MAKİNE ÖĞRENMESİ İLE PORTFÖY OPTİMİZASYONU: FTSE, DAX VE BIST UYGULAMALARI

Umut UYAR*

***Dr. Öğr. Üyesi, Pamukkale Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü**
uuyar@pau.edu.tr

ÖZET

Portföy optimizasyonu yarım yüzyılı aşkın süredir birçok araştırmacının ve yatırımcının test ettiği ve kullandığı bir teoridir. Ancak konu üzerine yapılan bazı çalışmalarda teoriye bir takım eleştiriler de getirilmektedir. Bu eleştirilerden bir tanesi de konsantrasyon problemidir. Temelinde çeşitlendirme olan bir teori ile elde edilen optimum portföylerin yatırım olanakları kümesindeki az sayıda varlığa yatırım yapılmasını önermesi önemli bir eleştiri noktasını oluşturmaktadır. Bu problemin çözümü üzerinde çalışan araştırmacılardan birisi olan Prado (2016; 2018), geliştirdiği Hiyerarşik Risk Paritesi (HRP) metodu ile makine öğrenmesi kullanılarak portföy optimizasyonu yapılabileceğini ifade etmektedir. Bu çalışmanın amacı HRP metodu kullanılarak BIST, FTSE ve DAX piyasalarında Temmuz 2005 - Haziran 2017 aralığında makine öğrenmesi algoritmalarının portföy performanslarını incelemektir. Analizler sonucunda, HRP metodunun BIST, FTSE piyasalarında başarılı olmamasına rağmen DAX piyasalarında başarılı performans sergilediğine ulaşılmıştır.

Anahtar Kelimeler: Makine Öğrenmesi, Portföy Optimizasyonu, FTSE, DAX, BIST

MACHINE LEARNING ASSET ALLOCATION: APPLICATIONS FROM FTSE, DAX AND BIST

ABSTRACT

Portfolio management is a theory used by researchers and investors more than half of the century. Although it is a commonly used theory, several studies are criticizing the theory. One of these criticisms is the concentration problem. The theory is a diversification base method. However, when you calculate the optimal portfolios, you usually get a few assets to invest. Prado (2016; 2018) searches the solution of that problem and he suggests Hierarchical Risk Parity (HRP) approach. HRP applies machine-learning techniques to build diversified portfolios based on the information contained in the covariance matrix. The aim of this study is to investigate the performance of HRP in BIST, FTSE, and DAX for the period July 2005 – June 2017. The results show that the performance of HRP in BIST and FTSE is not successful; on the other hand, it has a successful performance in DAX.

Keywords: Machine Learning, Portfolio Optimization, FTSE, DAX, BIST

1. GİRİŞ

Portföy optimizasyonu, modern finansal piyasaların doğuşundan bu güne yatırımcıların, aracı kurumların ve araştırmacıların önemli problemlerinden birisi olmuştur (Bachelier, 1900). Erken dönem finans çalışmalarında portföy teorisinin, daha çok çeşitlendirme üzerine kurulu olduğu görülmektedir. Tüm varlıkların tek sepete konulmaması, riskin dağıtılması gerektiği mantığına dayanan bu çalışmalar, modern portföy teorisinin temellerini oluşturmaktadır (Hicks, 1935; Marschak, 1938; Williams, 1938; Leavens, 1945). Nitekim portföyde yer alan varlıkları çeşitlendirmenin tek başına yeterli olmadığının anlaşılması çok uzun sürmemiştir. Varlıkların çeşitlendirilmesi riskin dağıtılmasına imkan sağlarken, varlıkların kendi aralarındaki ilişkiyi de dikkate almanın asıl çözüm olduğu 1952 yılında Harry Markowitz tarafından ortaya atılmıştır. Markowiz (1952), portföy teorisinde finansal varlıkları çeşitlendirmenin tek başına yeterli olmayacağını, aynı zamanda portföydeki varlıkların kendi aralarındaki ilişkisinin de (korelasyon-kovaryans) dikkate alınması gerektiğini ifade etmektedir. Portföy optimizasyonu üzerine çalışmaları sonucunda, aynı risk düzeyinde farklı getiri seviyelerine sahip portföyler olabileceği, ancak bunlardan sadece bir tanesinin en yüksek getiriye sahip olduğunu kanıtlamıştır. Farklı risk düzeylerindeki en yüksek getiriye sahip portföylerin bir koordinat düzleminde birleşmesi ile "Etkin Sınır" kavramını oluşturmuştur. Etkin sınır üzerinde yer alan portföyler de optimal portföyler olarak isimlendirilmiştir. Basit anlamda, belirli bir yatırım olanakları kümesi içerisindeki varlıklar kullanılarak, kabul edilen risk düzeyinde en yüksek getirinin elde edilebileceği portföylerin nasıl hesaplanacağı Markowitz sonrasında çözüme ulaşmıştır. Böylece modern portföy teorisi oluşmuştur.

Modern portföy teorisi üzerine aradan geçen yarım yüzyıldan fazla sürede yüzlerce araştırmacı ve yatırımcı çalışmalar yapmıştır. Yapılan analizlerde optimal portföylerin başarıları ve başarısızlıkları ortaya konulmaya çalışılmıştır. Lakin etkin sınır üzerinde yer alan optimal portföylerin önemli bir problemi bulunmaktadır. Tüm teorinin temelinde çeşitlendirme olmasına rağmen, optimal portföyler birçok durumda, belirlenen risk düzeyinde eldeki yatırım kümesindeki varlıkların oldukça azına yatırım yapmayı önermektedir. Başka bir deyişle, optimal portföyler sıklıkla birkaç varlık üzerine konsantre olmuş portföyler üretmektedir. Özellikle yüksek dalgalanmaya (volatiliteye) sahip varlıklardan oluşturulan portföylerin bu şekilde konsantre olması, yatırımcının beklenen getiriden ciddi biçimde sapmasına neden olmaktadır (Jobson ve Korkie, 1980; Michaud, 1998). Optimal portföylerin bahsi geçen problemine çözüm bulmak amacıyla birçok araştırmacı farklı yöntemler/yaklaşımlar önermiş ve uygulamıştır (Jorion, 1985; 1986; Frost ve Savarino, 1986; Michaud, 1989; Best ve Grauer, 1991; Ledoit ve Wolf, 2003; DeMiguel vd. 2005; Bera ve Park, 2008). Yatırım olanakları kümesini daraltmayı veya farklı teknikler kullanarak portföy ağırlıklarını belirlemeyi amaçlamış çalışmaların da başarılı ve başarısız olduğu durumlarla karşılaşılmıştır. Özellikle, optimal portföyleri daha fazla varlık ile oluşturmaya çalışıldığı durumlarda; eklenen varlıklar dikkat edilmesi gereken daha fazla korelasyona, korelasyonun çözümü daha fazla çeşitlendirmeye, fazla çeşitlendirme de istikrarsız portföyler oluşmasına neden olmuştur. Ne yazık ki sorunun çözüm önerileri yeni sorunlar oluşturarak önerdikleri çözümlerin faydasını ortadan kaldırmıştır. Prado (2018:223) çalışmasında, çözüm önerilerinin optimal portföyleri daha da sıkıntılı duruma yönlendirmesini ve çeşitlendirmenin sağlanamıyor olmasını "Markowitz'in Laneti" olarak isimlendirmiştir.

Prado (2016; 2018), yukarıda bahsedilen problemlerin çözümüne yönelik farklı bir bakış açısı ortaya koymuştur. Son yıllarda gelişen teknoloji ile birlikte sıklıkla kullanılan bilgisayar yazılımlarının Markowitz'in lanetine çözüm olabileceğini ifade etmiştir. Bu amaçla, matematik, grafik teorisi ve makine öğrenmesi kullanarak "Hiyerarşik Risk Paritesi (HRP)" metodu ile portföy optimizasyonu yapılabileceğini, kullandığı simülasyonlar aracılığı ile göstermiştir. HRP metodu ile portföy optimizasyonunda doğrudan veride yer alan bilgi kullanılmakta ve portföy optimizasyonu yöntemlerindeki matris cebir işlemlerine gerek kalmamaktadır. HRP, basit anlamda kovaryans matrisinden başlayan bir kümeleme algoritmasına dayanmaktadır. Yöntem, farklı sektörlerdeki ya da türlerdeki finansal varlıklara (örn. hisse senedi, tahvil, gayrimenkul, türev ürünler vs.) yatırım yapmayı amaçlayan bir yatırımcıya, varlıkların uzaklık matrislerini hesaplayarak makine öğrenmesi ile gruplama yapma imkanı sağlamaktadır. Markowitz portföy optimizasyonu yönteminde, oluşturulan kovaryans matrisinde eşit yatırım olasılığı bulunan bu varlıklar, HRP algoritmasında, hiyerarşik bir yatırım olasılığına sahip olmaktadır. Örneğin, Çin hükümet tahvili, Londra'daki bir yatırım fonu ya da Brezilya'daki telekomünikasyon hisseleri aynı küme içerisinde yer almayarak, hiyerarşik yatırım olasılıklarına sahip olmaktadır. Prado (2016) çalışmasında HRP ile Ortalama-Varyans ve Ters-Varyans Optimizasyon (IVP) metodlarını karşılaştırmış ve simülasyon sonuçlarında HRP'nin rakiplerine göre üstün olduğunu vurgulamıştır.

Bu çalışmanın amacı, makine öğrenmesi ile HRP metodu kullanılarak gerçek piyasa verisi ile tekniğin test edilmesindir. Bu amaçla Avrupa kıtasında karşılaştırma yapılmakta ve Borsa İstanbul Tüm Endeksi (BISTTUM), Frankfurt Borsası (DAX Classic All Share) ve Londra Borsası'nda (FTSE All Share) Temmuz 2005 ile Haziran 2017 dönemlerinde işlem gören hisse senetlerine ait aylık getiriler kullanılmaktadır. Seçilen üç piyasa, etkin piyasa hipotezine göre güçlü-etkin (FTSE), orta-etkin (DAX) ve zayıf-etkin (BIST) piyasa türlerini temsil etmektedir. Analizlerde Fama ve French'in (1993) yatırım metodolojisinden yararlanılmış ve her 12 ayda bir (Temmuz$_{t-1}$ – Haziran$_t$) güncellenen portföyler üretilmiştir. Portföylerin üretilmesinde HRP metodunun yanında kıyaslama yapılabilmesi için Minimum

Varyans ve Ters-Varyans portföyleri de oluşturulmuştur. Yöntemlerin performans testlerini yapabilmek amacıyla Sharpe rasyosu, düzeltilmiş Sharpe rasyosu, kesin-eşdeğer getiri ve maksimum düşüş performans kriterlerinden yararlanılmıştır.

Çalışmanın sonraki bölümünde literatür araştırması yer almaktadır. Üçüncü bölümde analiz metodolojisi detaylı bir şekilde incelenirken, dördüncü bölümde ise analizlerde kullanılacak veri setine ilişkin değerlendirmeler bulunmaktadır. Beşinci bölümde bulguların yorumuna, sonrasında ise sonuç bölümüne yer verilmiştir.

2. LİTERATÜR ARAŞTIRMASI

Makine öğrenmesi ile Hiyerarşik Risk Paritesi (HRP) metodu 1980'lerden itibaren çalışılmakta olan bir konudur. Özellikle, bilgisayar teknolojisinin yaygınlaşması ile mühendislik alanındaki birçok araştırmacının ilgisini çekmiştir. Ancak, konunun finans alanında, portföy optimizasyonu üzerine yapılan uygulanmaları ve uyarlamaları oldukça güncel ve yeni bir alan olarak dikkat çekmektedir.

Konu ile doğrudan ilgili ilk çalışma Bailey ve Prado (2012) tarafından yayımlanmıştır. Yazarlar, çalışmalarında öncelikle finansal varlıkların dengeli ağırlıklandırıldığı portföy yatırımlarının pratikte ne kadar önemli olduğuna dikkat çekmişlerdir. Çalışmada, dengeli portföy yatırımı planlamak amacıyla eşit riskli varlıklar, maksimum çeşitlendirme ve min-maks alt-set korelasyon yöntemleri olmak üzere üç faklı metodoloji kullanılmıştır. Analizlerde, herhangi bir yöntemin diğerine göre üstünlüğünü incelemek yerine, her bir yöntemin makine öğrenmesi açsından uygulama alanları ve detayları ortaya konulmuştur. Yazarlar 2013 yılındaki diğer çalışmalarında ise, alt ve üst ağırlık kısıtlarına sahip Markowitz ortalama-varyans optimizasyonu üzerine odaklanmışlardır (Bailey ve Prado, 2013). Portföy optimizasyonunun bu gibi durumları için Markowitz (1956) tarafından kritik doğru algoritmasının önerildiğini vurgulayan yazarlar, algoritmanın adım adım uygulama safhalarını içeren bütün bir eser ortaya koymayı amaçlamışlardır.

Prado (2016; 2018), hiyerarşik risk paritesi yöntemini sunduğu çalışmalarında, geleneksel portföy optimizasyonu yönteminin kararsızlık, düşük performans ve belli varlıklara konsantre olma problemlerini çözüme ulaştırdığını iddia etmiştir. Çalışmalarda uygulanan Monte-Carlo simülasyonları sonucunda HRP metodu ile oluşturulmuş portföylerde, Markowitz'in kritik doğru algoritması kullanılarak oluşturulanlara göre daha az beklenmeyen varyans oluştuğu gösterilmiştir. Hiyerarşik kümeleme temelli portföy optimizasyonu yöntemini kullanan Raffinot (2017), yöntemin portföy ağırlıklarını belirlerken araştırmacı tarafından önceden belirlenmiş bir formu değil, veriden elde ettiği formu kullandığını vurgulamıştır. S&P500 listesinde yer alan tüm hisse senetlerinin kullanıldığı çalışmada, önerilen yöntemin geleneksel portföy optimizasyonu karşısındaki performansını değerlendirmek amacıyla düzeltilmiş Sharpe rasyosu, kesin-eşdeğer getiri, maksimum düşüş, ortalama geri dönüş ve portföy ağırlıklarının kareleri toplamı kriterleri kullanılmıştır. Analizler sonucunda hiyerarşik kümeleme temelli portföy optimizasyonuna göre oluşturulmuş portföylerin mümkün olduğunca çeşitlendirilmiş ve istatistiksel olarak daha optimize risk performansına sahip olduğu sonucuna ulaşılmıştır. Jothimani (2017), çalışmasında Hindistan piyasasında yer alan finansal farklıkları üç farklı portföy optimizasyonu yöntemi kullanarak makine öğrenmesi çerçevesinde karşılaştırmayı ve portföylerin fiyat öngörülerini yapmayı amaçlamıştır. Yazar, iki bölüme ayırdığı çalışmasının ilk bölümünde, portföy ağırlıklarını belirlemek amacıyla Geber istatistikleri bazlı hiyerarşik risk paritesi yöntemini, tarihsel korelasyon bazlı minimum varyans modelini ve Geber istatistikleri bazlı minimum varyans modelini karşılaştırmıştır. Sonrasında, altı farklı öngörü tekniği kullanarak oluşturulan portföylerin fiyat öngörüleri hazırlanmıştır. Çalışmanın ikinci bölümünde ise Hindistan borsasındaki yatırımcılara bir farkındalık anketi uygulanmıştır. Araştırmanın ilk bölümünde, yatırımcılara al/sat/bekle sinyallerini içeren bilimsel bir yol haritası sunulurken, ikinci bölümünde ise yatırımcıların finansal teorileri ve matematiksel/istatistiksel yöntemleri neredeyse hiç kullanmadı sonucuna ulaşılmıştır.

Butler (2018), çalışmasında makine öğrenmesi ile optimizasyon metodunu kullanarak, gerçek hayatta portföy oluştururken analistlerin ve yatırımcıların karşılaştıkları sorunlara çözüm üretmeyi amaçlamıştır. Dokuz farklı portföy optimizasyonu yönteminin etkinliğinin makine öğrenmesi ile test edildiği çalışmanın uygulama bölümünde, Kenneth French'in internet sayfasından elde edilen portföyler de kontrol verisi kullanılmıştır. Simülasyonlar sonucunda, minimum varyans ve hiyerarşik minimum varyans yöntemlerinin diğer optimizasyon yaklaşımlarına göre daha iyi performans gösterdiği ortaya konulmuştur. Liu vd. (2018), çalışmalarında yeni bir yöntem olarak Yeterlik Paritesi Yöntemi isminde bir portföy optimizasyon tekniği sunmuşlardır. Yazarlar, risk paritesi yöntemi üzerine inşa edilen yeni yöntemlerinin, geleneksel optimizasyon yöntemlerine göre çok daha yüksek performansa sahip olduğunu ve pratikte uygulamasının oldukça kolay olduğunu iddia etmişlerdir. Çin yatırım fonu piyasasında uygulamasının yapıldığı yöntem, yatırımcılara, ortalama-varyans, maksimum Sharpe rasyosu, risk paritesi ve CSI300 endeks getirisine göre çok daha yüksek getiri sağlamaktadır. Çalışmada, yeni yöntemin tek sorununun yüksek devir oranına sahip olması gösterilirken, uzun vadede yüksek olan performansın bu sorunu ortadan kaldırdığı vurgulanmaktadır. Raffinot, 2018 yılında konu ile ilgili yayımladığı bir diğer çalışmasında, ilk çalışmasında kullandığı hiyerarşik risk paritesi yöntemi ile hiyerarşik kümeleme bazlı portföy optimizasyonu yöntemlerini birleştirerek makine öğrenmesi temelli yeni bir portföy optimizasyonu metodu ortaya atmıştır. Hiyerarşik Eşit Risk Paylı Portföy adının verildiği yöntemi risk ölçütleri açısından, koşullu riske maruz değer (conditional value at risk) ve koşullu risk düşüşü (conditional drawdown at risk) olmak üzere ikiye ayırmıştır. Çalışmanın uygulama bölümünde Avrupa ve ABD'de yer alan 18 farklı endeks verisi ve S&P500 listesinde yer alan 365 hisse senedine ait veri, iki ayrı test grubu oluşturularak analiz edilmiştir. Yazar, analizler sonucunda yeni metodun özellikle koşullu risk düşüşü ölçütünü kullanan versiyonunun riske göre düzeltilmiş performansının istatistiksel olarak oldukça yüksek olduğunu belirtmiştir.

3. METODOLOJİ

Hiyerarşik Risk Paritesi (HRP) metodu kullanılarak oluşturulan portföy yapıları, grafik teorisi ve makine öğrenmesini kullanmaktadır. HRP kullanılarak portföy ağırlıklarının elde edilmesinde doğrudan veride yer alan bilginin kullanılmasıyla, klasik portföy optimizasyonu yöntemlerine göre, matris cebiri işlemlerine gerek kalmamaktadır. HRP algoritması, Prado'nun (2018:224) çalışmasında detaylı bir şekilde açıklanmakta ve üç aşamadan oluşmaktadır: Ağaç kümeleme, yarı-köşegenleştirme ve ardışık ikili seçim.[28] Bölümün sonraki kısmında her bir aşama ve içerdiği işlemler ayrıntılı olarak ele alınmaktadır. Çalışmanın amacı gereği, HRP algoritmasının farklı versiyonları ve ortalama-varyans portföy optimizasyonuna göre oluşturulmuş portföylerin karşılaştırılabilmesi için bir takım performans kriterleri kullanılmaktadır. Literatürde portföy performanslarının değerlendirilebilmesi için geniş bir performans kriteri yelpazesi bulunmaktadır. Bu çalışma için Raffinot (2017), tarafından önerilen portföy performans değerlendirme kriterleri tercih edilmiştir.

3.1. Ağaç kümeleme

HRP algoritmasının ilk aşamasında ağaç kümeleme işlemi uygulanmaktadır. Bu işlem için öncelikle N adet serinin T zaman aralığı için "$T \times N$" boyutunda bir matris ele alınmaktadır. Vektör olarak belirlenen her bir seri için hiyerarşik bir kümeleme yapısı oluşturulmak istenmektedir. Bunun için ilk olarak, her bir elemanı $\rho = \{\rho_{i,j}\}_{i,j=1,\dots,N}$ olan $\rho_{i,j} = \rho[X_i, X_j]$ şeklinde "$N \times N$"lik korelasyon matrisi oluşturulmaktadır. Daha sonra korelasyonlar kullanılarak seriler arası uzaklık ölçüleri olarak d_{ij} Eşitlik 1'de gösterildiği şekli ile hesaplanmaktadır.

$$\begin{aligned} &d: (X_i, X_j) \subset B \to \mathbb{R} \in [0,1] \\ &d_{i,j} = d[X_i, X_j] = \sqrt{\frac{1}{2}(1 - \rho_{i,j})} \end{aligned} \tag{1}$$

Eşitlik 1'de yer alan B, $\{1, \dots, i, \dots, N\}$ kümesinin Kartezyen çarpımını ifade etmektedir. Uzaklık ölçüsü olan $d_{i,j}$'lerin hesaplanması ile kümeleme analizinde kullanılacak olan "$N \times N$" boyutundaki uzaklık matrisi ($X = \{d_{i,j}\}_{i,j=1\dots,N}$) oluşturulmaktadır. Hesaplamalarda, X matrisi negatif olmayan,

[28] Terimlerin orijinal karşılıkları: Tree clustering, quasi-diagonalization, recursive bisection.

tesadüfi, simetrik, mutlak ve uygun bir metrik uzayı ifade etmektedir. Ağaç kümeleme analizi için uzaklık matrisinin oluşturulmasından sonra, matrisin her iki sütunu arasındaki Öklidyen uzaklıkların hesaplanması işlemine geçilmektedir. Bu işlem için Eşitlik 2'den yararlanılmaktadır.

$$\tilde{d}: (D_i, D_j) \subset B \rightarrow \mathbb{R} \in \left[0, \sqrt{N}\right]$$
$$\tilde{d}_{i,j} = \tilde{d}\left[D_i, D_j\right] = \sqrt{\sum_{n=1}^{N} \left(d_{n,i} - d_{n,j}\right)^2} \qquad (2)$$

Eşitlik 2'de $d_{i,j}$, X matrisinin elemanları olarak tanımlanırken, $\tilde{d}_{i,j}$ ise, D matrisinin elemanları olarak tanımlanmaktadır. Ayrıca, $\tilde{d}$, D matrisinin tüm metrik uzayı için tanımlanmış bir uzaklık ölçüsüdür. Diğer bir açıklama ile korelasyon uzaklık değerleri arasındaki Öklidyen uzaklıkları ifade etmektedir.

Öklidyen uzaklıkların belirlenmesinin ardından, kümeleme algoritmasının uygulanmasına geçilmektedir. Bu aşamasında her bir sütun çifti i^* ve j^* kümeleri olarak çiftlere ayrılmakta ve belirlenen "*Bağlanım Kriteri*"ne (Linkage Criterion) göre dönüştürülerek $u[1]$ matrisi elde edilmektedir. Literatürde, bu işlem için kullanılabilecek yedi farklı bağlanım kriteri yer almaktadır. Bunlar *Tek, Tam, Ortalama, Ağırlıklandırılmış, Merkezi, Bölgesel* ve *Medyan*[29] kriterleridir (Day & Edelsbrunner, 1984; Bar-Joseph vd. 2001; Zhao & Karypis, 2002; Müllner, 2011; 2013; Prado, 2018:226).

En yakın nokta algoritması olarak bilinen Tek Bağlanım Kriteri, Eşitlik 3'te gösterildiği şekilde hesaplanmaktadır.

$$(i^*, j^*) = argmin(i,j)_{i \neq j}\{\tilde{d}_{i,j}\} \qquad (3)$$

En uzak nokta ya da Voor Hees algoritması olarak da bilinen Tam Bağlanım Kriteri Eşitlik 4'te gösterildiği şekilde hesaplanmaktadır.

$$(i^*, j^*) = argmaks(i,j)_{i \neq j}\{\tilde{d}_{i,j}\} \qquad (4)$$

Literatürdeki diğer adı UPGMA (Unweighted Pair Group Method with Arithmetic mean) algoritması olan Ortalama Bağlanım Kriteri Eşitlik 5'te gösterildiği şekilde hesaplanmaktadır. Eşitlik 5'te yer alan $|i^*|$ ve $|j^*|$ ifadeleri her bir kümeye ait nicel sayı olarak tanımlanmaktadır.

$$(i^*, j^*) = \frac{1}{|i^*||j^*|} \sum_{i,j} \{\tilde{d}_{i,j}\} \qquad (5)$$

Diğer adı WPGMA (Weighted Pair Group Method with Arithmetic mean) algoritması olan Ağırlıklandırılmış Bağlanım Kriteri Eşitlik 6'da gösterildiği şekilde hesaplanmaktadır. Eşitlik 6'da i^* kümesi s ve t kümeleri kullanılarak tekrar oluşturulurken, bu yeni kümeler j^* kümesi üzerinden ağırlıklandırılmaktadır.

$$(i^*, j^*) = \frac{\left(\{\tilde{d}_{s,j}\} + \{\tilde{d}_{t,j}\}\right)}{2} \qquad (6)$$

[29] Terimlerin orijinal karşılıkları: Single, Complete, Average, Weighted, Centroid, Ward, Median.

Diğer adı UPGMC (Unweighted Pair Group Method using Centroids) olan Merkezi Bağlanım Kriteri Eşitlik 7'de gösterildiği şekilde hesaplanmaktadır. Eşitlik 7'de $\vec{c}_i$ ve $\vec{c}_j$ sırasıyla i^* ve j^* kümelerinin ağırlık merkezlerini (Centroid) ifade etmektedir. Kriter, belirlenen ağırlık merkezleri arasındaki Öklidyen uzaklıkların hesaplanmasına dayanmaktadır.

$$(i^*, j^*) = \left\| \vec{c}_i - \vec{c}_j \right\|_2 \qquad (7)$$

Kademeli algoritma olarak da adlandırılabilen Bölgesel Bağlanım Kriteri Eşitlik 8'de gösterildiği şekilde hesaplanmaktadır.

$$(i^*, j^*) = \sqrt{\frac{2|i^*||j^*|}{|i^*| + |j^*|}} \left\| \vec{c}_i - \vec{c}_j \right\|_2 \qquad (8)$$

Son bağlanım kriteri olan ve WPGMC (Weighted Pair Group Method using Centroids) algoritması olarak da adlandırılan Medyan Bağlanım Kriteri ise Eşitlik 9'da gösterildiği şekilde hesaplanmaktadır. Bu kriterin hesaplama adımları merkezi bağlanım kriterine oldukça benzerlik göstermektedir. Eşitlik 9'da yer alan $\vec{\omega}_i$ ve $\vec{\omega}_j$ ifadeleri sırasıyla i^* ve j^* kümelerinin medyan değerlerini göstermektedir. Kriter, belirlenen medyan değerleri arasındaki Öklidyen uzaklıkların hesaplanmasına dayanmaktadır.

$$(i^*, j^*) = \left\| \vec{\omega}_i - \vec{\omega}_j \right\|_2 \qquad (9)$$

Bağlanım kriterinin belirlenmesinin ardından $\tilde{d}$ matrisinin ilk iki sütunundan oluşan $u[1] = (1,2)$ matrisleri oluşturulmaktadır. $\tilde{d}$ matrisinde yer alan ve gruplanmış olan sütunlar, her bir sütun çifti olan i^* ve j^* bağlanım kriteri ve $u[1]$ matrisi kullanılarak $\dot{d}_{i,u[1]}$ matrisine dönüştürülmektedir. Tek sütundan oluşan $\dot{d}_{i,u[1]}$ matrisi, sonraki aşamada $\tilde{d}$ matrisinin güncellenmesinde kullanılmaktadır. Ağaç kümeleme işlemlerinin son aşamasında ise güncellenmiş $\tilde{d}$ matrisinin son dört elemanından oluşan $u[2] = (3,4)$ matrisi oluşturulmakta ve kümeleme işlemi tamamlanmaktadır (Rokach & Maimon, 2005).

Ağaç kümeleme işleminin sonunda araştırmacı $(N-1)x4$ boyutunda ve $Y = \{(y_{m,1}, y_{m,2}, y_{m,3}, y_{m,4})\}_{m=1,\dots,N-1}$ şeklinde bir "*Bağlantı Matrisi*" elde etmektedir. Bağlantı matrisinde (Y) her bir küme için dört bağlantı verisi yer almaktadır. Matriste "$y_{m,1}, y_{m,2}$" elemanları bileşen verisini; $y_{m,3}$ elemanı bileşenler arasındaki uzaklığı ($y_{m,3} = \tilde{d}_{y_{m,1}y_{m,2}}$); $y_{m,4}$ ve diğer elemanlar ($y_{m,4} \leq N$) ise bağlantı kurulan orijinal eleman sayısını ifade etmektedir (Prado, 2018:229).

3.2. Yarı-Köşegenleştirme

HRP algoritmasının bu adımında, kovaryans matrisinde yer alan satır ve sütunlarında yer alan en büyük değerlerin köşegen boyunca yer alacak şekilde yeniden düzenlemesi işlemi bulunmaktadır. Gerçekleştirilen yarı-köşegenleştirme işlemi finansal anlamda benzer yatırımların bir araya getirilmesine; farklı yatırımların ise birbirinden ayrılmasına olanak sağlamaktadır. Bu şekilde oluşturulan kümeler, finansal yatırımlar bakımından da optimize edilmiş olmaktadır. Bu adımda gerçekleştirilecek ilk işlem bağlantı matrisinde (Y) yer alan elemanların kümeler yok edilinceye kadar birleştirilmesi ile başlamaktadır. Her bir y elemanı "$(y_{N-1,1}, y_{N-1,2})$" şeklinde gruplar haline birleştirilmektedir. Bu birleştirme işlemi kümeleri ortadan kaldırmasına rağmen, küme sırasının korunmasını sağlamaktadır. İşlem sonucunda elde edilen matris, orijinal (kümelenmemiş) değerlerin sıralı bir listesi şekildedir (Prado, 2016).

3.3. Ardışık İkili Seçim

HRP algoritmasının üçüncü ve son adımında, yarı-köşegen matrisi kullanılarak bir ardışık ikili seçim işlemi izlenmektedir. Ardışık ikili seçim işlemlerini maddeler halinde özetlemek mümkündür (Prado, 2018: 230):

1- İşlemlere geçmeden önce bir takım tanımlamaların yapılması gerekmektedir

(a) Bir eleman seti olarak $L = \{L_0\}$ ve $L_0 = \{n\}_{n=1,\ldots,N}$ tanımlanmaktadır.

(b) Tüm elemanların toplamı (ağırlıklar) ise $w_n = 1$ ve $\forall n = 1, \ldots, N$ şeklinde tanımlanmaktadır.

2- İşlemler döngüsünde $|L_i| = 1$ $(\forall L_i \in L)$ durumu için döngünün durması sağlanır.

3- Eğer $\forall L_i \in L$ için $|L_i| > 1$ ise;

(a) L_i iki ardışık alt sete ayrılır: $L_i^{(1)} \cup L_i^{(2)} = L_i$. Burada $\left|L_i^{(1)}\right| = int\left[\frac{1}{2}\left|L_i^{(1)}\right|\right]$ dir.

(b) $L_i^{(j)}$ $(j = 1,2)$ setinin varyansı $\widetilde{V}_i^{(j)} \equiv \widetilde{w}_i^{(j)'} V_i^{(j)} \widetilde{w}_i^{(j)}$ şeklinde tanımlanmaktadır. Burada $V_i^{(j)}$, $L_i^{(j)}$ ardışık setleri arasındaki kovaryans matrisini; $\widetilde{w}_i^{(j)}$, ise $\widetilde{w}_i^{(j)} = diag\left[V_i^{(j)}\right]^{-1} \frac{1}{tr\left[diag\left(V_i^{(j)}\right)^{-1}\right]}$ işlemini ifade etmektedir.[30]

(c) Ayrık faktörünün hesaplanması: $\alpha_i = 1 - \frac{\widetilde{V}_i^{(1)}}{\widetilde{V}_i^{(1)} + \widetilde{V}_i^{(2)}}$ $(0 \leq \alpha \leq 1)$ formülü ile gerçekleştirilir.

(d) Ağırlıklar (w_n), α_i faktörü kullanılarak yeniden hesaplanmaktadır $(\forall n \in L_i^{(1)})$.

(e) Ağırlıklar (w_n), $(1 - \alpha_i)$ faktörü kullanılarak yeniden hesaplanmaktadır $(\forall n \in L_i^{(2)})$.

4- İşlem ikinci adımdan itibaren tekrarlanır.

Maddeler halinde tanımlanan ardışık ikili seçim döngüsü, her tekrarında bir üst hiyerarşik gruptan alınan ağırlıkları iki sete ayırdığı için ağırlıkların sıfır ile bir arasında gerçekleşmesi ve tüm ağırlıklar toplamının bire eşit olması şartlarını sağlamaktadır ($0 \leq w_i \leq 1, \forall i = 1, \ldots, N, \sum_{i=1}^{N} w_i = 1$). Üçüncü ve son adım olan ardışık ikili seçimin tamamlanmasının ardından, portföydeki finansal varlıklar için HRP algoritması ile oluşturulmuş "$N\ x$ 1" matrisi şeklinde portföy ağırlıkları elde edilmektedir.

3.4. Kıyas Portföy Optimizasyon Yöntemleri

HRP algoritması ile makine öğrenmesi yöntemine göre oluşturulan portföylerin performans kıyaslamalarını yapabilmek amacıyla Butler (2018) ve Prado (2018) çalışmalarında kullanılan minimum varyans portföyü (IVP) ve ters-varyans optimizasyonu (GMV) yöntemlerinden yararlanılmıştır. İlgili yöntemlerim hesaplamaları sırasıyla Eşitlik 10 ve Eşitlik 11'de gösterilmektedir (Markowitz, 1952; Pennacchi, 2008:54; Butler, 2018).

$$w_{MV} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}\Sigma^{-1}\mathbf{1}} \tag{10}$$

$$w_i^{IVar} = \frac{\frac{1}{\sigma_i^2}}{\sum_{i=1}^{n}\left(\frac{1}{\sigma_i^2}\right)} \tag{11}$$

Eşitlik 10'da w_{MV}, minimum varyans portföyüne ait ağırlıkların yer aldığı matrisi; Σ, portöye dahil edilen finansal varlıklar arası kovaryans matrisini; $\mathbf{1}$ ise, bir değerlerinden oluşan sütun vektörünü ifade etmektedir Eşitlik 11'de ise, n, portföye dahil edilen varlık sayısını; w_i^{IVar}, i. varlığın ters-varyans optimizasyonuna göre hesaplanmış portföy içindeki ağırlıklarını; σ_i^2, i. varlığın araştırma zaman aralığındaki varyans değerini göstermektedir.

[30] $diag[.]$ ve $tr[.]$ sırasıyla köşegen ve iz işlemini ifade etmektedir.

3.5. Portföy Performans Değerlendirme Kriterleri

Çalışmanın amacı doğrultusunda oluşturulan portföylerin performansları değerlendirilirken bir takım performans ölçüm kriterlerine ihtiyaç duyulmaktadır. Portföy performansı literatüründe oldukça çok seçenek bulunmasına rağmen, bu çalışmada makine öğrenmesi konusuna uygun olması amacıyla Raffinot (2017), tarafından önerilen düzeltilmiş Sharpe rasyosu, kesin-eşdeğer getiri, maksimum düşüş ve portföy ağırlıklarının kareleri toplamı yöntemleri[31] tercih edilmiştir.

Düzeltilmiş Sharpe rasyosu (ASR), geleneksel Sharpe rasyosuna, negatif çarpıklığı ve basıklığı önlemek için bir ceza faktörü eklenmesi ile oluşturulmaktadır. Eşitlik 12'de Sharpe rasyosu (SR), Eşitlik 13'te ise düzeltilmiş Sharpe rasyosunun hesaplama formülleri gösterilmektedir (Sharpe, 1966; Pezier ve White, 2008).

$$SR = \frac{\mu - r_f}{\sigma} \tag{12}$$

$$ASR = SR\left[1 + \left(\frac{\mu_3}{6}\right)SR + \frac{(\mu_4 - 3)}{24}SR^2\right] \tag{13}$$

Eşitlik 10 ve 11'de, μ finansal varlığın ortalama (beklenen) getirisini; r_f, risksiz varlık getirisini[32]; σ, finansal varlığın standart sapmasını (risk); μ_3 ve μ_4 ise sırasıyla çarpıklık ve basıklığı ifade etmektedir. Kesin-eşdeğer getiri (CEQ) yatırımcının riskli yatırım stratejisi yerine kabul ettiği risksiz varlık getiri oranını ifade etmektedir. CEQ, Eşitlik 14'de gösterildiği şekilde hesaplanmaktadır (Frost ve Savarino, 1986; DeMiguel vd. 2009).

$$CEQ = \left(\mu - r_f\right) - \frac{\gamma}{2}\sigma^2 \tag{14}$$

Eşitlik 12'de γ, yatırımcının riskten kaçınma katsayısı olarak ifade edilmektedir. Araştırmacılar tarafından genellikle bir değeri verilen katsayı ($\gamma = 1$), farklı değerler de alabilmektedir (Levy, 2017; Raffinot, 2017). Maksimum düşüş (MDD) performans değerlendirme yöntemi yatırımcının karşı karşıya kaldığı kesin kayıp göstergesi olarak tanımlanmaktadır. Bir portföy yatırımının belirli bir zaman dilimindeki zirve noktasından dip noktasına kadar tek ve en büyük düşüşün kümülatif toplamı ile bulunmaktadır. MDD Eşitlik 15'te gösterildiği şekilde hesaplanmaktadır (Magdon-Ismail vd. 2004; Reveiz ve Leon, 2010).

$$MDD_{[0,T]} = min\left(\frac{V_T - V_{max}}{V_{max}}, MDD_{[0,T-1]}\right) \tag{15}$$

Eşitlik 13'te $[0, T]$ zaman periyodu için yatırımın $[0, T-1]$ dönemi sonundaki değeri V_T ile ifade edilirken; yatırımın $[0, T-1]$ dönemi içerisinde ulaştığı maksimum nokta V_{max} ile ifade edilmektedir.

[31] Adjusted Sharpe ratio (ASR), Certainty-Equivalent Return (CEQ), Max Drawdown (MDD), Sum Of Squared Portfolio Weights (SSPW)

[32] Çalışmada portföy optimizasyonu metodlarının karşılaştırılması amacıyla tüm hesaplamalarda risksiz varlık oranı sıfır olarak kabul edilmiştir.

4. VERİ

Çalışmada, makine öğrenmesi ile HRP metodu kullanılarak yapılacak olan analizlerde Borsa İstanbul Tüm Endeksi (BISTTUM), Frankfurt Borsası (DAX Classic All Share) ve Londra Borsası'nda (FTSE All Share) işlem gören hisse senetlerine ait aylık getiriler kullanılmaktadır. Temmuz 2005 ile Haziran 2017 dönemlerinde işlem gören tüm hisse senetlerinin aylık getirileri kullanılarak her bir borsa için Fama ve French (1993) metodolojisi ile her 12 ayda bir güncellenen portföyler üretilmiştir. Buna göre, her bir borsa için portföy ağırlıkları HRP metodu kullanılarak her yıl temmuz ayında güncellenmekte ve belirlenen ağırlıklarla sonraki haziran ayına kadar yatırım yapılmaktadır. Çalışmada kullanılan üç borsaya ait veri setinin tanımlayıcı istatistikleri Tablo 1'de gösterildiği gibidir.

Tablo 1. Tanımlayıcı İstatistikler

Araştırma Dönemi: Temmuz 2005 – Haziran 2017 **Gözlem Sayısı:** 158			
	BIST	**DAX**	**FTSE**
Varlık Sayısı	208	96	410
Maksimum Getiri*	0,1514	0,1520	0,1751
Minimum Getiri*	-0,3246	-0,2437	-0,2178
Ortalama Getiri*	0,0085	0,0054	0,0043
Standart Sapma*	0,0756	0,0485	0,0440
Çarpıklık*	-1,0732	-1,4749	-0,9727
Basıklık*	2,4048	5,3596	5,5026
JB Test	64,37 (0,000)	231,10 (0,000)	208,9 (0,000)

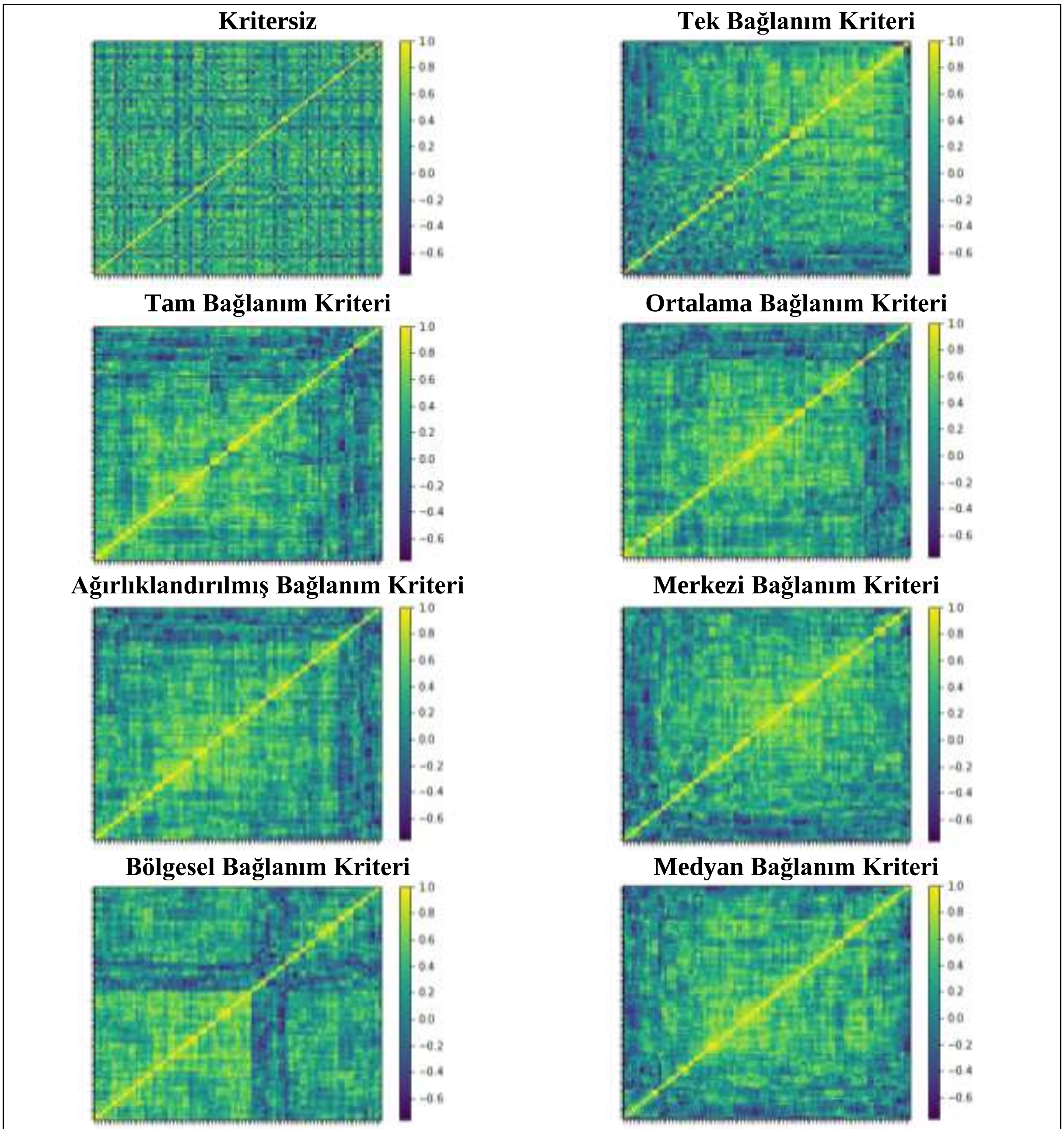
*İşlem gören tüm hisse senetlerinden eşit ağırlık kullanılarak oluşturulan portföy için hesaplanmıştır.
Parantez içerisindeki değerler testin olasılık değerlerini ifade etmektedir

Tablo 1'de yer alan ve ilgili borsadaki tüm hisse senetlerinden eşit ağırlıkla oluşan portföylere ait tanımlayıcı istatistikler incelendiğinde maksimum ve minimum getiriler açısından üç borsanın da yakın değerler aldığı görülmektedir. Ancak, ortalama getiriler açısından Borsa İstanbul portföyünün neredeyse Frankfurt ve Londra portföyleri toplamı kadar ortalama getiri sağladığı görülmektedir. Diğer yandan, yüksek ortalama getiriye sahip olan Borsa İstanbul portföyünün doğru orantılı bir şekilde standart sapmasının da yüksek olduğu göze çarpmaktadır. Jarque ve Bera (1987) tarafından geliştirilen normallik testi sonuçlarına göre her üç borsa portföyü de normal dağılıma uymamaktadır.

5. BULGULAR

Çalışmanın amacı doğrultusunda kullanılan üç farklı piyasa (Borsa İstanbul, Frankfurt Borsası ve Londra Borsası) için Temmuz 2005 ile Haziran 2017 dönemlerinde işlem gören tüm hisse senetlerinin aylık getirileri kullanılarak yapılan analiz sonuçları Tablo 2-4'de özetlenmektedir. Tablolarda yer alan gösterimlerde eşit ağırlık, minimum varyans portföyü (IVP) ve ters-varyans optimizasyonu (GMV) kıyaslama portföyleri olarak sunulurken; Tek, Tam, Ortalama, Ağırlıklandırılmış, Merkezi, Bölgesel ve Medyan algoritmaları kullanılarak oluşturulan portföyler ise makine öğrenmesi portföylerini göstermektedir. Elde edilen sonuçların karşılaştırılabilmesi ve objektif şekilde yorumu için her bir portföyün ortalama, minimum, maksimum getirisi, aralık ölçütü (range), Sharpe (SR) ve Düzeltilmiş Sharpe (ASR) rasyoları, kesin-eşdeğer getiri (CEQ) ölçütü ve son olarak maksimum düşüş (MDD) performans ölçütü kullanılmıştır.

Oluşturulan portföylere ilişkin sonuçların değerlendirilmesinden önce ise, makine öğrenmesi kullanılarak oluşturulan portföylere ilişkin korelasyonların farklı bağlanım kriterlerine göre nasıl bir değişim gösterdiğini açıklamak gerekmektedir. Bu aşamada her piyasa ve her yıl için korelasyon matrisleri oluşturulmaktadır. Oluşturulan bu matrislerin gösterimindeki zorluk nedeniyle ısı grafikleri kullanılarak özetlenebilmektedir. Her piyasa ve her yıl için ısı grafiklerinin sunulması yerine Şekil 1'de yer alan grafikler örnek olarak verilmektedir. Şekil 1'de Frankfurt Borsası için 2006 yılına ait veri kullanılarak Kritersiz, Tek, Tam, Ortalama, Ağırlıklandırılmış, Merkezi, Bölgesel ve Medyan bağlanım kriterlerine göre oluşturulan korelasyon matrisleri yer almaktadır.



Şekil 1. 2006 yılı Frankfurt Borsası için Bağlanım Kriterlerine göre Korelasyon Isı Grafikleri

Şekil 1'de yer alan korelasyon ısı grafikleri incelendiğinde, bağlanım kriterlerinin ya da başka bir deyişle makine öğrenmesi algoritmalarının, verinin kendisinden hareketle korelasyonları nasıl değiştirdiği ve kümeleme yaptığı açıkça görülmektedir. Her kriter kendi metodolojisindeki eşitlikleri kullanarak portföy içindeki varlıkların ağırlıklarını da bu yolla etkileyebilmektedir.

Tablo 2. Borsa İstanbul Portföy Sonuçları

BIST	Ortalama Getiri	Minimum Getiri	Maksimum Getiri	Aralık	SR	ASR	CEQ	MDD
Eşit Ağırlık	0.0073	-0.3246	0.1514	0.4760	13.76	6.82	0.96	1.07
IVP	0.0092	-0.3018	0.1399	0.4417	21.91	103.89	1.21	0.96
GMV	0.0067	-0.4071	0.2754	0.6825	170.96	837028.92	0.89	0.91
Ortalama	0.0084	-0.3006	0.1349	0.4355	15.97	-178.9872	1.11	1.02
Merkezi	0.0062	-0.3392	0.1362	0.4754	11.86	34.8461	0.82	1.16
Tam	0.0090	-0.3442	0.2257	0.5699	1.41	1.1062	1.19	1.05
Medyan	0.0075	-0.3560	0.1747	0.5307	13.95	72.9735	0.99	1.21
Tek	0.0087	-0.2631	0.1475	0.4106	16.29	-210.4189	1.14	1.01
Bölgesel	0.0081	-0.3103	0.2400	0.5504	14.59	-53.3065	1.06	1.16
Ağırlıklan.	0.0062	-0.3525	0.1436	0.4962	11.58	46.1031	0.81	1.07

Tablo 2'de yer alan Borsa İstanbul Tüm Endeksi kullanılarak oluşturulan portföy getirileri analiz sonuçları incelendiğinde, genel olarak en başarılı portföylerin makine öğrenmesi algoritması ile oluşturulanlar olmadığı görülmektedir. Sonuçlar detaylı incelendiğinde, IVP'nin neredeyse tüm ölçütlerde en yüksek getiriye ve en iyi performansa sahip portföy olduğu gözlenmektedir. Ayrıca, makine öğrenmesi kullanılarak oluşturulan portföylerin aralık değerleri incelendiğinde, IVP'ye göre daha geniş bir getiri aralığına sahip olduğu görülmektedir. Sharpe rasyoları incelendiğinde ise, kıyas portföylerinin makine öğrenmesi ile oluşturulan portföylere göre çok daha iyi performans sergilediği ortaya konulmaktadır. Elde edilen bulgular birlikte değerlendirildiğinde ise, Borsa İstanbul için çalışmaya dahil edilen veri aralığında makine öğrenmesi kullanılarak oluşturulan portföylerin başarılı bir performans sergilediğini söylemek mümkün değildir.

Tablo 3. Londra Borsası Portföy Sonuçları

FTSE	Ortalama Getiri	Minimum Getiri	Maksimum Getiri	Aralık	SR	ASR	CEQ	MDD
Eşit Ağırlık	0.0035	-0.2178	0.1751	0.3928	11.71	148.14	0.46	0.93
IVP	0.0049	-0.1578	0.1193	0.2771	25.42	104.46	0.65	0.61
GMV	0.0015	-0.1750	0.1103	0.2853	79.83	5109.32	0.20	0.82
Ortalama	0.0049	-0.1578	0.1193	0.2771	4.80	8.91	0.20	0.86
Merkezi	0.0033	-0.1804	0.1093	0.2897	8.17	-12.28	0.44	0.90
Tam	0.0017	-0.2378	0.2004	0.4382	5.73	26.40	0.23	0.98
Medyan	0.0021	-0.2304	0.1982	0.4286	6.57	31.35	0.28	1.11
Tek	0.0041	-0.1315	0.1112	0.2427	13.77	-171.00	0.54	0.72
Bölgesel	0.0029	-0.1941	0.1493	0.3434	8.44	0.54	0.38	0.90
Ağırlıklan.	0.0026	-0.1934	0.1427	0.3361	7.92	8.69	0.34	0.91

Tablo 3'te yer alan Londra Borsası Tüm Endeksi kullanılarak oluşturulan portföy getirileri analiz sonuçları incelendiğinde, Borsa İstanbul ile benzer şekilde ve makine öğrenmesi ile oluşturulan portföylerin başarılı olduğu söylenememektedir. Bu piyasada da IVP'nin neredeyse tüm performans ölçütlerinde başarılı olarak değerlendirildiği gözlenmektedir. Ancak, ortalama getiri açısından Ortalama Algoritması ile oluşturulan portföyün IVP ile aynı seviyede getiriye sahip olduğu da görülmektedir.

Diğer yandan, Sharpe rasyoları incelendiğinde makine öğrenmesi algoritmaları ile oluşturulan portföylerin performansları kıyas portföylere göre oldukça geride kalmaktadır.

Tablo 4. Frankfurt Borsası Portföy Sonuçları

DAX	Ortalama Getiri	Minimum Getiri	Maksimum Getiri	Aralık	SR	ASR	CEQ	MDD
Eşit Ağırlık	0.0037	-0.2437	0.1520	0.3957	11.74	125.03	0.49	1.02
IVP	0.0011	-0.2427	0.1382	0.3808	2.72	3.97	0.15	1.05
GMV	-0.0006	-0.2369	0.2107	0.4476	-20.29	-1872.60	-0.08	1.08
Ortalama	0.0025	-0.2336	0.1586	0.3923	7.70	33.83	0.33	1.08
Merkezi	0.0049	-0.1843	0.1717	0.3560	13.79	-84.13	0.65	0.96
Tam	0.0029	-0.1992	0.1395	0.3387	8.34	-19.47	0.38	0.99
Medyan	0.0058	-0.2296	0.1560	0.3856	18.16	99.05	0.77	0.95
Tek	0.0070	-0.2446	0.1291	0.3737	20.73	359.12	0.92	0.95
Bölgesel	0.0029	-0.1861	0.1451	0.3312	9.14	-34.62	0.39	0.97
Ağırlıklan.	0.0050	-0.1946	0.1306	0.3252	15.79	-76.26	0.66	0.86

Tablo 4'de yer alan son piyasa Frankfurt Borsası Tüm Endeksi kullanılarak oluşturulan portföy getirileri analiz sonuçları incelendiğinde ise, diğer piyasaların aksine bu piyasada makine öğrenmesi algoritmaları ile oluşturulan portföylerin kıyas portföylere göre başarılı performans sergilediği söylenebilmektedir. Özellikle Medyan ve Tek algoritma kullanılarak oluşturulan portföylerin hem ortalama getiri bazında hem de Sharpe rasyoları bazında en iyi performansı gösterdiği gözlenmektedir. Diğer yandan GMV portföyünün ise diğer piyasalarda genelde ikinci sırada olmasına karşın en kötü performansı sergileyen portföy olduğu sonucuna ulaşılmıştır.

Sonuçlar birlikte değerlendirildiğinde, makine öğrenmesi algoritmalarının zayıf-etkin (BIST) ve güçlü-etkin (FTSE) piyasalarda başarılı performans gösteremezken; orta-etkin (DAX) bir piyasada kıyas portföylerine göre daha başarılı bir performans sergilemesinin nedenleri üzerine odaklanmak gerekmektedir. Tanımlayıcı istatistikler incelendiğinde, DAX, tüm istatistiklerde iki uç noktada olan BIST ve FTSE'nin ortasında değerlere sahiptir. BIST tüm istatistiklerde yüksek standart sapma ve yüksek getiri uç noktasında yer alırken; FTSE aksi uçta yer almaktadır. DAX ise istatistiksel göstergeler açısından diğer iki piyasanın ortasında bulunmaktadır. Sadece ilgili veri aralığında kesintisiz verisi bulunan varlık sayısında, diğer piyasaların altında kalmaktadır. Başka bir deyişle, portföy oluşturmada kullanılabilecek daha az varlık sayısına sahiptir. Varlık sayısının az olmasının daha başarılı performansa etkisi olabileceği düşünülmektedir.

6. SONUÇ

Optimal portföy teorisi, Markowitz tarafından 1952 yılında ortaya atıldığından bu güne yüzlerce araştırmacı ve yatırımcı tarafından test edilmiş ve kullanılmıştır. Söz konusu çalışmalar içerisinde, teorinin başarılı sonuçlar verdiğini ortaya koyanların yanında; teoriye birçok açından eleştiri getiren, eksik bulan ve geliştirmek için adımlar atan araştırmalar da bulunmaktadır. Bu eleştirilerden bir tanesi de optimal portföylerin sıklıkla az sayıda varlık üzerine konsantre olmuş portföyler üretmesi olarak ifade edilmektedir. Konsantrasyon problemi olarak adlandırılan bu durum, temelinde çeşitlendirme olan bir teorinin, çeşitlendirmeye olanak sağlamamasına neden olmaktadır. Literatürde bu probleme "Markowitz'in Laneti" isimlendirmesi yapılmaktadır.

Prado (2016; 2018), bahsedilen konsantrasyon probleminin çözümüne yönelik farklı bir bakış açısı ortaya koymaktadır. Bu amaçla, matematik, grafik teorisi ve makine öğrenmesi kullanarak "Hiyerarşik Risk Paritesi (HRP)" metodu ile portföy optimizasyonu yapılabileceğini, kullandığı simülasyonlar aracılığı ile göstermiştir. HRP, basit anlamda kovaryans matrisinden başlayan bir kümeleme algoritmasına dayanmaktadır. Yöntem, farklı sektörlerdeki ya da türlerdeki finansal varlıklara (örn. hisse senedi, tahvil, gayrimenkul, türev ürünler vs.) yatırım yapmayı amaçlayan bir yatırımcıya, varlıkların uzaklık matrislerini hesaplayarak makine öğrenmesi ile gruplama yapma imkanı sağlamaktadır.

Bu çalışmanın amacı, makine öğrenmesi ile HRP metodu kullanılarak gerçek piyasa verisi ile Prado'nun ortaya attığı tekniğin test edilmesidir. Bu amaçla Avrupa kıtasında karşılaştırma yapılmış ve Borsa İstanbul, Frankfurt Borsası ve Londra Borsası'nın tüm hisselerini içeren endeksler için Temmuz 2005 ile Haziran 2017 dönemlerine ait aylık getiriler kullanılmıştır. Seçilen üç piyasa, etkin piyasa hipotezine göre güçlü-etkin (FTSE), orta-etkin (DAX) ve zayıf-etkin (BIST) piyasa türlerini temsil etmektedir. Analizlerde Fama ve French'in (1993) yatırım metodolojisinden yararlanılmış ve her 12 ayda bir (Temmuz_{t-1} – Haziran_{t}) güncellenen portföyler üretilmiştir. Portföylerin üretilmesinde HRP metodunun yanında kıyaslama yapılabilmesi için Minimum Varyans (IVP) ve Ters-Varyans portföyleri (GMV) de oluşturulmuştur. Yöntemlerin performans testlerini yapabilmek amacıyla Sharpe rasyosu, düzeltilmiş Sharpe rasyosu, kesin-eşdeğer getiri ve maksimum düşüş performans kriterlerinden yararlanılmıştır.

Yapılan analizler sonucunda, makine öğrenmesi algoritmalarının kıyas amaçlı oluşturulan portföylere göre, Borsa İstanbul ve Londra Borsası'nda başarılı sonuçlar göstermediği; ancak Frankfurt Borsası'nda başarılı sonuçlar gösterdiği bulgusuna ulaşılmıştır. Frankfurt Borsası'nın diğerlerinden farkları üzerine odaklanıldığında ise tanımlayıcı istatistiklere göre gerçekten iki uç noktada yer alan diğer piyasaların ortasında değerler aldığı gözlenmiştir. En önemli fark olarak analizlere dahil edilen varlık sayısının diğer piyasalardan daha az olduğu görülmüştür.

Sonuçlar genel olarak değerlendirildiğinde, HRP algoritması ile oluşturulan portföylerin Prado'nun simülasyonlarında konsantrasyon problemini önlediği ve başarılı performans sergilediği iddia edilmiş olmasına rağmen; gerçek piyasa verisi ile test edildiğinde, konsantrasyon problemini çözerken de istikrarsız portföyler oluşmasına neden olduğu söylenebilmektedir. Ancak, gelecekteki çalışmalar açısından gelişmekte olan makine öğrenmesi teknikleri kullanılarak, farklı piyasa ve zaman aralıkları için analizler yapılabilir.

KAYNAKLAR

Bachelier, L. (1900). *Théorie de la spéculation*. Gauthier-Villars.

Bailey, D. H., & Prado, M. L. (2013). An open-source implementation of the critical-line algorithm for portfolio optimization. *Algorithms*, *6*(1), 169-196.

Bailey, D., & Lopez de Prado, M. (2012). Balanced baskets: a new approach to trading and hedging risks. Working Paper. Harvard University-RCC. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2066170

Bar-Joseph, Z., Gifford, D. K., & Jaakkola, T. S. (2001). Fast optimal leaf ordering for hierarchical clustering. *Bioinformatics*, *17*(1), S22-S29.

Bera, A. K., & Park, S. Y. (2008). Optimal portfolio diversification using the maximum entropy principle. *Econometric Reviews*, *27*(4-6), 484-512.

Best, M. J., & Grauer, R. R. (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *The review of financial studies*, *4*(2), 315-342.

Butler, A. (2018). Portfolio Optimization: Simple versus Optimal Methods. Working Paper. Advisor Perspectives. https://www.advisorperspectives.com/articles/2018/10/15/portfolio-optimization-simple-versus-optimal-methods.pdf

Day, W. H., & Edelsbrunner, H. (1984). Efficient algorithms for agglomerative hierarchical clustering methods. *Journal of classification*, *1*(1), 7-24.

DeMiguel, V., Garlappi, L., & Uppal, R. (2007). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?. *The review of Financial studies*, *22*(5), 1915-1953.

DeMiguel, V., Garlappi, L., Uppal, R. (2005). How inefficient is the 1/N asset allocation strategy? Working Paper. London Business School, London, UK. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=785164

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, *33*(1), 3-56.

Frost, P. A., & Savarino, J. E. (1986). An empirical Bayes approach to efficient portfolio selection. *Journal of Financial and Quantitative Analysis*, *21*(3), 293-305.

Hicks, J.R. (1935). A Suggestion for Simplifying the Theory of Money, *Economica*, February:1-1.

Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review/Revue Internationale de Statistique*, 163-172.

Jobson, J. D., & Korkie, B. (1980). Estimation for Markowitz efficient portfolios. *Journal of the American Statistical Association*, *75*(371), 544-554.

Jorion, P. (1985). International portfolio diversification with estimation risk. *Journal of Business*, 58, 259-278.

Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, *21*(3), 279-292.

Jothimani, D. (2017). *Portfolio optimization in Indian market: a study using financial analytics* (Doctoral dissertation). Department of management studies Indian institute of technology, Delhi, India.

Leavens, D.H. (1945). Diversification of Investments, *Trusts and Estates*, *80*(5), 469-473.

Ledoit, O., & Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of empirical finance*, *10*(5), 603-621.

Levy, M. (2017). Measuring Portfolio Performance: Sharpe, Alpha, or the Geometric Mean?. *Journal of Investment Management*, *15*, 1-17.

Liu, W., Liu, Y., Luo, R., & Ding, Y. (2018). Ability Parity Model for Optimal Fund Allocation. Working Paper. Southwestern University of Finance and Economics, China. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3179390

Lopez de Prado, M. (2016). Building diversified portfolios that outperform out-of-sample. Working Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2713516

Magdon-Ismail, M., Atiya, A. F., Pratap, A., & Abu-Mostafa, Y. S. (2004). On the maximum drawdown of a Brownian motion. *Journal of applied probability*, *41*(1), 147-161.

Markowitz, H. (1952). Portfolio Selection, *Journal of Finance, 7*(1), 77-91.

Markowitz, H. (1956). The optimization of a quadratic function subject to linear constraints. *Naval research logistics quarterly*, *3*(1-2), 111-133.

Marschak, J. (1938). Money and the Theory of Assets, *Econometrica*, *6*, 311-325.

Michaud, R. O. (1989). The Markowitz optimization enigma: Is ‘optimized’optimal?. *Financial Analysts Journal*, *45*(1), 31-42.
Michaud, R. O. (1998). *Efficient Asset Management: A Pratical Guide to Stock Portfolio Optimization and Asset Allocation.* Boston: Harvard Business School Press.
Müllner, D. (2011). Modern hierarchical, agglomerative clustering algorithms. Working Paper. Cornell University, New York, USA. https://arxiv.org/abs/1109.2378
Müllner, D. (2013). fastcluster: Fast hierarchical, agglomerative clustering routines for R and Python. *Journal of Statistical Software*, *53*(9), 1-18.
Pennacchi, G. G. (2008). *Theory of asset pricing*. Boston: Pearson/Addison-Wesley.
Pézier, J., & White, A. (2008). The relative merits of alternative investments in passive portfolios. *The Journal of Alternative Investments*, *10*(4), 37.
Prado, M. L. (2018). *Advances in financial machine learning*. John Wiley & Sons.
Raffinot, T. (2016). Hierarchical clustering based asset allocation. Working Paper. PSL Research University, Paris, France. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2840729
Raffinot, T. (2018). The Hierarchical Equal Risk Contribution Portfolio. Working Paper. PSL Research University, Paris, France. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3237540
Reveiz, A., & León, C. (2010). Efficient portfolio optimization in the wealth creation and maximum drawdown space. In *Interest Rate Models, Asset Allocation and Quantitative Techniques for Central Banks and Sovereign Wealth Funds* (pp. 134-157). Palgrave Macmillan, London.
Rokach, L., & Maimon, O. (2005). Clustering methods. In *Data mining and knowledge discovery handbook* (pp. 321-352). Springer, Boston, MA.
Sharpe, W. F. (1966). Mutual fund performance. *The Journal of business*, *39*(1), 119-138.
Williams, J.B. (1938). *The Theory of Investment Value*, Cambridge, MA: Harvard University Press.
Zhao, Y., & Karypis, G. (2002, November). Evaluation of hierarchical clustering algorithms for document datasets. In *Proceedings of the eleventh international conference on Information and knowledge management* (pp. 515-524). ACM.