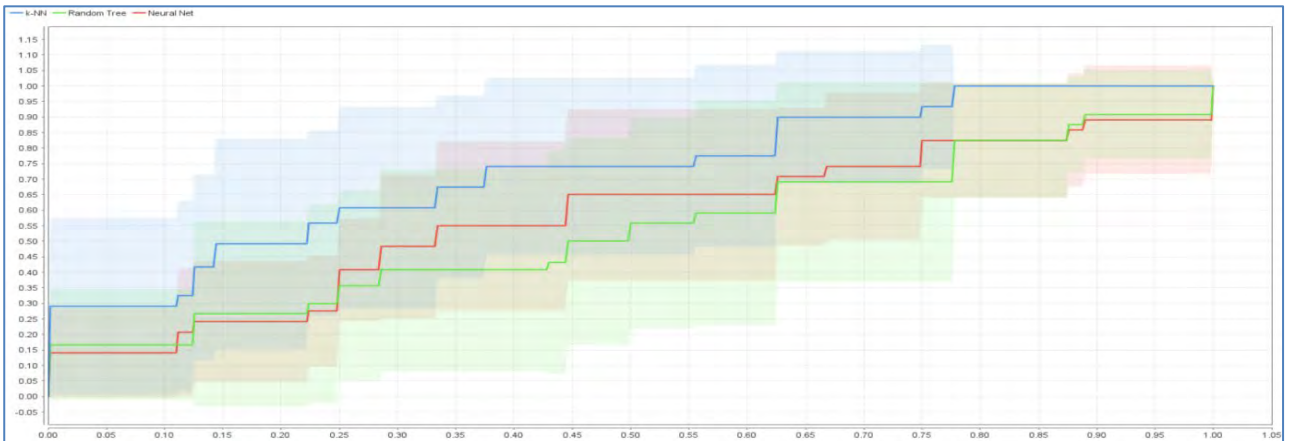


Şekil 4. Çalışmada Kullanılan Yöntemlerin Tahmin Performansının Karşılaştırılması

Patel, Shah, Thakkar, & Kotecha (2015) çalışmalarında pay senedi ve pay senedi fiyat endeksi hareketini tahmin etmek için ANN, SVM, Random Forest ve Naive-Bayes algoritmalarının tahmin performansını karşılaştırmıştır. Analiz sonucunda ANN %86,69, SVM %89,33, Random Forest %89,98 ve Naive Bayes %90,19'luk doğruluğa sahip olduğu belirtilmiştir. Bu sonuç bu çalışmada elde edilen ve bir karar ağacı algoritması olan CART'ın ANN'den daha yüksek sınıflandırma ve tahmin doğruluğu etmesi yönünden Patel vd. (2015) çalışmalarının bulgularıyla uyumludur.

Yakut ve Gemici (2017) çalışmasında LR (Logistic Regression) analizi %75,00, C5.0 algoritması %88,00, CART algoritması %89,80 ve SVM analizi %75,90'lık doğru sınıflandırma başarısı gerçekleştirmiştir. Yakut ve Gemici (2017) çalışmasında CART ilgili çalışmada en yüksek doğrulukla tahmin gerçekleştiren model olarak görülmektedir. Bu çalışmada da Yakut ve Gemici (2017) çalışma bulgusuyla uyumlu olarak CART karar ağacı ANN ve KNN yöntemleri ile kıyaslandığında en iyi tahmin modeli olarak bulunmuştur. Ayrıca ilgili çalışmada pozitif ve negatif pay senedi getiri sınıflandırma tahminine etki eden en önemli değişkenlerin "Piyasa/Defter Değeri Değişkeni", "TÜFE Değişkeni" ve "Brüt Kâr Marjı Değişkeni" olduğu belirtilmiştir. Bu çalışmada CART karar ağacının oluşturulmasında verileri bölen önemli değişkenler olarak "Özsermaye Kârlılık Oranı" ve "Piyasa Değeri /Defter Değeri" bulunmuştur.

Sınıflandırma performansının kapsamlı olarak değerlendirmesi ROC tarafından gerçekleştirilir. ROC eğrisi, modelin dikey ekseninde "isabetleri" (yani gerçek pozitifler) yüzdesini ve yatay ekseninde "yanlış alarmların" 1-özgüllüğü veya yüzde oranlarını çizer. Sonuç, 45° çizgisinden sol üst köşeye yükselen eğimli bir eğridir. Bükme keskinliği sol üst köşeye ne kadar yakınsa, modelin doğruluğu da o kadar yüksek olur. Eğri altındaki alan (AUC), çeşitli kesme noktalarının tüm olası seçimlerine göre yanlış sınıflandırma oranlarının ortalaması olarak kabul edilebilir (Gaganis, 2009: 222).



Şekil 5. ROC Eğrisi

6. Sonuç

Bu çalışmada pay senetleri Borsa İstanbul 30/100 ve sürdürülebilirlik endeksindeki imalat sanayi şirketlerinin yılsonu mali tablo verileri ve ekonomik göstergeler alınarak örnek kapsamındaki şirketlerin bir yıl sonraki ortalama pay senedi getirileri tahmin edilmiştir. Örnek kapsamındaki 13 şirketin 2010-2018 döneminde yılsonu mali tabloları ve yıllık ortalama ekonomik göstergeler alınarak 2011-2019 yıllık ortalama pay senedi getirileri Yapay sinir ağları (ANN), Sınıflandırma ve Regresyon Ağaçları (CART) ve K En Yakın Komşu Algoritması (KNN) yöntemleri ile tahmin edilmiştir. Çalışmada optimal veri dağılımı için 10 katlı çapraz doğrulama yöntemi ve verilen parametre aralıklarında en yüksek sınıflandırma ve tahmin başarısı gösteren modelin belirlenmesi amacıyla parametre optimizasyonu kullanılmıştır. Analiz sonucunda CART (%94,87), ANN (%94,02) ve KNN (%92,31) genel tahmin ve sınıflandırma doğruluğu elde etmiştir.

Örnek kapsamındaki şirketlerin t+1 döneminde pay senedi t dönemine göre %50'den fazla pozitif getiri elde eden örnekleri (4 ile kodlanmıştır) ANN (%92,31), CART (%84,62), KNN (%76,92) doğruluk oranları ile tahmin etmiştir. %20'den %50'ye kadar pozitif getiri elde eden örnekleri (3 ile kodlanmıştır) ANN (%96,55), CART (96,55), KNN (%100) oranında doğru tahmin etmiştir. 0'dan %20'ye kadar pozitif getiri elde örnekleri (2 ile kodlanmıştır) ANN (%97,56), CART (%95,12), KNN (%87,80) oranında doğru tahmin etmiştir. 0'dan %20 ye kadar negatif getiri sağlayan örnekleri (1 ile kodlanmıştır) ANN (%91,67), CART (%95,83), KNN (%95,83) oranında doğru tahmin etmiştir. %20'den %50 ye kadar negatif getiri sağlayan örnekleri (0 ile kodlanmıştır) ANN (%80,00), CART (%100,00), KNN (%100,00) oranında doğru tahmin etmiştir. ANN (%94,02), CART (94,87) ve KNN (%92,31) oranında genel tahmin doruluğu elde etmiştir. %20'den %50 ye kadar negatif getiri sağlayan örnekleri ANN (%80,00), CART (%100), KNN (%100) oranında doğru tahmin etmesi ve %20'den %50'ye kadar pozitif getiri tahmininde KNN (%100), ANN (%96,55), CART (96,55) tahmin doğruluğu elde etmesi dikkate değer bulunmuştur. Gaganis (2009) çalışmasında %75,00'in üzerinde tespit doğruluğunun sosyal bilimler alanında iyi bir sonuç olduğunu belirtmiştir. Gaganis (2009) sınıflandırmasına göre bu çalışmadaki ANN, CART ve KNN analiz sonuçları çok iyi olarak değerlendirilebilir.

Gelecekte BIST sürdürülebilirlik endeksindeki tüm şirketlerin endekste yer almayan şirketlerle t+1 döneminde pay senedi getirilerinin tahmin yöntemlerinin performansının karşılaştırılması üzerine çalışma yapılması ve sonuçların paylaşılmasının literatüre katkısının olacağı düşünülmektedir. Ayrıca daha kısa dönemlere ait mali tablo, makroekonomik veriler ve teknik göstergelerden oluşan değişkenlerin tahmin ve sınıflandırmadaki etkinlikleri incelenebilir. Yerli ve yurt dışında faaliyet gösteren ve sürdürülebilirlik endeksinde bulunan şirketlerin pay senedi getirileri hangi değişkenler ve tahmin yöntemleri kullanılarak en yüksek sınıflandırma doğruluğu elde edeceğine yönelik bir çalışma da literatüre değer katacaktır. Diğer taraftan sosyal sorumlu yatırım fonlarının, geleneksel yatırım fonlarından performans yönünden ayrışıp ayrışmadığı ile ilgili yapılacak çalışmalar literatüre değerli katkılar sunacaktır.

Kaynakça

- Adebiyi, A. A., Ayo, C. K., Adebiyi, M. O., & Otokiti, S. O. (2012). Stock Price Prediction using Neural Network with Hybridized Market Indicators. *Journal of Emerging Trends in Computing and Information Sciences*, 3 (1), 3 1-9.
- Akcan, A., & Kartal, C. (2011). İMKB Sigorta Endeksini Oluşturan Şirketlerin Hisse Senedi Fiyatlarının Yapay Sinir Ağları ile Tahmini. *Muhasebe ve Finansman Dergisi*, Temmuz, 27-40.
- Akpınar, H. (2014). *Data Veri Madenciliği Veri Analizi*. İstanbul: Papatya Yayınları.
- Altınay, A., Kaki, B., Kestane, A., Soba, M., Dinçer, Ö., & Şık, E. (2017). Sürdürülebilirlik Endeksinin Bankacılık Sektörü Hisse Senedi Değerlerine Etkileri, BİST Sürdürülebilirlik Endeksi Üzerine Bir İnceleme. *Sosyal Ekonomik Araştırmalar Dergisi*, 17 (34), 208-229.
- Anbalagan, T., & Maheswari, S. U. (2015). Classification and prediction of stock market index based on fuzzy metagraph. *Procedia Computer Science*, (47), 214-221.
- Boyacioglu, M. A., & Avci, D. (2010). An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the Prediction of Stock Market Return: The Case of the Istanbul Stock Exchange. *Expert Systems with Applications*, (37), 7908-7912.

- Charlo, M. J., Moya, I., & Muñoz, A. M. (2015). Sustainable Development and Corporate Financial Performance: A Study Based on the FTSE4Good IBEX Index. *Business Strategy and the Environment*, (24), 277.
- Çelik, U., Akçetin, E. ve Gök, M. (2017). *Rapidminer ile Veri Madenciliği*. İstanbul: Pusula Yayınları.
- Dener, M., Dörterler, M. ve Orman, A. (2009). Açık Kaynak Kodlu Veri Madenciliği Programları: WEKA'da Örnek Uygulama. Conference: XI. Akademik Bilişim Konferansı (s. 1-11). Şanlıurfa: Conference: XI. Akademik Bilişim Konferansı.
- Dhanalakshmi, S. & Subramanian, C. (2014). An Analysis of Data Mining Applications for Fraud Detection in Securities Market. *International Journal of Data Mining Techniques and Applications*, (3), 326-335.
- Enke, D., & Thawornwong, S. (2005). The use of Data Mining and Neural Networks for Forecasting Stock Market Returns. *Expert Systems with Applications*, (29), 927-940.
- Gaganis, C. (2009). Classification Techniques for the Identification of Falsified Financial Statements: A Comparative Analysis. *Intelligent Systems in Accounting, Finance and Management*, (16), 207-229.
- Gestel, T. V., Martens, D., Baesens, B., Feremans, D., Huysmans, J. Ve Vanthienen, J. (2007). Forecasting and Analyzing Insurance Companies Ratings, *International Journal of Forecasting*, (23): 513-529.
- Gök, İ. Y., & Özdemir, O. (2017). Borsa İstanbul Sürdürülebilirlik Endeksinin Performans Karakteristiği. *Sosyoekonomi*, 25 (34), 87-105.
- Huang, W., Nakamori, Y., & Wang, S.-Y. (2005). Forecasting Stock Market Movement Direction with Support Vector Machine. *Computers & Operations Research*, (32), 2513-2522.
- Imandoust, S. B., & Bolandraftar, M. (2014). Forecasting the Direction of Stock Market Index Movement using Three Data Mining Techniques: The Case of Tehran Stock Exchange. *Int. Journal of Engineering Research and Applications*, 106-117.
- Jardin, P. d. (2016). A Two-Stage Classification Technique for Bankruptcy Prediction. *European Journal of Operational Research*, (254), 236-252.
- Kara, Y., Boyacioglu, M. A., & Baykan, Ö. K. (2011). Predicting Direction of Stock Price Index Movement using Artificial Neural Networks and Support Vector Machines: The Sample of The Istanbul Stock Exchange. *Expert Systems with Applications*, (38), 5311-5319.
- Kılıç, S., (2015). Kappa Testi, *Journal of Mood Disorders*, 5(3), 142-144.
- Liang, D., Tsai, C.-F. & Wu, H.-T. (2015). The Effect of Feature Selection on Financial Distress Prediction. *Knowledge-Based Systems*, (73), 289-297.
- Lo'pez, M. V., Garcia, A., & Rodriguez, L. (2007). Sustainable Development and Corporate Performance: A Study Based on the Dow Jones Sustainability Index. *Journal of Business Ethics*, (75), 285-300.
- Nair, B. B., Mohandas, V., & Sakthivel, N. R. (2010). A Decision Tree- Rough Set Hybrid System for Stock Market Trend Prediction. *International Journal of Computer Applications*, 6 (9), 1-6.
- Oliveira, F. A., Nobre, C. N., & Zárate, L. E. (2013). Applying Artificial Neural Networks to Prediction of Stock Price and Improvement of Directional Prediction Index – Case Study of PETR4, Petrobras, Brazil. *Expert Systems with Applications*, (40), 7596-7606.
- Önder, Ş. (2017). İşletme Kârlılığına Kurumsal Sürdürülebilirliğin Etkisi: BIST'de Bir Uygulama. *Muhasebe Bilim Dünyası Dergisi*, 19 (4), 937-956.
- Özçalıcı, M. (2017). Aşırı Öğrenme Makineleri ile Hisse Senedi Fiyat Tahmini. *Hacettepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 35 (1), 67-88.
- Özdemir, Z., & Pamukçu, F. (2016). Kurumsal Sürdürülebilir Raporlama Sisteminin Borsa İstanbul Sürdürülebilirlik Endeksi Kapsamındaki İşletmelerde Analizi. *Mali Çözüm*, Mart-Nisan, 13-35.
- Özdemir, A. K., Tolun, S., & Demirci, E. (2011). Endeks Getirisi Yönünün İkili Sınıflandırma Yöntemiyle Tahmin Edilmesi: İMKB 100 Endeksi Örneği. *Niğde Üniversitesi İİBF Dergisi*, 4 (2), 45-59.

- Özkan, Y. (2016). *Veri Madenciliği Yöntemleri*, İstanbul: Papatya Yayınları.
- Öztemel, E. (2012). *Yapay Sinir Ağları*. İstanbul: Papatya Yayınları.
- Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting Stock and Stock Price Index Movement using Trend Deterministic Data Preparation and Machine Learning Techniques. *Expert Systems with Applications*, (42), 259-268.
- Silahtaroglu, G. (2016). *Veri Madenciliği Kavram ve Algoritmaları*. İstanbul: Papatya Yayınları.
- Sudha, S. (2015). Risk-Return and Volatility Analysis of Sustainability Index in India. *Environ Dev Sustain*, (17), 1329-1342.
- Yang, B., Gong, Z. J., & Yang, W. (2017). Stock Market Index Prediction using Deep Neural Network Ensemble. Proceedings of the 36th Chinese Control Conference, July 26-28, 2017, Dalian, China, 3882-3887.
- Yakut, E., & Gemici, E. (2017). LR, C5.0, CART, DVM Yöntemlerini Kullanarak Hisse Senedi Getiri Sınıflandırma Tahmini Yapılması ve Kullanılan Yöntemlerin Karşılaştırılması: Türkiye’de BIST’de Bir Uygulama. *Ege Akademik Bakış*, 17 (4), 461-479.
- Yetiş, Y., Kaplan, H., Jamshidi, M., & IEEE, F. (2014). Stock Market Prediction by using Artificial Neural Network. World Automation Congress (s. 1-5). *San Antonio, Texas: World Automation Congress*.
- Zhong, X., & Enke, D. (2017). Forecasting Daily Stock Market Return using Dimensionality Reduction. *Expert Systems With Applications*, (67), 126-139.
- <https://tr.investing.com>. (Erişim Tarihi, 10.07.2020).
- <https://www.kap.org.tr>. (Erişim Tarihi, 10.07.2020).
- <https://www.kpmg.vergi.com>. (Erişim Tarihi: 17.07.2020).

Ekler

Ek 1. ANN Ağırlıkları

Girdi Katmanı	Gizli Katman										Threshold
	Düğüm 1	Düğüm 2	Düğüm 3	Düğüm 4	Düğüm 5	Düğüm 6	Düğüm 7	Düğüm 8	Düğüm 9	Düğüm 10	
	Sigmoid										
X1	1,449	11,376	-11,169	0,555	6,742	-1,814	1,023	-3,740	3,602	-6,633	
X2	-1,065	-0,747	3,933	3,222	4,204	-2,186	2,142	0,214	-0,232	-1,178	
X3	4,846	0,729	-2,143	0,611	-0,836	6,627	4,198	1,708	2,483	8,199	
X4	3,293	-5,793	1,962	-0,094	-3,329	2,119	-1,151	2,021	-6,684	-1,482	
X5	10,006	-2,856	9,941	-2,140	-3,701	-1,104	2,502	-1,367	5,762	1,329	
X6	-12,840	1,960	1,728	8,989	-1,535	-11,093	-2,960	2,931	7,363	9,580	
X7	-1,116	3,087	-3,333	4,723	1,137	-2,370	-2,764	-3,514	-2,529	5,294	
X8	2,726	2,775	4,106	-4,152	9,410	-5,768	0,730	-4,363	1,202	-0,411	
X9	2,977	-1,073	-2,591	-7,970	-0,039	4,520	7,848	-2,862	1,635	11,752	
X10	-0,111	1,080	-1,330	-8,255	4,810	-2,611	-2,869	-3,265	0,363	2,534	
X12	-2,263	-4,082	0,858	7,994	-2,343	-1,961	-2,465	5,736	-9,107	3,536	
X13Ln	-1,501	-1,705	1,368	0,837	1,910	9,372	6,072	-1,772	-11,238	-3,238	
X14Ln	7,004	4,246	-4,189	0,570	-0,135	6,389	-1,250	-8,131	3,841	-1,333	
Bias	-1,419	-2,350	-4,450	-3,289	-0,251	2,821	-1,126	-1,195	-2,744	1,591	
Çıktı katmanı (4)	-4,268	3,432	-2,782	2,236	-6,662	-11,008	-2,173	-1,072	-8,802	2,096	1,472
Çıktı katmanı (3)	-8,742	-16,974	-10,483	8,746	7,315	-3,572	2,445	-7,915	6,795	-4,254	0,885
Çıktı katmanı (2)	-7,678	9,830	0,129	-9,171	-2,155	12,840	-7,269	-1,246	5,219	10,885	-9,890
Çıktı katmanı (1)	9,650	-0,130	4,596	-2,781	-6,703	1,017	10,041	8,432	-1,948	-8,896	-3,582
Çıktı katmanı (0)	5,637	1,945	3,934	3,665	5,884	-5,930	-1,357	-5,522	-6,211	1,017	-5,423

Ek 2. CART Analizi Karar Kuralı

Tree

```

x7 > 0.006
  x7 > 0.032
    x7 > 0.297
      x10 > 0.004
        x7 > 0.338
          x7 > 0.427: 2 {3=0, 1=0, 2=2, 0=0, 4=0}
          x7 <= 0.427
            x7 > 0.408: 3 {3=1, 1=0, 2=0, 0=1, 4=0}
            x7 <= 0.408: 2 {3=0, 1=0, 2=2, 0=0, 4=0}
            x7 > 0.338: 4 {3=1, 1=0, 2=0, 0=0, 4=1}
            x7 <= 0.338
              x10 > 0.004
                x7 > 0.312
                  x7 > 0.448: 3 {3=2, 1=0, 2=0, 0=0, 4=0}
                  x7 <= 0.448
                    x10 > 0.003: 3 {3=1, 1=0, 2=1, 0=0, 4=1}
                    x10 <= 0.003
                      x7 > 0.366: 4 {3=0, 1=0, 2=0, 0=0, 4=3}
                      x7 <= 0.366: 3 {3=1, 1=0, 2=0, 0=0, 4=1}
                      x7 > 0.312: 3 {3=2, 1=0, 2=0, 0=0, 4=0}
                x7 <= 0.312
                  x10 > 0.003
                    x10 > 0.005: 2 {3=0, 1=0, 2=3, 0=0, 4=0}
                    x10 <= 0.005
                      x10 > 0.004
                        x7 > 0.155: 2 {3=0, 1=1, 2=2, 0=0, 4=0}
                        x7 <= 0.155: 4 {3=0, 1=1, 2=0, 0=0, 4=1}
                      x10 > 0.004
                        x7 > 0.223: 2 {3=0, 1=0, 2=3, 0=0, 4=0}
                        x7 <= 0.223
                          x7 > 0.132: 1 {3=1, 1=2, 2=1, 0=0, 4=0}
                          x7 <= 0.132: 2 {3=0, 1=0, 2=2, 0=0, 4=0}
                  x10 <= 0.003
                    x7 > 0.207

```

```

| | | X10 ≤ 0.003
| | | X7 > 0.207
| | | X7 > 0.254: 3 {3=3, 1=0, 2=0, 0=0, 4=0}
| | | X7 ≤ 0.254: 4 {3=1, 1=0, 2=0, 0=0, 4=1}
| | | X7 ≤ 0.207
| | | X7 > 0.127: 2 {3=2, 1=5, 2=9, 0=1, 4=1}
| | | X7 ≤ 0.127
| | | X7 > 0.105
| | | X10 > 0.002: 4 {3=0, 1=0, 2=0, 0=1, 4=1}
| | | X10 ≤ 0.002: 3 {3=5, 1=1, 2=0, 0=1, 4=2}
| | | X7 ≤ 0.105
| | | X7 > 0.041
| | | X7 > 0.083: 2 {3=2, 1=3, 2=5, 0=1, 4=0}
| | | X7 ≤ 0.083: 3 {3=5, 1=1, 2=2, 0=0, 4=0}
| | | X7 ≤ 0.041: 1 {3=0, 1=2, 2=1, 0=1, 4=0}
| X7 ≤ 0.032: 2 {3=0, 1=0, 2=6, 0=0, 4=0}
X7 ≤ 0.006
| X10 > 0.000
| X7 > -0.011: 2 {3=0, 1=1, 2=1, 0=0, 4=0}
| X7 ≤ -0.011: 1 {3=0, 1=6, 2=0, 0=0, 4=1}
| X10 ≤ 0.000
| X7 > -0.965: 0 {3=0, 1=0, 2=1, 0=2, 4=0}
| X7 ≤ -0.965: 3 {3=2, 1=1, 2=0, 0=0, 4=0}

```

Ek 3. Korelasyon Matrisi

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13Ln	X14Ln	X15Ln
X1	1,00	0,04	-0,04	0,07	0,15	0,23	0,31	-0,32	-0,03	0,12	0,29	0,03	0,00	-0,02	-0,03
X2		1,00	0,51	0,28	-0,33	-0,11	0,23	-0,13	0,12	0,12	0,27	-0,05	-0,07	-0,12	-0,09
X3			1,00	0,07	-0,30	-0,05	0,09	-0,45	0,09	-0,15	0,07	0,03	-0,01	-0,03	-0,01
X4				1,00	0,05	0,08	0,25	-0,18	0,10	0,40	0,32	0,09	0,17	0,17	0,18
X5					1,00	0,58	0,32	-0,28	-0,08	0,24	0,35	0,08	0,09	0,08	0,09
X6						1,00	0,65	-0,42	-0,01	0,10	0,58	-0,06	-0,06	-0,09	-0,07
X7							1,00	-0,36	0,07	0,34	0,84	0,09	0,01	0,02	0,06
X8								1,00	-0,01	0,24	-0,28	0,09	0,16	0,17	0,17
X9									1,00	0,08	0,05	0,13	-0,05	0,12	0,12
X10										1,00	0,49	0,07	0,16	0,24	0,20
X11											1,00	0,12	0,02	0,03	0,07
X12												1,00	0,55	0,71	0,78
X13Ln													1,00	0,70	0,81
X14Ln														1,00	0,92
X15Ln															1,00