Haziran, 2026

İSTATİSTİK ALANINDA TEORİ VE UYGULAMALAR

EDİTÖR
DOÇ. DR. DURMUŞ YARIMPABUÇ

DOI:10.5281/zenodo.21118273

İstatistik Alanında Teori ve Uygulamalar

Editör

Doç. Dr. Durmuş Yarımpabuç

İmtiyaz Sahibi
Platanus Publishing®

Editör
Doç. Dr. Durmuş Yarımpabuç
Kapak & Mizanpaj & Sosyal Medya
Platanus Yayın Grubu

Birinci Basım
Haziran, 2026

Yayımcı Sertifika No
45813

ISBN
978-625-6185-65-7

©copyright
Bu kitabın yayım hakkı Platanus Publishing'e aittir. Kaynak gösterilmeden alıntı yapılamaz, izin alınmadan hiçbir yolla çoğaltılamaz.

Adres: Natoyolu Cad. Fahri Korutürk Mah. 157/B, 06480, Mamak, Ankara, Türkiye.

Telefon: +90 312 390 1 118
web: www.platanuspublishing.com
e-mail: platanuskitap@gmail.com

Platanus Publishing®

İÇİNDEKİLER


BÖLÜM 1

Metal Emtia Piyasalarında Doğrusal Piyasa Modeli ile Genelleştirilmiş Toplamsal Modelin Performans Karşılaştırması

Serdar Neslihanoğlu[1]

1. GİRİŞ

Son yıllarda küreselleşme ve dijitalleşmenin hız kazanmasıyla metal emtia piyasalarının finansal piyasalarla bütünleşme seviyesi gün geçtikçe artmaktadır. Metal emtia fiyatları; endüstriyel talep, para politikaları, jeopolitik riskler ve döviz kurundaki dalgalanmalar gibi makroekonomik faktörlerden etkilenmektedir. Bu fiyatların en düşük tahmin hatasıyla modellenmesi ve gelecek tahmini, hem portföy yöneticileri hem de araştırmacılar açısından kritik önem taşımaktadır. Finansal piyasalarda portföy yönetiminin temelinde Markowitz (1952) tarafından geliştirilen Modern Portföy Teorisi (MPT) yer almaktadır. Bu teoriyi temel alan ve Sharpe (1964), Lintner (1965) ile Mossin (1966) tarafından geliştirilen Sermaye Varlıkları Fiyatlandırma Modeli (Capital Asset Pricing Model, CAPM), özellikle uygulama kolaylığı nedeniyle portföy yönetiminde sıklıkla tercih edilmektedir. CAPM ile tutarlı olan ve temel varsayımı piyasa portföyü ile varlık getirileri arasındaki doğrusal ilişkiye ve durağan beta parametresine dayanan Doğrusal Piyasa Modeli (DPM), finansal piyasa analizlerinde yaygın olarak kullanılmakta olup parametreleri sıklıkla En Küçük Kareler (EKK) yöntemiyle tahmin edilmektedir (Mergner ve Bulla, 2008; Choudhry ve Wu, 2009; Neslihanoğlu vd., 2017; Neslihanoğlu ve Aksoy, 2020; Neslihanoğlu ve Paker, 2021).

Araştırmanın temel modeli olan DPM, metal emtia piyasalarındaki portföy ile emtia getirileri arasındaki doğrusal ilişkinin analizinde kullanılmaktadır. Örneğin, Ibrahim ve Jayeola (2018), değerli metal yatırımlarında CAPM çerçevesinde sistematik risk (beta) ile beklenen getiri arasında pozitif bir doğrusal ilişki bulunduğunu göstermiştir. Urom, Chevallier ve Zhu (2020), enerji ve emtia piyasalarında (altın, bakır ve alüminyum dâhil) durağan betanın zaman içinde değişkenlik gösterdiğini ve zamanla değişen betalı rejim-değişimli modellerin statik CAPM'e göre daha iyi performans sergilediğini ortaya koyarak durağan beta varsayımının geçerliliğini ve doğrusal ilişki durumunu sorgulamıştır. Mensi,

[1] Doç. Dr., Eskişehir Osmangazi Üniversitesi, Fen Fakültesi, İstatistik Bölümü
ORCID ID: 0000-0001-8451-8023

Nekhili, Vo ve Kang (2021), değerli ve endüstriyel metal vadeli işlem getirileri arasındaki bağımlılık yapısını farklı piyasa koşulları ve yatırım ufukları altında inceleyerek bu ilişkilerin kantillere ve yatırım ufkuna göre değiştiğini, dolayısıyla doğrusal olmayan ve asimetrik bir yapı sergilediğini göstermiştir. Schischke ve Rathgeber (2025), endüstriyel metal piyasalarında fiyatlar arasındaki bağımlılıkların zaman içinde değiştiğini ve zamanla değişen modellerin, statik kıyaslama modellerine göre bu doğrusal olmayan etkileşimleri daha başarılı biçimde temsil ettiğini ortaya koymuştur. Sonuç olarak literatürde, DPM'nin temel varsayımı olan piyasa portföyü ile varlık getirileri arasındaki doğrusal ilişkinin geçerliliği sorgulanmakta; getiriler arasındaki ilişkinin doğrusal olmayan bir yapı sergilediğini gösteren kanıtlar doğrusal olmayan modellerin önemine işaret etmektedir.

DPM'nin doğrusallık varsayımına esneklik kazandırmak amacıyla öne çıkan yaklaşımlardan biri, Hastie ve Tibshirani (1990) tarafından geliştirilen Genelleştirilmiş Toplamsal Model'dir (GAM). GAM'ın, piyasa portföyü ile varlık getirileri arasındaki ilişkiyi veriden tahmin edilen düzgünleştirici fonksiyonlarla modelleyerek doğrusal olmayan ilişkileri ve aykırı değerlere bağlı dinamikleri açıklayabileceği öngörülmektedir (Wood, 2006). GAM'ın varlık getirilerinin modellenmesinde kullanımına ilişkin literatür özellikle hisse senedi ve enerji piyasalarıyla sınırlı kalmıştır. Örneğin, Neslihanoğlu, Sogiakas, McColl ve Lee (2017), gelişmiş ve gelişmekte olan hisse senedi piyasalarında doğrusal piyasa modelinin doğrusallık varsayımına esneklik kazandırmak amacıyla GAM'ı önermiş ve piyasa ile varlık getirileri arasındaki ilişkide doğrusal olmayan örüntülerin varlığını ortaya koymuştur. Moreno, Figuerola-Ferretti ve Muñoz (2024) ise petrol fiyatı tahmininde GAM ile doğrusal transfer fonksiyonunu birleştiren hibrit bir yaklaşım kullanarak fiyatı belirleyen faktörler arasındaki ilişkilerin doğrusal modellerce tam olarak yakalanamadığını ve önerdikleri modelin tahmin doğruluğunun vadeli fiyatlar dâhil kıyaslama yöntemlerini geçtiğini göstermiştir. Buna karşılık, metal emtia piyasalarındaki portföy ile emtia getirileri arasındaki doğrusal olmayan ilişkinin GAM ile analizi araştırılan literatürde henüz ele alınmamış olup bu konuda önemli bir araştırma boşluğu bulunmaktadır.

Bu araştırma, DPM ile GAM'ın metal emtia piyasalarıyla ilgili literatürde henüz incelenmemiş olan sistematik bir performans karşılaştırmasını gerçekleştirmektedir. Buna ek olarak, GAM'daki düzgünleştirme fonksiyonu metal emtia piyasaları açısından yorumlanarak getirilerin doğrusal olmayan dinamiklere sahip olduğunun ampirik olarak gösterilmesi hedeflenmektedir. Ayrıca bu çerçevede, metal emtia portföyü yönetimi süreçlerinde model seçiminin önemi ile varlık bazında farklılaşan doğrusal olmayan ilişki yapısının dikkate alınmasının gerekliliği vurgulanmaktadır. Bu amaç doğrultusunda, Bloomberg Emtia Endeksi (BCOM) piyasa portföyü ile Alüminyum, Altın, Bakır, Çinko, Gümüş ve Nikel emtialarının Mayıs 2016–Mayıs 2026 tarihleri

arasındaki haftalık endeks fiyatları kullanılarak, Doğrusal Piyasa Modeli (DPM) ile Genelleştirilmiş Toplamsal Model'in (GAM) modelleme ve üç aylık gelecek tahmini performansları karşılaştırılmaktadır.

Araştırmanın geri kalanı şu şekilde düzenlenmiştir: Bölüm 2'de kullanılan emtia piyasaları verileri detaylandırılmakta ve Bölüm 3'te araştırmada kullanılan DPM ve GAM'ın metodolojileri açıklanmaktadır. Araştırma bulgularına Bölüm 4'te, sonuçlarına ise Bölüm 5'te yer verilmektedir.

2. VERİ

Bu araştırmada, Bloomberg Emtia Endeksi (BCOM) kapsamında işlem hacmi en yüksek olan 6 adet endüstriyel ve değerli metal emtianın (Alüminyum, Altın, Bakır, Çinko, Gümüş ve Nikel) Mayıs 2016 ile Mayıs 2026 tarihleri arasındaki haftalık Amerikan Doları (USD) cinsinden endeks fiyatları verisi tercih edilmiştir. Veriler Investing.com (2026) veri bankasından elde edilmiştir. Her bir metal emtianın ve BCOM endeksinin getirisi (R_{it}), t ve t-1 zamanındaki kapanış fiyatlarının (P_{it} ve P_{it-1}) logaritmik farkı alınarak hesaplanmıştır. BCOM ve metal emtiaların endeks kodları ile haftalık logaritmik getirilerinin betimleyici istatistikleri Tablo 1'de verilmiştir.

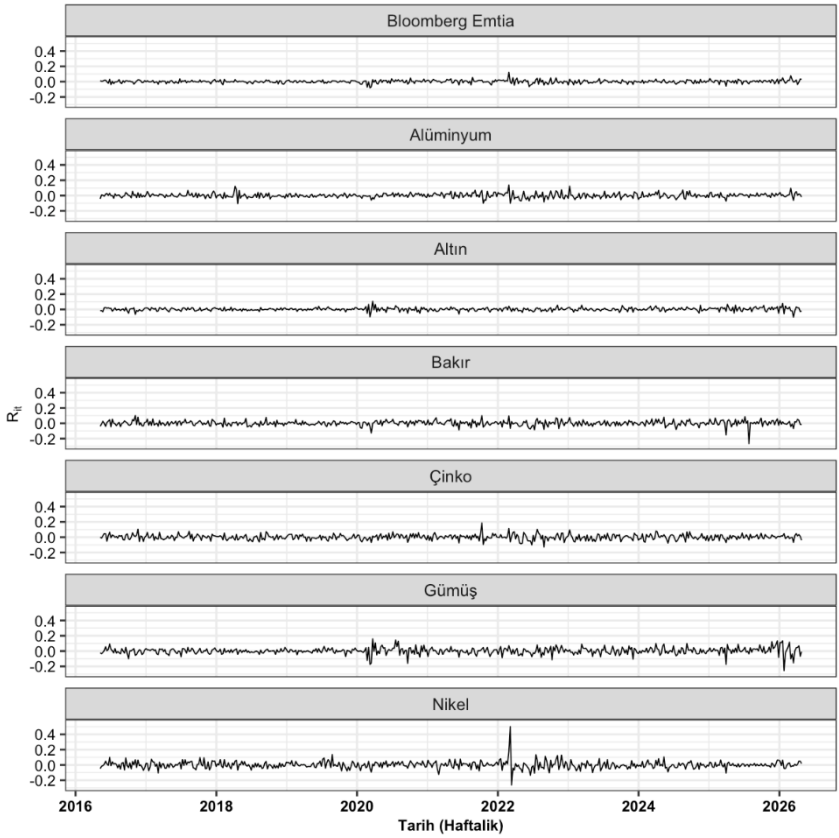
Tablo 1 Betimleyici İstatistikler

Endeks	Endeks Kodu	Ortalama	Standart Sapma	Çarpıklık	Basıklık	Jarque-Bera	Ljung-Box
Bloomberg Emtia	BCOM	0.100	2.019	0.102	6.371	247.63*	22.289
Alüminyum	BCOMAL	0.090	2.942	0.363	5.253	121.63*	27.121
Altın	BCOMGC	0.181	2.162	-0.137	5.904	184.74*	18.478
Bakır	BCOMHG	0.148	3.336	-1.101	11.539	1687.87*	24.834
Çinko	BCOMZS	0.116	3.487	0.178	4.621	59.77*	24.882
Gümüş	BCOMSI	0.215	4.389	-0.534	7.201	407.97*	31.339
Nikel	BCOMNI	0.094	4.808	1.984	26.515	12345.15*	28.918

Not: Jarque-Bera normallik testi (H_0: seri normal dağılıma sahiptir); Ljung-Box otokorelasyon testi (H_0: tüm otokorelasyon katsayıları sıfırdır) için "*" işareti, H_0'ın %95 güven düzeyinde reddedildiğini ifade eder.

Tablo 1'deki betimleyici istatistikler incelendiğinde, BCOM (Bloomberg Emtia) endeksinin ve metal emtiaların ortalama getirilerinin pozitif olduğu görülmekte; bu da araştırma periyodunda yatırımcının finansal kazanç elde ettiğini ifade etmektedir. Standart sapma değerlerine göre, en yüksek dalgalanmaya (volatiliteye) Nikel sahipken en düşük dalgalanmaya BCOM sahiptir. Bu durumda, standart sapma riskin bir ölçüsü olarak ele alındığında, Nikel'in diğer emtialara göre daha riskli olduğu söylenebilir. BCOM, Alüminyum, Çinko ve Nikel pozitif çarpıklık katsayısına sahipken Altın, Bakır ve Gümüş negatif çarpıklık katsayısına sahiptir. Bu durum, söz konusu negatif çarpıklık katsayısına sahip emtiaların getirilerinde sık görülen küçük artışların

yanında ani büyük düşüşlerin de yaşandığına işaret etmektedir. Buna ek olarak, BCOM ve metal emtia getirilerinin leptokurtik olması, getirilerin normal dağılıma göre daha kalın kuyruklara sahip olduğuna ve tüm değişkenler için aşırı sonuç olasılığının daha yüksek olduğuna işaret etmektedir; bu da metal emtia yatırımcıları için aşırı finansal kazanç veya kayıp olasılığının daha yüksek olduğunu göstermektedir. BCOM ve metal emtiaların normalliği, Jarque-Bera testi aracılığıyla %5 anlamlılık düzeyinde reddedilmektedir. Ljung-Box test istatistikleri açısından ise yalnızca Gümüş emtiası için istatistiksel olarak anlamlı bir otokorelasyon tespit edilmiştir. Özetle, bu araştırma verilerinin temel özellikleri pozitif ortalamalar, dalgalanma, asimetri (sola ve sağa çarpıklık) ve leptokurtozdur (kalın kuyruklar). Bu bulgular, BCOM endeksi ile metal emtialar arasındaki doğrusallık varsayımının genişletilmesi gerektiğine işaret etmektedir.



Şekil 1 Araştırma Verileri Zaman Serisi Grafikleri

Şekil 1'de, BCOM endeksi ve 6 metal emtianın (Alüminyum, Altın, Bakır, Çinko, Gümüş ve Nikel) haftalık logaritmik getirilerinin zaman serisi grafikleri gösterilmiştir. Şekil incelendiğinde, söz konusu araştırma verilerine ait serilerin durağan olduğu ve dönemsel olarak dalgalanmalar gösterdiği görülmektedir. BCOM endeksi sınırlı dalgalanmayla en durağan görünüme sahiptir. Nikel'de 2022 yılının başında ani bir dalgalanma görülmüş olup bu durum, Mart 2022'de Londra Metal Borsası'nın (LME) tarihi kısa pozisyon kapama (short squeeze) krizi nedeniyle Nikel piyasasındaki işlemleri geçici olarak durdurmasını yansıtmaktadır. Tablo 1 ve Şekil 1'deki tutarlı sonuçlar, doğrusallık varsayımına dayanan geleneksel DPM'nin bu tür dalgalanma rejimlerini modellemekte yetersiz olabileceğine işaret etmektedir. Bu bulgular, DPM'ye düzgünleştirici fonksiyonlarla esneklik kazandıran GAM'a gereksinim duyulduğunu güçlü biçimde desteklemektedir.

3. METODOLOJİ

3.1. Doğrusal Piyasa Modeli (DPM)

Araştırmanın temel modeli olan Doğrusal Piyasa Modeli (DPM), BCOM endeksi ile metal emtia getirileri arasındaki ilişkinin doğrusal olduğunu varsaymakta ve aşağıdaki şekilde tanımlamaktadır.

$$R_{it} = \alpha_i + \beta_{im} R_{mt} + \varepsilon_{it} \qquad \varepsilon_{it} \sim N(0, \sigma_i^2)\ t = 1 \dots, N \qquad (1)$$

Burada, R_{it}, $i.$ metal emtiasının t zamandaki getirisini; R_{mt} ise BCOM endeksinin t zamandaki getirisini göstermektedir. β_{im}, $i.$ metal emtiasının sistematik riskini ifade etmektedir. ε_{it} ise $i.$ metal emtiasının t zamandaki $N(0, \sigma_i^2)$ koşulunu sağlayan ve birbirinden bağımsız olan hata terimleridir. Bu araştırmada DPM'nin sabit terimi α_i sıfır olarak kabul edilmiş ve β_{im} parametresi En Küçük Kareler (EKK) yöntemiyle tahmin edilmektedir. (Neslihanoğlu ve Aksoy, 2020).

3.2. Genelleştirilmiş Toplamsal Model (GAM)

Genelleştirilmiş Toplamsal Model (GAM; Hastie ve Tibshirani, 1990), BCOM endeksi ve metal emtia getirileri arasındaki ilişkiyi düzgünleştirici fonksiyonlar aracılığıyla modelleyen ve DPM'nin doğrusallık varsayımına esneklik kazandıran istatistiksel bir yöntem olup aşağıdaki şekilde tanımlanmaktadır.

$$R_{it} = \alpha_i + f_i(R_{mt}) + \varepsilon_{it} \qquad \varepsilon_{it} \sim N(0, \sigma_i^2) \qquad (2)$$

Burada, $f_i(R_{mt})$ düzgünleştirici bir fonksiyonu temsil etmektedir. Bu fonksiyon, araştırma verilerinden tahmin edilmekte olup, BCOM endeksi ile metal emtianın getirisi arasındaki doğrusal olmayan ilişkiyi tanımlamaktadır. ε_{it} ise $i.$ metal emtiasının t zamandaki $N(0, \sigma_i^2)$ dağılımına sahip ve birbirinden bağımsız olan hata terimleridir. GAM parametre tahmin prosedürü Wood (2006)

tarafından özetlenmiş ve *R* programlama dilindeki *mgcv* paketi ile tahmin edilmiştir.

3.3. Model Karşılaştırma Kriterleri

Araştırma modelleri olan DPM ve GAM'ın modelleme ve gelecek tahmini performanslarının karşılaştırılmasında kullanılan Ortalama Mutlak Hata (MAE) ve Hata Kareler Ortalaması (MSE) kriterleri aşağıdaki şekilde tanımlanmaktadır.

$$\text{MAE} = \frac{1}{N}\sum_{t=1}^{N}\left|\widehat{R_{it}} - R_{it}\right| \tag{3}$$

$$\text{MSE} = \frac{1}{N}\sum_{t=1}^{N}\left(\widehat{R_{it}} - R_{it}\right)^2 \tag{4}$$

Burada, R_{it} ve $\widehat{R_{it}}$ sırasıyla *i.* metal emtia getirisinin *t* zamanındaki gözlem değerini ve tercih edilen model tarafından üretilen modelleme (ya da gelecek tahmin) tahminini ifade etmektedir. Bu kriterlere göre daha küçük değerlere sahip olan modelin daha iyi bir performans gösterdiği anlaşılmaktadır.

4. BULGULAR

Bu araştırmada modelleme aşamasında tüm veriler kullanılarak DPM ve GAM'ın sonuçları elde edilmiştir. Gelecek tahmini aşamasında ise 8 yıllık kayan pencere genişliği ile gelecek 3 aylık tahmin sonuçları her bir model için ayrı ayrı hesaplanmıştır. DPM ve GAM'ın modelleme ve gelecek tahmini performanslarına ait model karşılaştırma kriterleri Tablo 2'de sunulmuştur.

Tablo 2 DPM ve GAM'ın Modelleme ve Gelecek Tahmini Performanslarının Karşılaştırması

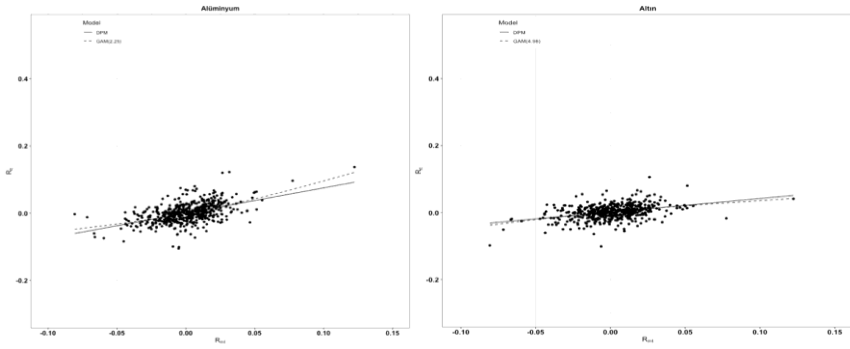
Modelleme				
	MAE ($x10^2$)		MSE ($x10^4$)	
Emtia	DPM	GAM	DPM	GAM
Alüminyum	1.891	1.883	6.341	6.268
Altın	1.483	1.471	4.017	3.915
Bakır	2.084	2.081	8.324	8.240
Çinko	2.387	2.384	9.751	9.750
Gümüş	2.768	2.717	15.110	14.084
Nikel	2.878	2.877	18.987	18.870

Ortalama	*2.249*	*2.236*	*10.422*	*10.188*
Değişim(%)		*% 0.58*		*% 2.24*
Gelecek Tahmini				
	MAE ($x10^2$)		MSE ($x10^4$)	
Emtia	DPM	GAM	DPM	GAM
Alüminyum	2.212	2.128	7.833	7.065
Altın	4.047	3.321	23.575	16.959
Bakır	3.508	3.155	21.603	15.773
Çinko	2.828	2.820	12.430	12.315
Gümüş	6.683	6.076	70.507	59.970
Nikel	2.701	2.691	11.927	10.911
Ortalama	*3.663*	*3.365*	*24.646*	*20.499*
Değişim(%)		*% 8.14*		*% 16.83*

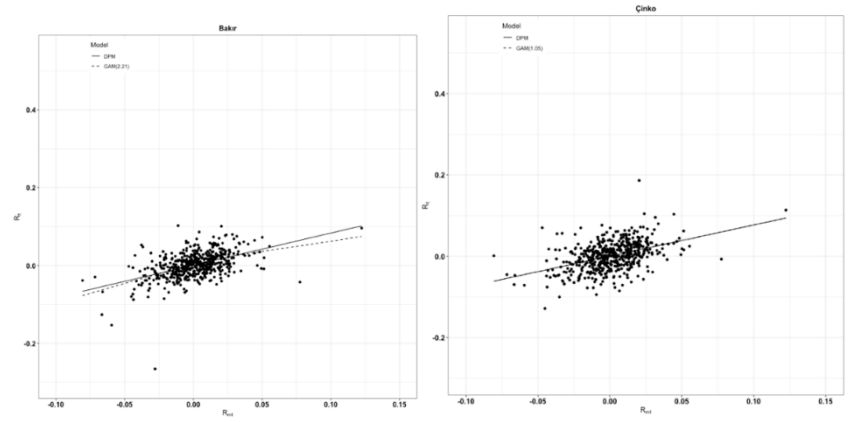
Tablo 2'deki modelleme performansı sonuçları incelendiğinde, GAM modelinin her bir metal emtia getirisi için DPM'ye göre daha iyi performansa sahip olduğu görülmüştür. Buna ek olarak, metal emtia ortalamasına göre GAM'ın DPM'ye göre MAE (MSE) kriterinde %0.58 (%2.24) daha iyi performans gösterdiği belirlenmiştir.

Gelecek tahmini sonuçları incelendiğinde ise, GAM'ın metal emtia getirilerinin gelecek tahmini sürecinde DPM'ye göre yine daha iyi performansa sahip olduğu gözlenmiştir. Bu kapsamda GAM, MAE (MSE) kriterine göre %8.14 (%16.83)'lük bir performans artışı sağlamıştır.

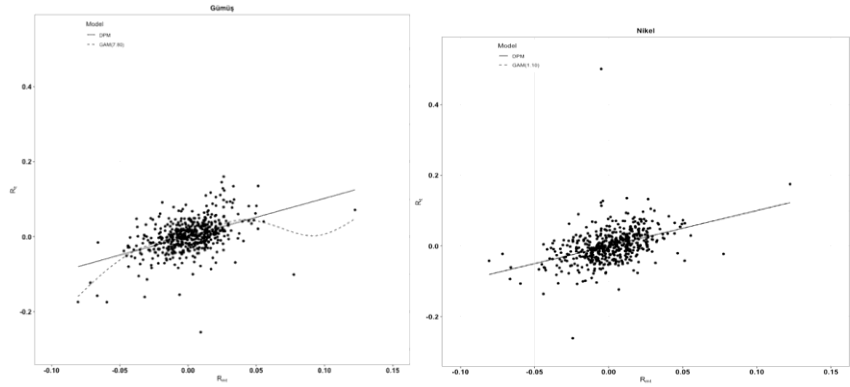
Sonuç olarak, BCOM endeksi ile metal emtia getirisi arasındaki ilişkinin doğrusal bir yapı sergilemediği ve GAM'ın söz konusu metal emtia getirilerinde gözlenen dalgalanma ve aykırı değerlerin modellenmesinde DPM'ye kıyasla daha etkin olduğu sonucuna ulaşılmıştır.



a) Alüminyum b)Altın



c) Bakır ç)Çinko



d) Gümüş e)Nikel

Şekil 2 Metal Emtialarının Haftalık Getirilerine ait Dağılım Grafikleri

BCOM endeksi ile metal emtia getirileri arasındaki ilişkiyi ve DPM ile GAM'ın modelleme performanslarını göstermek amacıyla Şekil 2'de metal emtiaların uyum eğrilerinin dağılım grafikleri oluşturulmuştur. Burada GAM(λ) ifadesi düzgünleştirme parametresini (λ) göstermekte ve GAM'ın esnek yapısını yansıtmaktadır. Bu grafikler incelendiğinde, BCOM endeksi ile her bir metal emtia getirisi arasındaki ilişkinin pozitif yönlü olduğu görülmektedir. Gümüş emtiasının getirisi GAM(λ=7,80) ile en yüksek düzgünleştirme parametresine sahip olup en belirgin doğrusal olmayan ilişkiyi sergilemektedir. Bu durum, Gümüş'ün hem değerli hem de endüstriyel metal özelliğinden kaynaklanan asimetrik piyasa tepkisini yansıtmaktadır. Çinko GAM(λ=1,10) ve Nikel GAM(λ=1,05) emtia getirileri için ise GAM ile DPM'nin benzer ilişki grafiğine sahip oldukları görülmüştür. Bu bulgulara göre, düzgünleştirme parametresi yüksek olan emtialarda (Gümüş, Altın) GAM'ın görsel olarak DPM'den belirgin biçimde ayrıştığı; düşük parametreli emtialarda (Çinko, Nikel) ise iki modelin pratikte eşdeğer olduğu görülmektedir. Buna ek olarak, GAM'ın verideki ani dalgalanmaların modellenmesinde DPM'ye göre daha etkin olduğu açıkça görülmektedir. Sonuç olarak bu durum, metal emtia getirilerinde model seçiminin önemini vurgulamaktadır.

5. SONUÇ

Bu araştırmada, DPM ile DPM'nin doğrusallık varsayımının araştırılması amacıyla tercih edilen GAM modelleri, endüstriyel ve değerli metallerden oluşan 6 emtianın (Alüminyum, Altın, Bakır, Çinko, Gümüş ve Nikel) Mayıs 2016 ile Mayıs 2026 tarihleri arasındaki haftalık verileri kullanılarak modelleme ve 3 aylık gelecek tahmini performansları açısından karşılaştırmalı olarak incelenmiştir.

Araştırma bulguları, GAM'ın metal emtia getirilerinin modellenmesi ve gelecek tahmininde DPM'ye göre daha iyi bir performansa sahip olduğunu ortaya koymuştur. Bu durum, BCOM endeksi ile metal emtia getirileri arasındaki ilişkinin, geleneksel DPM'nin doğrusallık varsayımının aksine, doğrusallık varsayımına esneklik kazandıran GAM'ı desteklediğine işaret etmektedir. Buna ek olarak, model uyum sürecinde GAM'ın, metal emtia getirilerindeki dalgalanma ve aykırı değerlerin modellenmesinde DPM'ye kıyasla daha etkin olduğu belirlenmiştir. Bu bağlamda araştırma, metal emtia getirilerinin açıklanmasında geleneksel piyasa modelinin doğrusal olmayan genişletmelerinin önemini ortaya koymaktadır. Ayrıca GAM'ın gelecek tahminindeki performans üstünlüğü, özellikle Gümüş ve Altın gibi belirgin doğrusal olmayan ilişki sergileyen emtialarda riskten korunma oranlarının ve portföy ağırlıklarının daha ayrıntılı belirlenmesine olanak tanımaktadır.

Sonuç olarak araştırma, esnek modelleme yaklaşımlarının doğrusal modellerce gözden kaçırılan yapısal ilişkileri ortaya çıkarma potansiyeline sahip olduğunu göstermektedir. Gelecek çalışmalarda doğrusallık varsayımına ilişkin

araştırmaların derinleştirilmesi, farklı emtia sınıflarına ve veri setlerine odaklanılması ve zamanla değişen model genişletmelerinin de araştırılması önerilmektedir.

Kaynakça

Choudhry, T. ve Wu, H. (2009). Forecasting ability of GARCH vs Kalman filter method: Evidence from daily UK time-varying beta. *The European Journal of Finance, 15*(4), 437–444.

Hastie, T. ve Tibshirani, R. (1990). *Generalized additive models*. Chapman and Hall.

Ibrahim, O. M. ve Jayeola, D. (2018). On the effect of capital asset pricing model on precious metals and crude oil investments. *Control Science and Engineering, 2*(2), 66–70.

Investing.com. (2026). *Investing.com*. https://www.investing.com/ (Erişim Tarihi: 10.05.2026)

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics, 47*(1), 13–37.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance, 7*(1), 77–91.

Mensi, W., Nekhili, R., Vo, X. V. ve Kang, S. H. (2021). Quantile dependencies between precious and industrial metals futures and portfolio management. *Resources Policy, 73(C)*, 102230.

Mergner, S. ve Bulla, J. (2008). Time-varying beta risk of pan-European industry portfolios: A comparison of alternative modeling techniques. *The European Journal of Finance, 14*(8), 771–802.

Moreno, P., Figuerola-Ferretti, I. ve Muñoz, A. (2024). Forecasting oil prices with non-linear dynamic regression modeling. *Energies, 17*(9), 2182.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, 34*(4), 768–783.

Neslihanoğlu, S. ve Aksoy, T. (2020). BİST'teki ulaştırma sektörü firmalarının verilerinin modellenmesi ve gelecek tahmini için Doğrusal Piyasa Modeli yeterli mi? *Yönetim ve Ekonomi Araştırmaları Dergisi, 18*(4), 54–72.

Neslihanoğlu, S. ve Paker, M. (2021). Beta risklerinin modellenmesi ve tahmini: Türkiye'deki döviz portföyü örneği. *Çankırı Karatekin Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 11*(2), 467–491.

Neslihanoğlu, S., Sogiakas, V., McColl, J. H. ve Lee, D. (2017). Nonlinearities in the CAPM: Evidence from developed and emerging markets. *Journal of Forecasting, 36*(8), 867–897.

R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Schischke, A. ve Rathgeber, A. (2025). Time-varying spillover effects within and between industrial metal markets. *Mineral Economics, 38*(4), 911–939.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*(3), 425–442.

Urom, C., Chevallier, J. ve Zhu, B. (2020). A dynamic conditional regime-switching GARCH CAPM for energy and financial markets. *Energy Economics, 85*, 104577.

Wood, S. (2006). *Generalized additive models: An introduction with R*. Chapman and Hall/CRC.


BÖLÜM 2

A Comparative Multi-Criteria Analysis of PCA, t-SNE, and UMAP for High-Dimensional Data

Ayse Bugatekın[1] & Mehmet Gurcan[2]

1. Introduction.

The number of variables, or in other words, the number of features in a dataset, is referred to as dimensionality. Dimensionality reduction can be simply defined as the process of reducing the number of features in a dataset. The analysis and interpretation of datasets in which the number of features exceeds the number of observations are often difficult and complex. Most statistical analyses are designed for situations where the number of observations is greater than the number of features. In high-dimensional datasets, features are often highly correlated and may therefore contain redundant information. An increased number of features leads to a larger number of model parameters, which in turn makes the model more difficult to interpret. Dimensionality reduction methods are applied to make datasets more suitable for statistical analysis and to simplify their interpretation. The primary goal of dimensionality reduction is to represent the original dataset in a lower-dimensional space with minimal loss of information. Reducing dimensionality by projecting the data onto a lower-dimensional subspace that captures the essential structure of the dataset is generally beneficial [1].

Dimensionality reduction has become an essential tool for analyzing and visualizing high-dimensional data arising in fields such as bioinformatics, machine learning, and pattern recognition. Among the most widely used techniques are linear approaches such as Principal Component Analysis (PCA) and nonlinear manifold learning methods including t-SNE and UMAP.

PCA, extensively studied in classical multivariate statistics, provides a linear projection that maximizes variance and offers computational efficiency and interpretability. However, its inability to capture nonlinear structures limits its effectiveness for complex datasets [2–4]. To overcome these limitations, several nonlinear dimensionality reduction methods have been proposed, including Isomap, Locally Linear Embedding (LLE), and Laplacian Eigenmaps, which aim

[1] Doç. Dr., Department of Science, Firat University, ORCID: 0000-0001-8949-8367
[2] Prof. Dr., Department of Science, Firat University, ORCID: 000-0002-3641-8113

to preserve intrinsic geometric structures of the data manifold [5–8]. Among modern approaches, t-SNE has gained widespread popularity due to its ability to preserve local neighborhood structures by modeling pairwise similarities through probability distributions and minimizing divergence in the embedding space [9]. Despite its effectiveness in revealing cluster structures, t-SNE is known to suffer from limitations such as high computational cost, sensitivity to hyperparameters, and lack of global structure preservation [10–12].

More recently, UMAP has emerged as a powerful alternative grounded in manifold learning and topological data analysis. UMAP aims to preserve both local and global structures and has demonstrated improved scalability and stability compared to t-SNE [13]. Several studies have reported that UMAP provides more interpretable embeddings and better reflects global relationships among data points [14–16].

A growing body of literature has focused on the comparative evaluation of dimensionality reduction techniques. Existing studies typically assess methods based on visualization quality, clustering performance, or application-specific outcomes [17–19]. For instance, comparative analyses in bioinformatics have highlighted the advantages of UMAP in single-cell data visualization, while other works have emphasized the importance of parameter tuning in t-SNE. In addition, recent studies have introduced quantitative evaluation metrics such as trustworthiness, continuity, and clustering indices to assess embedding quality [20, 21]. Despite these contributions, most existing works evaluate dimensionality reduction methods using a limited set of criteria or focus on specific application domains. In particular, the joint assessment of cluster separability, neighborhood preservation, structural distortion, and embedding stability remains relatively underexplored.

In contrast, this study proposes a unified evaluation framework that integrates multiple statistical criteria, including clustering quality, neighborhood preservation, distance distortion, and stability analysis. By combining these aspects, the study provides a more comprehensive understanding of the reliability and effectiveness of PCA, t-SNE, and UMAP in representing high-dimensional data structures.

2. DIMENSIONALITY REDUCTION TECHNIQUES

2.1. Principal Component Analysis (PCA).

Principal Component Analysis (PCA) is a linear dimensionality reduction technique that transforms a high-dimensional dataset into a lower-dimensional space while preserving as much variance as possible. Therefore, PCA can be used to reduce a higher-dimensional matrix m to a smaller-dimensional matrix n [22,

23]. Given a data matrix $X \in R^{nxp}$, the first step involves centering the data by subtracting the mean vector from each observation. In cases where variables have different scales, standardization is also applied.

The covariance matrix of the centered data is then computed as:

$$\sum = \frac{1}{n} X^T X \tag{2.1}$$

PCA is based on the eigenvalue decomposition of the covariance matrix:

$$\sum v_i = \lambda_i v_i \tag{2.2}$$

where λ_i represents the variance explained by the corresponding eigenvector v_i. The eigenvalues are sorted in descending order, and the first k eigenvectors corresponding to the largest eigenvalues are selected to form the projection matrix.

The reduced representation is obtained as:

$$Z = XV_k \tag{2.3}$$

The number of components k is typically chosen such that a predefined proportion of the total variance (e.g., 90% or 95%) is retained:

$$\frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{p} \lambda_i} \geq \tau \tag{2.4}$$

where τ is the desired variance threshold.

2.2. t-Distributed Stochastic Neighbor Embedding

t-distributed stochastic neighbor embedding (t-SNE) is a statistical method for visualizing high-dimensional data by giving each datapoint a location in a two or three-dimensional map. t-SNE, particularly favored for data visualization, makes complex relationships and clustering structures between data more understandable by visualizing them in two or three-dimensional spaces. Developed in 2008 by Laurens van der Maaten and Geoffrey Hinton [9], this method has widespread use in machine learning and data science.

Given data points $x_i, x_j \in R^p$, the conditional similarities in the high-dimensional space are defined using a Gaussian distribution as follows:

$$p_{j/i} = \frac{\exp\left(-\left\|x_i - x_j\right\|^2 / 2\sigma_i^2\right)}{\sum_{k \neq i} \exp\left(-\|x_i - x_k\|^2 / 2\sigma_i^2\right)} \tag{2.5}$$

In this formulation, the parameter σ_i determines the local scale for each data point and is typically adjusted using the perplexity parameter. The symmetric joint probability is then defined as:

$$p_{ij} = \frac{p_{j/i} + p_{i/j}}{2n} \tag{2.6}$$

In the low-dimensional space, where $y_i, y_j \in R^2$, similarities are modeled using a Student-t distribution:

$$q_{ij} = \frac{\left(1 + \left\|y_i - y_j\right\|^2\right)^{-1}}{\sum_{k \neq l} (1 + \|y_k - y_l\|^2)^{-1}} \tag{2.7}$$

The main objective of t-SNE is to minimize the difference between the probability distributions in the high-dimensional and low-dimensional spaces. This is achieved by minimizing the Kullback–Leibler (KL) divergence:

$$KL(P\|Q) = \sum_{i \neq j} p_{ij} log \frac{p_{ij}}{q_{ij}} \tag{2.8}$$

This optimization problem is typically solved using gradient descent methods. t-SNE demonstrates strong performance in preserving local neighborhood structures, while it may be less effective in maintaining the global structure of the data.

2.3. Uniform Manifold Approximation and Projection (UMAP)

UMAP is a dimensionality reduction technique used to reduce high-dimensional datasets to lower-dimensional spaces. UMAP aims to learn the manifold structure of the data and create a low-dimensional representation that preserves this structure. This process requires understanding and preserving both the local neighborhood structures and the general topological properties of the data [24,25]. UMAP essentially has two stages: learning the local neighborhood structure and projecting it into a low-dimensional space. In the high-dimensional space, the neighborhood relationships between data points are modeled using a weighted graph structure. In this graph, the connection weight between two points is defined as:

$$w_{ij} = \exp\left(-\frac{d(x_i, x_j) - \rho_i}{\sigma_i}\right) \tag{2.9}$$

where ρ_i represents the minimum distance parameter for each data point, and σ_i denotes the local scale parameter. This structure characterizes the local manifold geometry of the data.

In the low-dimensional space, similarities are modeled as follows:

$$v_{ij} = \frac{1}{1+a\|y_i - y_j\|^{2b}} \tag{2.10}$$

where the parameters a and b control the shape of the distance function in the low-dimensional space.

UMAP minimizes the discrepancy between the high-dimensional and low-dimensional graph structures using a cross-entropy–based objective function:

$$C = \sum_{i \neq j}[w_{ij} log\, v_{ij} + (1 - w_{ij}) \log (1 - v_{ij})] \tag{2.11}$$

As a result of this optimization process, the resulting embedding is able to preserve both local neighborhood relationships and the global structure of the data in a balanced manner. Due to its computational efficiency and ability to produce stable embeddings, UMAP is widely preferred, especially for large-scale datasets.

3. Analysis of Datasets.

The class labels in the datasets used in this study represent the categorical structures specific to each dataset. In the Iris dataset, the classes setosa, versicolor, and virginica correspond to three different flower species. In the Pima Indians Diabetes dataset, the labels pos (positive) and neg (negative) represent individuals with and without diabetes, respectively. In the Breast Cancer dataset, the classes benign and malignant indicate the nature of the tumor. These class labels were used as reference points in evaluating the low-dimensional representations obtained through dimensionality reduction methods. In the visualizations, they were represented using different colors, enabling a clear visual assessment of class separation. Throughout the experimental process, the same analysis steps were applied to each dataset, and all methods were implemented under identical conditions to ensure comparability of the results. This multi-criteria evaluation approach enabled a comprehensive assessment of dimensionality reduction methods not only in terms of their visual performance, but also their ability to preserve data structure, achieve cluster separation, and produce reliable results. In this way, a fair comparison across datasets and evaluation metrics was ensured.

In this study, the performance of dimensionality reduction methods for obtaining low-dimensional representations of high-dimensional datasets was systematically evaluated using multiple criteria. The primary reason for selecting PCA, t-SNE, and UMAP is that these methods are among the most widely used techniques in the dimensionality reduction literature and represent different methodological approaches. While PCA, as a linear method, focuses on preserving the global structure of the data, t-SNE emphasizes local neighborhood relationships through a nonlinear approach. UMAP, on the other hand, provides a more recent framework that balances both local and global structures while offering computational efficiency.

Therefore, analyzing these three methods together enables a comprehensive evaluation of dimensionality reduction techniques under different data structures. The datasets used in this analysis were selected to reflect varying structural characteristics and consist of the Iris, Pima Indians Diabetes, and Breast Cancer datasets. These datasets represent low-, medium-, and high-dimensional data structures, allowing the investigation of method performance across different levels of complexity. Specifically, the datasets used are Iris (n = 150, p = 4), Pima Indians Diabetes (n ≈ 768, p = 8), and Breast Cancer (n ≈ 569, p = 30).

Prior to the analysis, all datasets were standardized. In this process, each observation was normalized using the mean and standard deviation of the corresponding variable. In other words, each observation x_{ij} was standardized using the transformation $x_{ij}^{*} = \frac{x_{ij}-\mu_j}{\sigma_j}$ where μ_j and σ_j denote the mean and standard deviation of the *j-th* variable, respectively. This transformation ensures that all variables have zero mean and unit variance, thereby eliminating the influence of differences in scale across variables.

Let $X \in R^{nxp}$ denote the standardized data matrix. Each dimensionality reduction method was applied to this matrix, and the data points were projected into a two-dimensional space. This process can be expressed mathematically as a transformation $f: R^{p} \rightarrow R^{2}$. In all analyses, the output dimensionality was fixed to *d=2* in order to represent the data points in a two-dimensional space. For PCA, this was achieved by selecting the eigenvectors corresponding to the two largest eigenvalues obtained from the eigenvalue decomposition of the covariance matrix and projecting the data matrix onto these components. In contrast, t-SNE and UMAP do not perform a direct projection; instead, they learn a two-dimensional embedding space that preserves neighborhood relationships in the high-dimensional space. Therefore, the two-dimensional representations in t-SNE and UMAP are obtained through an optimization process that determines the coordinates of the data points.

3.1. Visual Analysis (Examination of Low-Dimensional Representations)

To further assess the performance of the dimensionality reduction methods, the resulting low-dimensional embeddings were visually examined for each dataset. The analysis focuses on how effectively the methods separate classes, preserve neighborhood structures, and reflect the overall data distribution in the embedding space.

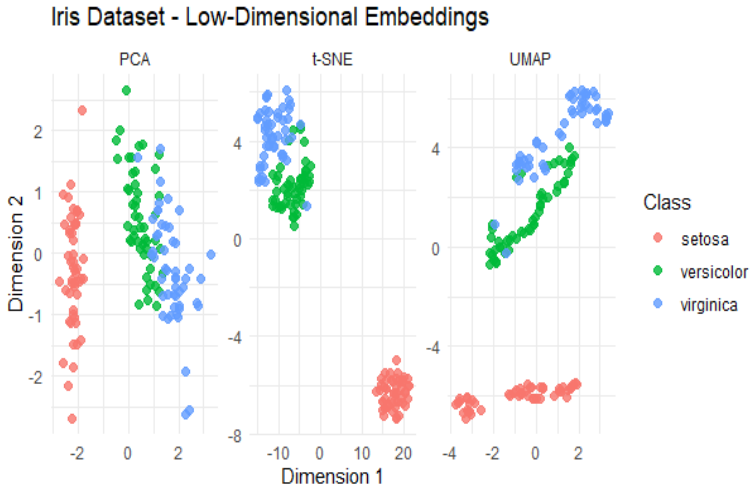


Figure 1. Iris Dataset – Low-Dimensional Embeddings

When the low-dimensional representations obtained for the Iris dataset in Figure 1 are examined, it is observed that all methods are able to separate the classes to a certain extent. Since PCA provides a linear projection, it clearly separates the setosa class from the others, while a partial overlap is observed between the versicolor and virginica classes. In contrast, t-SNE better preserves local neighborhood relationships and produces well-defined clusters for all three classes. UMAP also achieves strong cluster separation similar to t-SNE; however, it provides a more balanced and organized distribution, preserving both local and, to some extent, global structures. These results indicate that nonlinear methods are more effective for this dataset.

Figure 2. Pima Dataset – Low-Dimensional Embeddings

When the embeddings obtained for the Pima Indians Diabetes dataset in Figure 2 are examined, it is observed that the data structure is more complex and that clear class separation is difficult to achieve. Since PCA is based on a linear structure, it fails to produce a distinct separation between classes, resulting in a substantial overlap of data points. Although t-SNE improves class separation in certain regions by preserving local neighborhood relationships, a significant degree of overlap between clusters remains overall. UMAP exhibits a similar pattern to t-SNE and produces a more balanced distribution; however, due to the inherent nature of the dataset, clear separation between classes is still not achieved. This suggests that the dataset intrinsically has low separability.

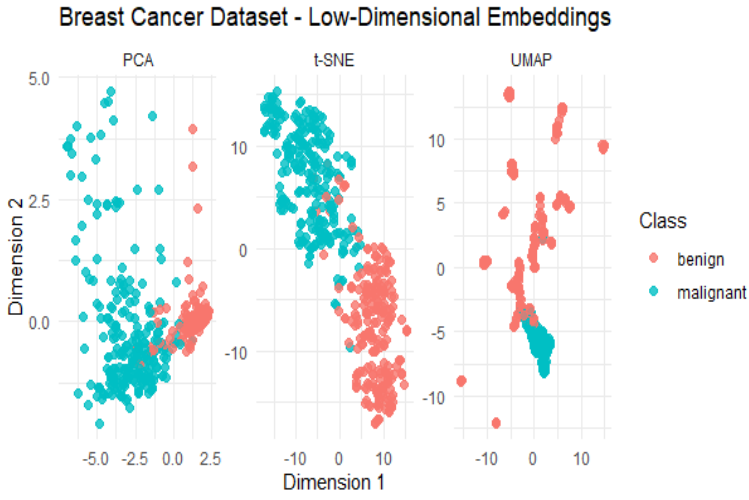


Figure 3. Breast Cancer Dataset – Low-Dimensional Embeddings

The results obtained for the Breast Cancer dataset in Figure 3 reveal the performance differences among dimensionality reduction methods more clearly. While PCA performs well in preserving the global structure, it leads to some overlap between class points. In contrast, t-SNE significantly improves class separation by representing the data points as more compact and well-separated clusters. UMAP not only enhances cluster separation but also provides a more regular and stable distribution. In particular, the more homogeneous and balanced structure produced by UMAP indicates that it better preserves both local and global relationships.

Instead of relying on a single metric, four different performance criteria were jointly used to evaluate the obtained low-dimensional representations. First, the Silhouette coefficient was computed to measure cluster separation. This metric assesses how well each observation is assigned to its own cluster by comparing its proximity to points within the same cluster against its distance to the nearest neighboring cluster.

Second, the trustworthiness metric was employed to evaluate the extent to which neighborhood relationships in the high-dimensional space are preserved in the low-dimensional embedding. This measure penalizes points that appear as neighbors in the low-dimensional space but are not neighbors in the original space, thereby providing insight into local structure preservation.

Third, distance preservation (distortion) analysis was conducted to assess how well the global structure of the data is maintained. In this analysis, the correlation between pairwise distances in the high-dimensional space and those in the low-

dimensional embedding was calculated and used to compare the methods in terms of global structure preservation.

Finally, stability analysis was performed to evaluate the reliability of the methods, particularly those involving stochastic components. In this context, each method was executed multiple times with different random initializations, and the consistency of the resulting distance structures was assessed by computing the correlation between runs.

3.2. Multi-Criteria Evaluation

3.2.1. Cluster Separation Analysis (Silhouette Score)

To evaluate how well dimensionality reduction methods preserve class separation, the Silhouette coefficient was computed for the two-dimensional embeddings obtained using PCA, t-SNE, and UMAP. The analysis was conducted on three datasets with different structural characteristics: Iris, Pima Indians Diabetes, and Breast Cancer. The Silhouette score measures the quality of cluster separation by comparing the proximity of each observation to its own cluster with its distance to other clusters, where higher values indicate better-separated clusters. The obtained results are presented in Table 1.

Table 1. Cluster separation performance based on silhouette scores

Dataset	PCA	t-SNE	UMAP
Iris	0.401	**0.533**	0.463
Pima	**0.117**	0.093	0.081
Cancer	**0.702**	0.560	0.298

An examination of Table 1 shows that the performance of the methods varies considerably across datasets. For the Iris dataset, t-SNE achieves the highest Silhouette score (0.533), indicating that it better captures the nonlinear structure of the data by preserving local neighborhood relationships. UMAP also performs well (0.463), while PCA yields a lower score (0.401) due to its linear nature. In contrast, PCA achieves the best performance on the Breast Cancer dataset (0.702), suggesting that the data structure is largely linear, whereas t-SNE and UMAP produce lower scores (0.560 and 0.298, respectively). For the Pima Indians Diabetes dataset, all methods yield low Silhouette scores, indicating weak class separability and a complex or overlapping data structure.

Overall, the results demonstrate that no single method consistently outperforms others across all datasets. PCA is more effective for linearly structured data, whereas nonlinear methods such as t-SNE and UMAP perform better on more complex data. These findings highlight that the effectiveness of

dimensionality reduction methods strongly depends on the underlying data structure, and the choice of method should be made accordingly.

In Figure 4, the cluster separation performance (Silhouette score) of PCA, t-SNE, and UMAP is compared across datasets. The results show that PCA achieves the best performance on the Breast Cancer dataset, while t-SNE provides the highest separation on the Iris dataset. In contrast, all methods yield low scores on the Pima dataset, indicating weak class separability and a more complex data structure.

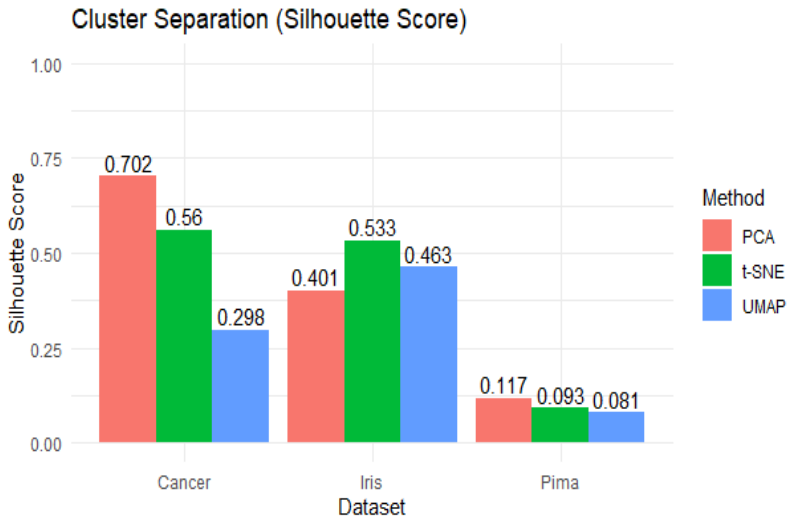


Figure 4. Cluster Separation Performance (Silhouette Score)

3.2.2. Neighborhood Preservation Analysis (Trustworthiness)

To assess the ability of dimensionality reduction methods to maintain neighborhood relationships from the high-dimensional space, the trustworthiness metric was utilized. This measure takes values between 0 and 1, where higher values indicate a better preservation of local structures in the low-dimensional embedding. The comparison was carried out on the Iris, Pima Indians Diabetes, and Breast Cancer datasets to evaluate the local structure preservation performance of PCA, t-SNE, and UMAP. The results are reported in Table 2.

Table 2. Neighborhood preservation performance based on trustworthiness

Dataset	PCA	t-SNE	UMAP
Iris	0.974	**0.991**	0.980
Pima	0.811	**0.979**	0.954
Cancer	0.904	**0.975**	0.962

According to Table 2, t-SNE achieves the highest trustworthiness scores across all datasets, indicating its strong ability to preserve local neighborhood relationships. UMAP also performs well but yields slightly lower values compared to t-SNE, while PCA consistently produces lower scores, particularly on the Pima dataset. At the dataset level, all methods show high performance on the Iris dataset, whereas nonlinear methods clearly outperform PCA on the Pima dataset. For the Breast Cancer dataset, all methods perform well, with t-SNE again leading.

Figure 5 visually supports these findings by illustrating the relative differences in trustworthiness scores across methods and datasets. The figure clearly shows the dominance of t-SNE in all cases, followed by UMAP, while PCA remains lower, especially for the Pima dataset. The consistency of results across datasets further emphasizes the effectiveness of nonlinear methods in preserving local structures.

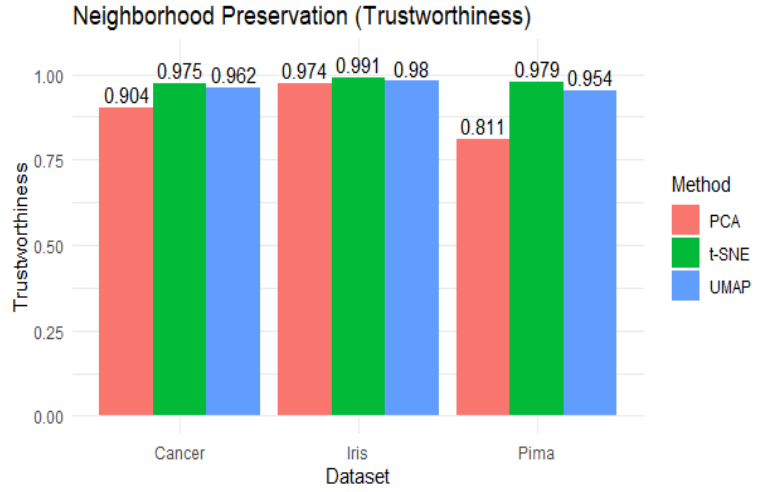


Figure 5. Neighborhood Preservation Performance (Trustworthiness)

3.2.3. Distance Preservation Analysis (Distortion).

To evaluate how well dimensionality reduction methods preserve distance structures, distortion analysis was performed. This involved computing the correlation between pairwise distances in the high-dimensional space and those in the low-dimensional embedding. Values closer to 1 indicate better preservation of the original distance relationships. The analysis was conducted on the Iris, Pima Indians Diabetes, and Breast Cancer datasets using PCA, t-SNE, and UMAP. The results are presented in Table 3.

Table 3. Distance preservation performance based on distortion

Dataset	PCA	t-SNE	UMAP
Iris	**0.995**	0.881	0.914
Pima	**0.793**	0.621	0.566
Cancer	**0.968**	0.785	0.109

According to Table 3, PCA achieves the highest distance preservation scores across all datasets, confirming its effectiveness in maintaining global data structure due to its linear nature. This is particularly evident for the Iris (0.995) and Breast Cancer (0.968) datasets. In contrast, t-SNE and UMAP show lower performance in preserving global distances. While t-SNE provides moderate results for the Iris dataset (0.881), its performance decreases for the Pima dataset (0.621). UMAP yields especially low values for the Breast Cancer dataset (0.109), indicating a weaker preservation of global structure. Overall, the results highlight the advantage of linear methods in maintaining distance relationships.

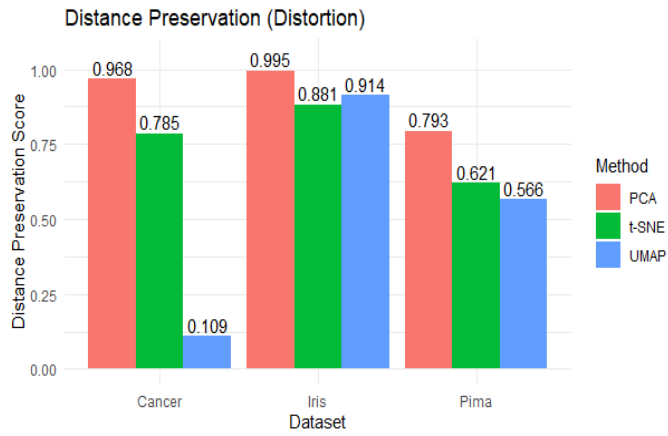


Figure 6. Neighborhood Preservation Performance (Trustworthiness)

Figure 6 visually illustrates these differences in distance preservation performance. The dominance of PCA across all datasets is clearly observed, while t-SNE and UMAP display comparatively lower values. The figure also highlights the consistently low scores for the Pima dataset, reflecting its complex structure. These patterns support the conclusion that nonlinear methods tend to sacrifice global distance preservation in favor of capturing local structures.

3.2.4. StabilityAnalysis.

To evaluate the reliability of dimensionality reduction methods, stability analysis was conducted. In this analysis, each method was executed 20 times on the same dataset using different random initializations, and the correlation between the resulting distance structures was computed. Values close to 1 indicate that the method produces consistent results across multiple runs and is therefore highly reliable. The results are presented in Table 4.

Table 4. Stability performance

Dataset	t-SNE	UMAP
Iris	**0.984**	0.942
Pima	0.843	**0.951**
Cancer	**0.972**	0.965

According to Table 4, both t-SNE and UMAP generally produce high stability values, indicating that they generate consistent results across multiple runs. However, performance varies depending on the dataset. For the Iris dataset, t-SNE achieves the highest stability (0.984), while UMAP also performs well (0.942). In contrast, UMAP shows higher stability for the Pima dataset (0.951), whereas t-SNE yields a lower value (0.843). For the Breast Cancer dataset, both methods exhibit similarly high stability (t-SNE: 0.972, UMAP: 0.965). Overall, these results suggest that both methods are reliable, although their relative performance depends on the dataset. It should also be noted that stability analysis is not applicable to PCA, as it is a deterministic method that produces identical results across runs.

Stability analysis is not applicable to PCA since it is a deterministic method that produces identical results across multiple runs. Therefore, stability is evaluated only for stochastic methods such as t-SNE and UMAP.

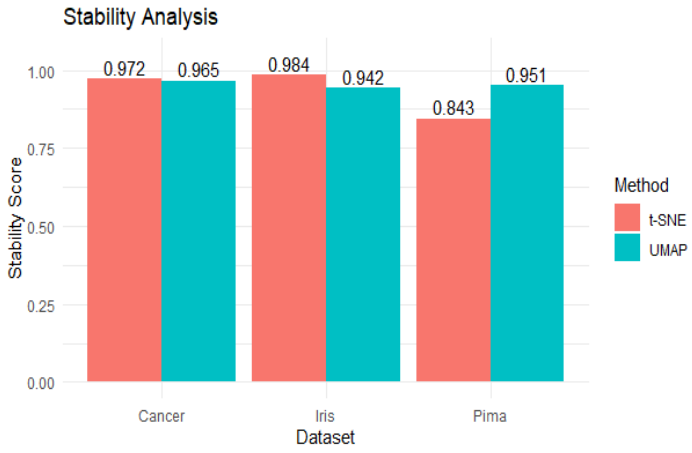


Figure 7. Neighborhood Preservation Performance (Trustworthiness)

Figure 7 visually confirms these findings by illustrating the stability scores of t-SNE and UMAP across datasets. The figure highlights the consistently high stability of both methods, while also revealing dataset-specific differences. In particular, the higher stability of t-SNE for the Iris dataset and the stronger performance of UMAP for the Pima dataset are clearly observable. These patterns emphasize the importance of stability as a key criterion when evaluating stochastic dimensionality reduction methods.

Conclusion

In this study, PCA, t-SNE, and UMAP were systematically evaluated using multiple statistical criteria, including cluster separation, neighborhood preservation, distance preservation, and stability. The findings clearly demonstrate that the performance of dimensionality reduction methods largely depends on the structural characteristics of the dataset.

PCA was found to be particularly effective in preserving global distance relationships and performing well on linearly structured data. In contrast, t-SNE shows superior performance in cluster separation and local neighborhood preservation, making it more suitable for complex and nonlinear data structures. UMAP provides a balanced performance by ensuring both local structure preservation and high stability.

These results indicate that relying on a single performance metric may lead to misleading conclusions. In this context, the proposed multi-criteria evaluation approach offers a more comprehensive and reliable assessment by considering multiple aspects of dimensionality reduction methods simultaneously.

References

[1] K. P. Murphy, Machine Learning: A Probabilistic Perspective. Cambridge, MA, USA: MIT Press, 2012.

[2] I. T. Jolliffe, Principal Component Analysis, 2nd ed. New York, NY, USA: Springer, 2002.

[3] K. Pearson, "On lines and planes of closest fit to systems of points in space," Philosophical Magazine, vol. 2, pp. 559–572, 1901.

[4] H. Hotelling, "Analysis of a complex of statistical variables into principal components," Journal of Educational Psychology, vol. 24, pp. 417–441, 1933.

[5] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," Science, vol. 290, no. 5500, pp. 2319–2323, 2000.

[6] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," Science, vol. 290, no. 5500, pp. 2323–2326, 2000.

[7] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," Neural Computation, vol. 15, no. 6, pp. 1373–1396, 2003.

[8] L. van der Maaten, E. Postma, and J. van den Herik, "Dimensionality reduction: A comparative review," Journal of Machine Learning Research, vol. 10, pp. 66–71, 2009.

[9] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," Journal of Machine Learning Research, vol. 9, pp. 2579–2605, 2008.

[10] G. C. Linderman and S. Steinerberger, "Clustering with t-SNE, provably," SIAM Journal on Mathematics of Data Science, vol. 1, no. 2, pp. 313–332, 2019.

[11] M. Wattenberg, F. Viégas, and I. Johnson, "How to use t-SNE effectively," Distill, 2016.

[12] D. Kobak and P. Berens, "The art of using t-SNE for single-cell transcriptomics," Nature Communications, vol. 10, no. 1, 2019.

[13] L. McInnes, J. Healy, and J. Melville, "UMAP: Uniform manifold approximation and projection for dimension reduction," arXiv:1802.03426, 2018.

[14] E. Becht et al., "Dimensionality reduction for visualizing single-cell data using UMAP," Nature Biotechnology, vol. 37, pp. 38–44, 2019.

[15] M. Diaz-Papkovich, L. Anderson-Trocmé, and S. Gravel, "UMAP reveals cryptic population structure and phenotype heterogeneity," Nature Communications, vol. 10, 2019.

[16] J. He, L. Li, and Y. Xu, “Comparative analysis of PCA, t-SNE, and UMAP in machine learning,” IEEE Access, vol. 11, 2023.

[17] R. M. Rauber, A. X. Falcão, and A. C. Telea, “Visualizing the hidden structure of data,” IEEE Transactions on Visualization and Computer Graphics, vol. 23, no. 1, pp. 101–110, 2017.

[18] A. Espadoto et al., “On the evaluation of dimensionality reduction techniques,” IEEE Transactions on Visualization and Computer Graphics, vol. 27, no. 2, pp. 1179–1190, 2021.

[19] A. Narayan et al., “Assessing single-cell transcriptomic variability through embedding techniques,” Nature Methods, vol. 18, pp. 45–52, 2021.

[20] J. Venna and S. Kaski, “Neighborhood preservation in nonlinear projection methods,” in Proc. ICANN, 2001, pp. 485–491.

[21] J. Venna et al., “Information retrieval perspective to nonlinear dimensionality reduction,” Journal of Machine Learning Research, vol. 11, pp. 451–490, 2010.

[22] C. Narkhede and U. Nagaraj, “Comparative study of PCA and LDA,” Pune, India, 2013.

[23] R. Ramachandran, G. Ravichandran, and A. Raveendran, “Evaluation of dimensionality reduction techniques for big data,” in Proc. ICCCMC, 2020, pp. 226–231.

[24] L. McInnes, J. Healy, and J. Melville, “UMAP: Uniform manifold approximation and projection for dimension reduction,” arXiv:1802.03426, 2018.

[25] E. Becht et al., “Dimensionality reduction for visualizing single-cell data using UMAP,” Nature Biotechnology, vol. 37, no. 1, pp. 38–44, 2019. bu kaynakları APA 6.0 a göre düzenle