

TOPSIS VE K-MEDOIDS YÖNTEMLERİYLE KARAR VERME: R PROGRAMLAMA DİLİ İLE MÜŞTERİ GERİ KAZANMA UYGULAMASI**Dr. Öğr. Üyesi M. Fevzi ESEN*** **Dr. Öğr. Üyesi Emrah BİLGİÇ*** **ÖZET**

Büyük işletmeler, kendisinden ayrılan ve rakip işletmelere geçen eski müşterilerine yönelik geri kazanma faaliyetleri düzenleyeceği zaman, kârlılığa en çok katkı yapan müşterileri tercih etmektedir. Bu seçim, çoğu zaman maliyet ve zaman kısıtları altında birçok değişkenin aynı anda değerlendirilmesini gerektirmektedir. TOPSIS yöntemi, çok kriterli problemler için oluşturulmuş bir karar verme tekniği olup, tedarikçi ve kuruluş yeri seçimi, üretim sistemleri, pazarlama yönetimi, sağlık, güvenlik ve çevre yönetimi gibi geniş bir uygulama alanında kullanılmaktadır. Bu çalışmada, bir telekomünikasyon işletmesinin rakip firmalara geçen 1145 müşterisine ait ses, veri ve katma değerli servilerle alakalı kullanım bilgileri TOPSIS yöntemi ile analiz edilmiştir. Elde edilen TOPSIS skorları işletme kârlılığına katkı sağlayan en değerli müşterileri tespit etmek amacıyla, k-medoids (k-ortaylar) kümeleme algoritması kullanılarak dört segmente ayrılmıştır. Çalışma sonucunda, dört segmente ayrılan müşterilerden üçüncü segmentte yer alan toplam 37 müşteri altın müşteri olarak tespit edilmiş ve geri kazanma faaliyetlerine bu müşterilerden başlanması gerektiği ortaya çıkmıştır.

Anahtar Kelimeler: TOPSIS, K-Medoids Algoritması, Müşteri Geri Kazanma

Jel Sınıflandırması: M2, M3, C38

DECISION MAKING WITH TOPSIS AND K-MEDOIDS: AN APPLICATION ON CUSTOMER RECOVERY WITH R PROGRAMMING LANGUAGE**ABSTRACT**

Large companies prefer the most profitable customers when they try to recover their old customers who left the company and acquired by the competitors. This choice requires many variables to be evaluated simultaneously under the constraints of cost and time. TOPSIS method is a decision-making technique for multi-criteria problems and is used in a wide range of applications such as: supplier and facility location selection, production systems, enterprise resource planning and marketing

* İstanbul Medeniyet Üniversitesi, Turizm Fakültesi, Turizm İşletmeciliği Bölümü, 34862. Kartal, İstanbul, Türkiye, e-mail: fevzi.esen@medeniyet.edu.tr

* Muş Alparslan Üniversitesi, İ.İ.B.F., Sayısal Yöntemler Bölümü, 49250 Muş, Türkiye, e-mail: kyildiz@bandirma.edu.tr

management, health, safety and environmental management problems etc. In this study, data of 1145 customers such as usage of voice, data service and other value added services usage are analyzed by TOPSIS method. The customers analyzed in this study are churned by a competitor of a telecommunication company. In order to determine the most valuable customers to company's profit, TOPSIS scores are clustered into four segments by k-medoids clustering algorithm. As a result of the analysis, customers are segmented into four groups and totally 37 customers which were placed in the third segment identified as golden customers, thus the company should start churn activities with the customers placing in this segment.

Key Words: TOPSIS, K-Medoids Algorithm, Customer Recovery

Jel Codes: M2, M3, C38.

1. GİRİŞ

Karar verme, en genel anlamda iki veya daha fazla alternatif arasından seçim yapma süreci olarak tanımlanmaktadır (Duncan, 1978). Bir işletme probleminin çözümünde çeşitli seçenekler arasından faydayı maksimize edecek olanın seçilmesi, karar verme sürecinin nihai amacıdır. Kabul görmeyen ve seçim dışı kalan diğer alternatifler ise işletmeler için fırsat maliyeti doğurmaktadır. Bu açıdan karar vericinin birçok alternatif arasından en uygun olan çözümü tercih etmesi gerekmektedir.

Çok kriterli karar verme yöntemlerinde, değerlendirmeye alınabilecek alternatifler karar vericiler tarafından çözüm sürecinin başında bilinmektedir. Amaç, kullanılacak olan matematiksel modellerin çözülerek alternatifler arasından en iyi seçimin yapılmasıdır. AHP, ANP, Macbeth, Promethee, Electre gibi çok kriterli karar verme yöntemlerinden farklı olarak, TOPSIS yönteminde ideal çözüm için seçilen alternatifin fayda kriterini maksimize eden çözüme uzaklığının en kısa mesafede, maliyet kriterini minimize eden çözüme ise (negatif ideal) en uzak mesafede olması gerektiği hususu ön plana çıkmaktadır (Hwang ve Yoon, 1981).

Kümeleme analizi, özellikle pazarlama literatüründeki müşteri veya pazar segmentasyonunda etkin bir şekilde kullanılmaktadır. Segmentasyondaki amaç, benzer özelliklere sahip gözlemlerin yer aldığı gruplar içerisinde homojenliğin sağlanması, farklı gruplar arasındaki ayrışmanın ise maksimum derecede gerçekleştirilmesini sağlamaktır. Heterojen niteliklere sahip çok boyutlu veri setlerinde kullanılabilirlik özelliği ve klasik kümeleme yaklaşımlarına kıyasla daha iyi grup içi benzerlik üretebilen özelliği sayesinde k-medoids (k-ortaylar) algoritması sağlam (robust) bir yöntem olup, oluşturulan kümeler için veri setinin özelliğini etkili bir şekilde yansıtmaktadır (Harikumar ve Pv, 2015).

Telekomünikasyon sektöründe müşterilerin ses, veri, sms gibi temel servis ve müzik, içerik, logo gibi katma değerli servis kullanımlarında heterojen bir örüntü mevcut olup, özellikle müşteri segmentlerinin tespit edilerek bunlara yönelik pazarlama stratejilerinin oluşturulması önem arz etmektedir (Delloite, 2015; Bond ve Shatz, 2016). Büyük, heterojen, yapılandırılmış ve yapılandırılmamış müşteri verilerinin analizi, yaşam boyu değeri yüksek olan sadık müşterilerin tespit

edilmesi, farklı demografik düzeyde, farklı ihtiyaçları olan müşteriler için özel ürün / hizmet sunumuyla müşteri kayıplarının engellenmesi ve uzun - kısa dönemli müşterilerin veya potansiyel müşteri gruplarının tespit edilmesini sağlamaktadır. Bu açıdan, müşterilerin birbirinden farklı nitelik (demografik nitelikler, kullanılan servis türü vb.) ve nicelikteki (kullanım miktarı ve sıklığı vb.) kullanım davranışlarının benzer gruplara ayrılması, yani segmentasyonu, işletmelerin uygun pazarlama stratejileri geliştirebilmesi açısından önemli bir rol oynamaktadır. Müşteri segmentasyonu ile, nominal, kategorik, aralık – oran veya karma niteliğe sahip değişkenleri içeren müşteri veri setlerinde benzer gözlem gruplarının tespiti ve bu gruplarda yer alan müşterilerin karakteristiklerinin belirlenmesi amaçlanmaktadır (Saglam vd., 2006; Teichert vd., 2008).

Bu çalışmada, bir telekomünikasyon firmasının müşterilerine ait ses, veri ve diğer katma değerli servis kullanımlarına ilişkin veriler, alanlarında uzman kişilerin görüşleri alınarak TOPSIS yöntemi için ağırlıklandırılarak karar matrisi oluşturulmuş olup, müşterilere ilişkin TOPSIS skorları elde edilmiştir. Müşteri segmentasyonu gerçekleştirirken ise, uzman görüşleri doğrultusunda telekomünikasyon sektöründe müşterinin kıymet derecesini belirleme açısından önemli olan ses, veri hizmetleri ve katma değerli servis kullanımı değişkenlerine ilişkin elde edilen TOPSIS skorları, klasik kümeleme yaklaşımlarından farklı olarak, gürültülü ve aykırı gözlemlere karşı daha dayanıklı bir algoritma olan k-medoids (k-ortaylar) yöntemi ile kümelendirilmiştir. Bu açıdan çalışmanın literatüre katkısı, gerçek müşteri verisi kullanılmasının yanı sıra, müşteri skorlarının çok kriterli karar verme tekniklerinden biri olan TOPSIS’le hesaplanıp, k-medoids kümeleme algoritması kullanılarak, geri kazanıldığında işletme kârına en çok katkıyı sağlayabilecek müşterilerin tespit edilmesidir.

2. TOPSIS YÖNTEMİ

TOPSIS yönteminin temelinde, m sayıda alternatifi ve n sayıda kriteri olan bir problemin n –boyutlu uzayda m noktaları ile ifade edilmesi ve her bir alternatifin ideal ve negatif idealden öklid uzaklığının hesaplanması yatmaktadır. Buna göre, karar kriterlerine dayalı olarak alternatiflerin ideal çözüme yakınlıkları hesap edildiği gibi, negatif ideal çözüme olan uzaklıkları da hesaplanmaktadır. TOPSIS yöntemi günümüzde, tedarikçi ve kuruluş yeri seçimi, servis kalitesi ölçümü, tedarik zinciri yönetimi ve atama problemleri gibi sıralama gerektiren çeşitli uygulama alanlarında kullanılmaktadır.

TOPSIS yönteminde ilk etapta, satırlar karar alternatiflerini (karar noktalarını), sütunlar ise karar kriterlerini (değerlendirmede kullanılacak faktörleri) gösterecek şekilde oluşturulan başlangıç karar matrisinin (A_{ij}) her bir elemanının, ait oldukları sütun toplamalarının kareköküne bölünerek normalizasyon işlemi gerçekleştirilmektedir. Buna göre karar matrisi (A_{ij}),

$$A_{ij} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad (1)$$

ve normalize edilmiş karar matrisi N_{ij} ,

$$N_{ij} = \frac{a_{ij}}{\sqrt{\sum_{i=1}^m a_{ij}^2}} \quad (i = 1, \dots, m \text{ ve } j = 1, \dots, n), \quad (2)$$

$$N_{ij} = \begin{bmatrix} n_{11} & \dots & n_{1n} \\ \vdots & \ddots & \vdots \\ n_{m1} & \dots & n_{mn} \end{bmatrix},$$

olarak oluşturulmaktadır. Karar kriterlerinin önem derecesi göz önünde bulundurularak her bir faktörün önem derecesine (w_{ij}) göre hesaplanan ağırlıklandırılmış normalize matris ise (v_{ij}) aşağıdaki gibi elde edilmektedir:

$$\sum_{i=1}^n w_i = 1$$

$$V_{ij} = \begin{bmatrix} w_1 n_{11} & \dots & w_n n_{1n} \\ \vdots & \ddots & \vdots \\ w_1 n_{m1} & \dots & w_n n_{mn} \end{bmatrix} \quad (3)$$

Her bir karar noktasının ideal (A^+) ve negatif ideal (A^-) çözüm değerleri ise:

$$A^+ = \left\{ \left(\max_i v_{ij} \mid \text{her bir sütüne ait maksimum değerler} \right) \right\}$$

için çözüm seti,

$$A^+ = \{v_1^+, v_2^+, \dots, v_n^+\},$$

olarak hesaplanırken,

$$A^- = \left\{ \left(\min_i v_{ij} \mid \text{her bir sütüne ait minimum değerler} \right) \right\}$$

olarak belirtilen negatif ideal için çözüm seti,

$$A^- = \{v_1^-, v_2^-, \dots, v_n^-\},$$

olarak hesaplanmaktadır. İdeal çözüm setlerinden olan uzaklıkların belirlenmesinde Öklid uzaklık ölçütünden yararlanılmaktadır. Buna göre, ideal S^+ ve negatif ideal S^- uzaklıklar,

$$s_i^+ = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^+)^2} \quad (4)$$

$$s_i^- = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^-)^2} \quad (5)$$

olarak elde edilmektedir. Karar alternatiflerinin ideal çözüme göreli yakınlığı (C_i) ise:

$$C_i = \frac{s_i^-}{s_i^- + s_i^+} \quad (6)$$

olarak hesaplanmaktadır. $0 \leq c_i \leq 1$ olup, sıfır noktasına yakınlık, karar noktasının negatif ideale yakınlığını ifade etmektedir. Hesaplanan göreli yakınlıklar sıralandırılıp, en yüksek göreli yakınlık değeri (C_i) sahip alternatif tercih edilmektedir (Hwang ve Yoon, 1981). Yöntemin avantajı, hesaplama

kolaylığının yanı sıra en iyi ve en kötü karar alternatiflerini ölçülebilir bir değerle açıklayabilirliği ve tüm alternatiflerin çok boyutlu olarak görselleştirilebilmesidir (Kim vd., 1997).

3. K-MEDOİDS ALGORİTMASI

Kümeleme analizi, veri noktalarını küme adı verilen doğal gruplara ayırma işidir. Bu ayırma sonucunda herhangi bir grup içerisindeki veri noktaları birbirine çok benzemekte iken, farklı kümelerde yer alan veri noktaları ise birbirine mümkün olduğunca benzememektedir (Zaki ve Meira Jr, 2014). Buna göre, kümeleme analizi sonlu bir grup nesneye yapılan bir sınıflandırma çeşidi olarak tanımlanabilir.

Nesneler arasındaki ilişki yani, nesnelerin birbirine benzemesi veya benzememesi durumu, oluşturulacak yakınlık matrisinden anlaşılabilir. Eğer söz konusu nesneler d –boyutlu ölçü uzayındaki örnekler veya noktalar ise, yakınlıklar nokta çiftleri arasındaki uzaklık (örneğin, öklidyen uzaklık) olarak düşünülebilir. Benzerlik (yakınlık) matrisinin kümeleme analizinin yapılabilmesi için tek girdi olmasından dolayı, nokta çiftleri arasında uzaklık veya benzerliğin olmadığı durumlarda kümeleme analizi gerçekleştirilememektedir (Jain ve Dubes, 1988). Bu sebeple, kümeleme analizi ham veri kümesi ile değil, kümelemesi yapılacak gözlem çiftlerinin benzerlik (yakınlık) veya benzeşmezlik (uzaklık) değerlerinin oluşturduğu matris ile yapılmaktadır.

Kümeleme analizinin aşamaları en genel anlamda; değişkenlerin seçimi, uzaklık fonksiyonu ile kümeleme algoritmasının seçimi, kümelemenin geçerlemesi ve sonuçların yorumlanması şeklinde sıralanabilir (Xu ve Wunsch, 2009).

Kümeleme analizinde ilk adımda, kümeleme analizinde kullanılacak olan nesneleri (örnekleri) birbirinden ayırt edici özelliği bulunan değişkenler seçimi yapılmaktadır. İkinci adım, ham veriden bir yakınlık (proximity), benzerlik veya ilişki göstergesinin hesaplanmasıdır (Aggarwal, 2015). Yakınlıktan kastedilen, literatürde sıkça karşılaşılan benzeşmezlik (dissimilarity) veya benzerlik (similarity) kavramlarına karşılık gelmektedir (Jain ve Dubes, 1988). Benzerlik, iki nesnenin birbirine ne kadar yakın olduğunun ölçüsü iken, benzeşmezlik ise aksine ne kadar uzak olduğunun ölçüsüdür. Benzeşmezlik genel olarak Öklid, Manhattan, Canberra ve Minkowski gibi bazı bilinen uzaklık ölçütleriyle hesaplanmaktadır.

Kümeleme analizinde üçüncü adımda ise, genel olarak hiyerarşik ve hiyerarşik olmayan kümeleme teknikleri şeklinde sınıflandırılan tekniklerden biri seçilmektedir. Hiyerarşik olmayan kümeleme tekniklerinde, küme sayısı (k) daha önce araştırmacı tarafından bilinmektedir. K -ortalamalar, k -medyanlar, k -medoids algoritmaları en bilinen hiyerarşik olmayan teknikler olarak sıralanabilir. Hiyerarşik algoritmalar ise, veri kümesine uzaklıkların oluşturduğu matrisi kullanarak kümeleme analizini uygulamaktadır. Tek bağlantı yöntemi, tam bağlantı yöntemi, ortalama bağlantı tekniği ve Ward'ın yöntemi bu grupta incelenmektedir. Son aşamada ise, yapılan kümelemenin tesadüfi çözümden farklı olup olmadığının belirlenmesi, yani geçerlemesi, dışsal, içsel veya göreceli ölçütlerle

gerçekleştirilmekte olup (Halkidi, Batistakis ve Vazirgiannis, 2001), sonuçlar dendogram veya diğer görsel tekniklerle, uzmanlarca yorumlanabilecek şekilde sunulmaktadır.

Veri seti içerisinde çok büyük değer alan gözlemler, analiz sonucunda dâhil olacağı kümenin merkez noktasını ve ortalamasını değiştirebileceği için, kümede bulunan gözlemlerin ortalamasını hesaplamak yerine, kümede orta noktaya en yakında konumlanmış gözlem kullanılarak kümeleme işlemi yapılmaktadır. Özellikle gürültülü ve aykırı gözlemlere sahip veri setlerinde küme sayısının tespit edilmesinde k-medoids ve k-means algoritmalarının kullanılması tavsiye edilmektedir (Lord vd., 2017). Kümeleme problemlerinde, daha önce vurgulandığı gibi göreceli kullanım kolaylığı, hesaplama performansı ve büyük veri setlerinde kullanımı açısından k-medoids algoritmasının kullanımı önerilmektedir (Park ve Jun, 2009; Arora vd., 2016).

K-medoids algoritması daha detaylı bir şekilde şöyle açıklanabilir. Bu yöntemde temsilci her zaman veri setinden seçilmektedir (Kaufman ve Rousseeuw 1987). Veri setinin yapısal özelliğini temsil eden k adet temsilci veri noktası seçilince, her bir kümede yer alan veri noktalarının ortalamasını hesaplamak yerine, her bir küme için her bir iterasyonda temsilci bir veri noktası veya medoid (ortay) belirlenmektedir. Her bir küme için medoidler şu şekilde seçilmektedir (Reynolds vd., 2004).

Bir kümede aşağıdaki toplamı minimize edecek herhangi bir nokta (i) seçilir.

$$\sum_{j \in C_i} d(i, j)$$

Burada C_i kümesi i noktasını içeren küme olmakla beraber $d(i, j)$ ise i ve j noktaları arasındaki uzaklıktır. K-medoids algoritması şu şekildedir:

1. k adet veri noktası rassal olarak ilk küme merkezleri olması açısından seçilir,
2. Veri noktaları kendisine en yakın olan merkezin (medoid) bulunduğu kümeye atanır,
3. k adet medoidin pozisyonu tekrar hesaplanır,
4. Adım 2 ve adım 3 medoidler tam yerlerini buluncaya kadar tekrarlanır,

Adım 3, her bir iterasyonda yukarıdaki formül yardımıyla hesaplanabilir.

K-medoids algoritmasının diğer algoritmalara tercih edilmesinin nedenleri daha önce vurgulandığı gibi, k-ortalamar algoritmasındaki temsilcinin o kümede yer alan aykırı bir değer sebebiyle sapmış olabileceği ihtimali ve karma tipteki veri noktalarının, ortalama veya medyan ile temsilci seçiminin bazen zor olması olarak sıralanabilir. Bu yüzden, k-medoids algoritmasının en önemli özelliği veri kümesine uygun benzerlik veya uzaklık fonksiyonu tanımlanabildiği sürece, algoritmanın her veri tipi için kullanılabilir olmasıdır (Aggarwal, 2015; Zaki ve Meira Jr, 2014). K-medoids algoritmasının k-ortalamar algoritmasından önemli farkı ise, temsilciyi orijinal veri tabanından seçerken, her bir boyutta bağımsız olarak seçilecek medyanların, d boyuttaki temsilcisi orijinal veri kümesine ait olmayabileceğidir.

K-medoids kümelemesi için Kaufman ve Roussew'in (1990) medoidler etrafında bölümlenme (PAM) algoritması en güçlü teknik olarak karşımıza çıkmaktadır, fakat bu algoritma büyük veri setlerinde çok etkin çalışmamaktadır (Han vd., 2011). PAM algoritmasında küme sayısı (k) önceden belirlenmekte olup, her bir iterasyonda en iyi kümelemeyi elde etmek için medoid (ortay) nokta olan i ve medoid olmayan nokta j tespit edilmektedir. Amaç fonksiyonu, her bir noktanın en yakın ortaya olan uzaklıkları toplamıdır. Algoritma ilk aşamada sıfır medoid ile başlayıp, her bir iterasyonda bir tane medoid bularak k adet medoid oluşana kadar devam etmektedir. Her bir iterasyonda yeni bir medoid aranmasının sebebi, amaç fonksiyonunu minimize etmektir.

Kaufman ve Roussew (1990) PAM metodunu uygulayan CLARA isimli algoritmayı önermiştir. Bu çalışmadan sonra k-medoids tekniğini geliştirmeye yönelik çalışmalar olmuştur. Lucasius, Dane ve Kateman (2004) küme sayısının artması ile birlikte k-medoids algoritmasının etkin çalışmadığını tespit etmiş, yeni bir k-medoids algoritmasını genetik algoritmalar kullanarak geliştirmişlerdir. Wei, Lee ve Hsu da (2003) CLARA algoritmasının performansını büyük veri setleri için denemiş olup; Laan, Pollard ve Bryan ise (2003) medoidlere olan toplam uzaklığı minimize etmek yerine PAM algoritmasının sunduğu siluet (silhouette) ölçüsünü maksimize etmeyi amaçlamıştır.

4. SİLÜETLER (SILHOUETTES)

Siluetler veri noktalarının kümelerinde ne derecede iyi veya kötü yer aldıklarını gösteren yani elde edilen kümelemenin kalitesini ölçmeye yarayan bir ölçü olarak tanımlanmaktadır (Kaufman ve Rousseeuw, 1987). Örneğin; her bir i veri noktası için, $a(i)$, i veri noktasının C_i kümesindeki diğer veri noktalarından olan ortalama uzaklığı olsun. Diğer bütün kümeler ($C \neq C_i$) için $d(i,C)$, i veri noktasının C 'deki diğer veri noktalarından olan ortalama uzaklığı olsun. Kümelerin hepsi için ($C \neq C_i$), $d(i,C)$ hesaplandıktan sonra $b(i)$ aralarındaki en küçük değere sahip olan olsun. Bu en küçük uzaklığın bulunduğu küme i veri noktasının en yakın komşusu olarak adlandırılır. Aşağıda formülü verilen $s(i)$ değeri i veri noktasının C_i kümesine, komşu kümeye nazaran ne derece uyduğunu (fit) ölçen bir ölçüdür.

$$s_i = \frac{b(i)-a(i)}{\max(a(i),b(i))} \quad (7)$$

Eğer $s(i)$ değeri bire yakınsa, i veri noktası iyi bir kümede yer almıştır. Fakat $s(i)$ değeri sıfıra yakınsa i veri noktasının C_i kümesine mi yoksa komşu kümeye mi ait olduğu hususu kuşkuludur. $s(i)$ değerinin negatif olması ise i veri noktasının yanlış kümede yer aldığını göstermektedir.

Bir kümedeki bütün i veri noktaları için $s(i)$ değerinin ortalaması, o kümenin ortalama siluet genişliğidir. Bütün veri noktaları için $s(i)$ değerinin ortalaması, bütün veri seti için ortalama siluet genişliğini ifade etmektedir ve $\bar{s}(k)$ ile gösterilmektedir, buradaki k ise küme sayısıdır.

Kaufman ve Rousseeuw (1987) en iyi k değerinin, maksimum $\bar{s}(k)$ değerinin seçilmesi ile bulunulabileceğini ifade etmektedir.

5. TOPSIS VE K-MEDOIDS YÖNTEMLERİ İŞLETME LİTERATÜRÜ

TOPSIS yönteminin işletme literatüründe uygulama alanına bakıldığında, pazarlama, insan kaynakları yönetimi, üretim ve lojistik, muhasebe - finans uygulamalarında son beş sene yapılan çalışmaların sayısı sınırlı kalmaktadır. Beikhhakhian vd. (2015) firmaların çeşitli kriterler altında tedarikçi seçimine yönelik TOPSIS yöntemini uygulamıştır. Kusumawardani ve Agintiara (2015) işletmenin değer yaratan faaliyetlerindeki rolleri göz önünde bulundurularak, işletme içerisindeki insan kaynakları yönetimini TOPSIS yöntemi ile gerçekleştirmiş; Sekhar vd. (2015) küçük ölçekli işletmelerin global ekonomide rekabet avantajı sağlayabilmesinde gerekli olan entelektüel sermaye için önemli göstergelerin belirlenmesinde AHP ve TOPSIS yöntemini kullanmıştır. Kang vd. (2016) çok kriterli karar verme yöntemlerinden biri olan TOPSIS yöntemiyle, firmadan müşteriye pazarlama e-ticaret sitelerinin hizmet kalitesi değerlendirilmesinde, kullanıcıların e-ticaret sitesinden bekledikleri ve algıladıkları hizmet arasındaki farklıları önem sırasına göre tespit etmiştir. Mavi vd. (2016) tedarik zinciri risk yönetimine yönelik zamanında teslim, performans geçmişi, arz ve talep riski, lojistik riski gibi çeşitli kalite kriterinin ağırlıklarını kullanarak, tedarikçilerin sıralandırılmasında TOPSIS yöntemini uygulamıştır. Zyoud ve Hanusch (2017) TOPSIS ve AHP teknikleriyle geniş bir uygulama alanında gerçekleştirilen çalışmalara ilişkin geniş bir bibliyografik çalışma gerçekleştirmiştir. Buna göre, TOPSIS'le gerçekleştirilen yaklaşık 2412 makale taranmış, söz konusu çalışmaların büyük çoğunluğunun coğrafik bilgi sistemleri, risk modellemesi ve tedarik zinciri yönetimi alanlarında gerçekleştirildiği sonucuna ulaşılmıştır. Sirisawat ve Kiatcharoenpol (2018) işletmelerde kurumsal kaynak planlamaya yönelik ürün geri dönüşümü ve yeniden kullanım – yeniden üretim süreçlerini içeren tersine lojistiğin adaptasyonu için önemli faktörlerin belirlenmesi ve tersine lojistik uygulama çözümlerini sıralamada bulanık AHP ve TOPSIS'den yararlanmıştır. Han ve Trimi (2018) e-ticaret platformlarının kurumsal sosyal sorumluluk, tersine lojistik, yeşil ve çevreci aktiviteler çerçevesinde re-organizasyonu için belirlenecek faaliyetlerin skorlarının oluşturulması ve önem derecelerine göre sınıflandırılmasında TOPSIS yöntemini kullanmıştır.

Brusco v.d. (2012) müşterilerin yaşam stili dinamikleri, satış promosyon analizi ve kampanya yönetimi, müşteri ilişkileri yönetimi uygulama alanlarında kullanımı gelişmekte olan ve çeşitli yazılımlara entegresi gerekliliğine dikkat çekilen kümeleme algoritmalarından birinin k-medoid olduğunu belirtmiştir. Çalışmada ayrıca kümeleme algoritmalarının çok boyutlu müşteri ve pazar segmentasyonu ile marka konumlandırma uygulamalarındaki etkinliğine dikkat çekilmiştir. Brito v.d. (2015) müşteriye / müşteri gruplarına özel çok çeşitli ürün gamının sunulduğu moda endüstrisinde üretim planlama ve pazarlama problemlerine çözüm getirmek üzere, müşteri ürün tercihlerindeki örüntülerini segmente etmek amacıyla k-medoids ve CN2-SD kümeleme algoritmalarını kullanmıştır. Çalışma sonucunda, söz konusu tekniklerin müşterilerin beğeni ve tercihlerinin anlaşılmasında ve rekabet avantajının sağlanmasında etkin olarak kullanılabilceği sonucuna ulaşılmıştır. Ghnemat ve Jaser (2015), telekomünikasyon sektöründe müşteri segmentasyonu oluşturmada, kümeleme

yöntemlerinden yararlanmışlardır. Çalışmada, öz-düzenleyici haritalar yöntemi ile tespit edilen müşteri segment sayısına dayalı olarak, müşterilerin mobil telefon kullanım özelliklerinin altında yatan örüntüleri tespit etmede kümeleme algoritmalarından yararlanmışır.

Bianchi v.d. (2016) pazarlama politikaları geliştirme ve müşterilerin ihtiyaçlarına uygun yeni iletişim hatlarının inşası amacıyla, bir mobil operatörünün müşteri arama veri setini kullanarak, müşterilerin servis kullanım alışkanlıklarını anlama ve kullanıcı profillerini çıkarmak üzere k-medoids yöntemini içeren PROCLUS kümeleme algoritmasını kullanmıştır. Fu v.d. (2017) online oyun oynayan kişiler için müşteri elde tutma stratejileri belirlemek üzere, kişilerin oyun oynama sıklığı ve oyunda kaldığı süre, gerçekleştirdiği görevler ve kazandığı başarılar, oyun içerisindeki rolü ve sanal arkadaş sayısı gibi çeşitli oyun içi bağıllık, performans ve sosyal etkileşim değişkenlerini k-medoids, bulanık c-means kümeleme algoritmalarını kullanarak müşteri segmentasyonu gerçekleştirmiştir. Berget (2018) müşteri ve pazar segmentasyonunda kullanılabilir istatistiksel yaklaşımları irdeleyerek, k-medoids algoritmasının ilk adım çözümünde merkez seçiminin analize etkisi olmadığı ve daha kararlı sonuçlar elde edilebileceği sonucuna ulaşmıştır.

6. VERİ SETİ VE UYGULAMA

Bir telekomünikasyon firmasının, Ocak 2016 – Aralık 2016 tarihleri arasında rakip firmaya geçen 1145 müşterisine ait verilerin kullanılarak, hangi müşteriler için geri kazanma faaliyetleri gerçekleştirilmesi gerektiğinin TOPSIS yöntemi ve kümeleme analizi ile tespitinin sağlanması amaçlanmaktadır. Müşterilere ait veri setinde, demografik bilgilerin yanı sıra; ses, veri ve katma değerli servis kullanımlarına ilişkin, müşterilerin aylık ortalama telefon görüşmesi sayısı ve süresi, uluslararası görüşmelerin diğer görüşmeler içerisindeki oranı, sabit hatlarla yapılan görüşmelerin yüzdesi, cep telefonları ile yapılan görüşmelerin yüzdesi, veri hizmetleri kullanım süresi ve diğer katma değerli servis kullanımını tanımlayan 23 farklı değişken bulunmaktadır. Veri setinde yer alan değişkenlerden, en iyi müşterileri belirlemesi açısından, 17 tanesi seçilmiş ve bu değişkenlerin TOPSIS skorlarına pozitif yönlü katkı yaptığı düşünülmüştür. Bu değişkenler ve aldıkları değerlerin bazıları Tablo 1’de gösterilmiştir.

Söz konusu değişkenler ve ağırlıkları (hepsine eşit şekilde ağırlık verilmiştir) temel alınarak R programlama dilinde TOPSIS yöntemi uygulanmıştır (Kodlar için bakınız: Ek-1). R programlama dilinde “*topsis*” paketinde yer alan “*as.matrix*” işlevi kullanılarak öncelikle veri seti bir matris şekline dönüştürülmüş; değişkenlerin ağırlıkları ve değişkenlerin TOPSIS skoruna pozitif katkı sağlayacakları programa girilmiştir. Sonrasında, her bir gözleme ait TOPSIS skorları hesaplanmış olup, TOPSIS skorları en yüksek on müşteri Tablo 2’de verilmiştir.

Tablo 1. TOPSIS Skorları İçin Kullanılan Değişkenler ve Bazı Müşteriler için Aldıkları Değerler

Değişkenler	Müşteri1	Müşteri2	Müşteri3	Müşteri4
Son ayda yapılan görüşme sayısı	0	0	0	1
Son iki ayda yapılan ortalama görüşme sayısı	0	0	0.5	0.5

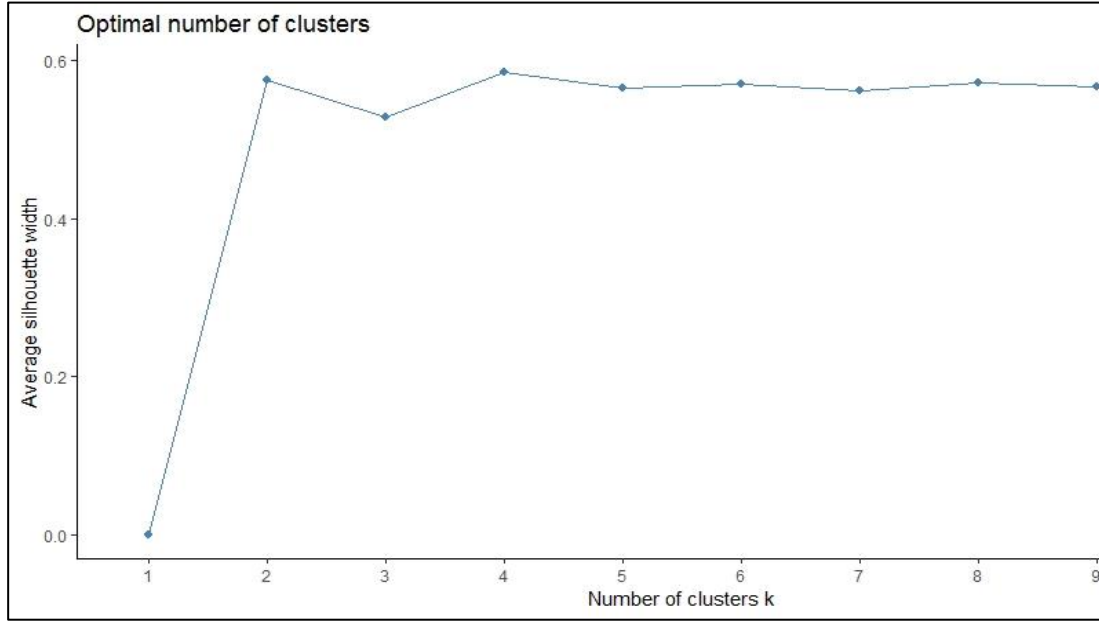
Son ayda kullanılan bütün hizmetlerin toplam tutarı (TL)	31.675	63.607	31.4	93.337
Son iki ayda kullanılan bütün hizmetlerin tutar ortalaması (TL)	32.766	63.36	34.89	77.694
Son ayda kullanılan sesli görüşme tutarı (TL)	26.245	63.457	31.4	81.36
Son iki ayda kullanılan görüşme miktarı ortalaması (TL)	29.506	63.15	34.86	69.068
Son ayda kullanılan verinin tutarı (TL)	5.43	0.15	0	11.977
Son 6 ayda kullanılan verinin ortalama tutarı (TL)	3.21	0.21	0.03	8.626
Son ayda kullanılan katma değerli diğer hizmetler tutarı (TL)	0	0	0	0
Son iki ayda kullanılan katma değerli diğer hizmetlerin ortalama tutarı (TL)	0.05	0	0	0
Son iki ayda yapılan ortalama görüşme süresi	774	1381.5	768.5	3141.5
Son ayda yapılan sabit hat görüşme süresi	0.168	0.008	0.021	0
Son ayda yapılan şebeke içi görüşme süresi	0.566	0.907	0.947	0.305
Son ayda yapılan diğer hatlar görüşme süresi	0.185	0.072	0.001	0.007
Son iki ayda yapılan ortalama roaming araması	0	0	0	0.007
Son iki ayda yapılan uluslararası görüşmelerin ortalama yüzdesi	0	0	0	0
Son iki ayda yapılan bedava görüşmelerin ortalama yüzdesi	0	0.008	0.029	0

Elde edilen TOPSIS skorları kümeleme analizine tâbi tutularak, müşteri segmentasyonu gerçekleştirilmiştir. Müşteri segmentasyonundaki amaç, geri kazanma faaliyetleri yapılacak müşteri grubunun elde edilmesidir. R programlama dilinde yer alan “cluster” paketindeki “pam” işlevi kullanılarak, k-medoids algoritmasıyla gerçekleştirilen kümeleme analizi sonucunda, her bir küme sayısı için (k=1, k=2, k=3, k=4...) ortalama siluet değerleri hesaplanmış ve veri setinin, ortalama siluet genişliğinin en yüksek olduğu dört kümeye (k=4), ayrılması uygun görülmüştür (Grafik 1).

Tablo 2. TOPSIS Skorları En Yüksek 10 Müşteri

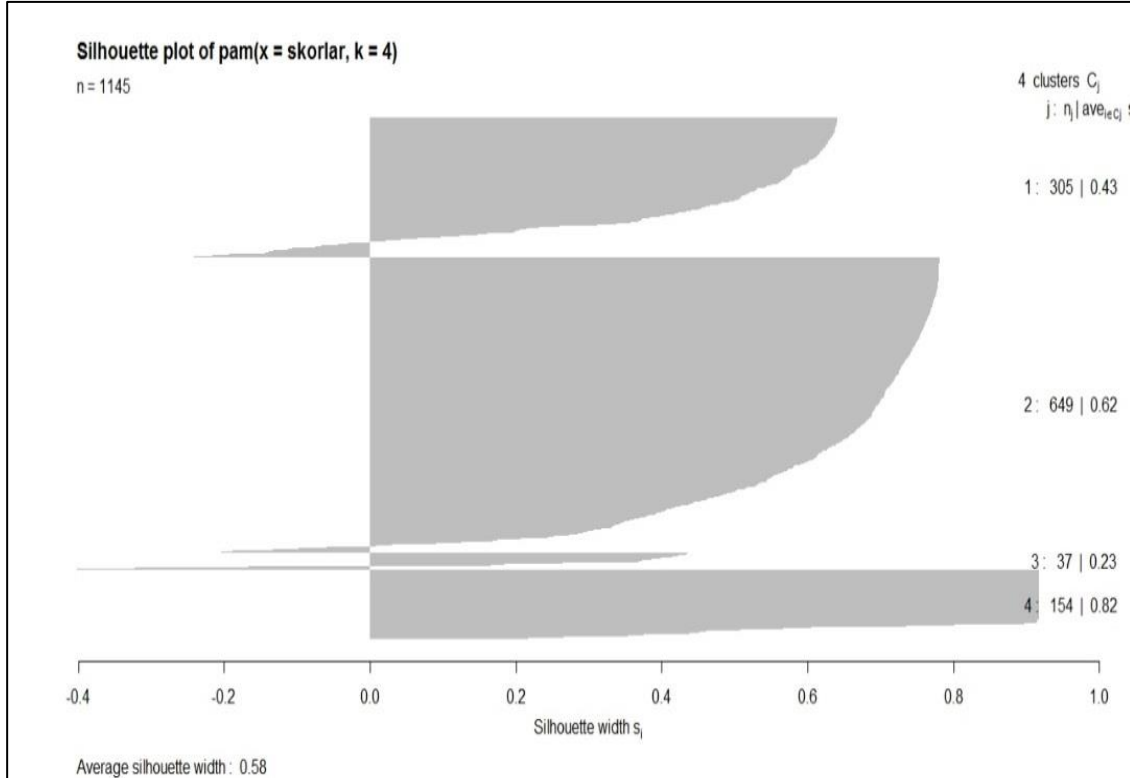
Müşteri sıra no	Müşteri numarası	TOPSIS skoru
7955929..	216	0.542895048
7955929..	217	0.337113889
7955936..	258	0.289238525
7955924..	95	0.21616948
7955938..	297	0.197290582
7955925..	107	0.173761471
7955922..	59	0.167498169
7958991..	940	0.156493827
7955926..	152	0.132244861
7957869..	644	0.12734715

Grafik 1. Farklı Küme Değerleri İçin K-Medoids Sonucu Ortaya Çıkan Ortalama Siluet Değerleri



Veri seti, diğer bir ifade ile 1145 müşteri dört segmente ayrıldığında karşımıza çıkan siluet şekli ise Grafik 2’de verilmiştir. Dört küme için, ortalama siluet genişliği 0,58 olup, Kaufman ve Rousseeuw’e (1990) göre gücü orta seviyede olan bir küme yapısının ortaya çıktığı söylenebilir. Gerçek hayat verisinin kullanıldığı bir çalışmada bu sonucun elde edilmesinin olumlu olduğu düşünülmektedir.

Grafik 2: Dört Küme Değeri İçin Ortaya Çıkan Siluet Grafiği



Veri seti, kümeleme analizi sonucu belirlenen dört kümeye ayrıldığında Tablo 3’de, detaylı bir şekilde gösterilen müşteri özellikleri ortaya çıkmıştır. Tablo incelendiğinde üçüncü segmentte yer alan

37 kişilik müşteri grubunun firma için en değerli grup olduğu düşünülebilir. Bu segmentte yer alan müşterilerin TOPSİS skor ortalamalarının diğer gruplara göre çok yüksek olmasının sebebi bu segmentteki her türlü kullanım miktarı ve tutarlarının diğer segmentlere göre çok yüksek olmasından kaynaklanmaktadır. Dolayısıyla firma geri kazanma faaliyetlerine, işletme kârlılığına yaptıkları pozitif ve yüksek katkı açısından bu segmentten başlamalı ayrıca kaynaklar yeterli ise faaliyetlere birinci segmentten devam etmelidir.

Tablo 3. Segmentlerin Özellikleri

Segment	Müşteri Sayısı	TOPSİS Skor Ortalaması	Müşterilerin Baskın Özellikleri	Segment İsmi
1	305	0,049	Toplam fatura tutarı ortalamasının ve toplam sesli görüşme tutarı ortalamasının yaklaşık 25 TL olduğu, ortalama veri kullanım tutarının 3 TL kadar olduğu ve konuşma süresinin ortalama yaklaşık 465 dakika, uluslararası görüşme yüzdesinin ise ortalama %11 olduğu müşteri grubu	Gümüş
2	649	0,022	Toplam fatura tutarı ortalamasının ve toplam sesli görüşme tutarı ortalamasının yaklaşık 10 TL olduğu, ortalama veri kullanım tutarının 1 TL olduğu, konuşma süresinin ortalama yaklaşık 200 dakika olduğu ve uluslararası görüşme yüzdesinin ortalama %3 olduğu müşteri grubu	Bronz
3	37	0,142	Toplam fatura tutarı ortalamasının 158 TL, toplam sesli görüşme tutarı ortalamasının 147 TL, ortalama veri kullanımının 10 TL olduğu, konuşma süresinin ortalama 1500 dakika olduğu ve uluslararası görüşme yüzdesinin ortalama %3 olduğu müşteri grubu	Altın
4	154	0,0018	Her tipteki kullanım verilerinin sıfır veya sıfıra çok yakın olduğu müşteri grubu	-

7. SONUÇ VE ÖNERİLER

Günümüzde teknoloji ile birlikte gelişen veri elde etme ve depolama faaliyetleri ile birlikte firmalar tarafından sahip olunan büyük veri, firmalar tarafından yönetilme ve analizini etme ihtiyacını doğurmuştur. Büyük veri terimi, çok büyük boyutlarla ifade edilen veri için kullanılmakta olup ayrıca içeriğinde firmanın müşterilerine, paydaşlarına ait karma tipteki birçok değişkeni de barındırmaktadır. Müşterilere ilişkin yapısal ve yapısal olmayan verilerin yönetilmesi konusu müşteri ilişkileri yönetimi açısından önemli olduğu gibi, söz konusu verilerin içerisindeki örüntülerin keşfedilmesi, firmaların geleceğe yönelik tahminlerini geliştirmeye katkıda bulunmaktadır. Müşteri kaybının önüne geçilebilmesi için mevcut müşterilerin elde tutulması ya da kaybedilmiş müşterilere ilişkin müşteri geri

kazanma faaliyetlerinin gerçekleştirilmesi gibi pazarlama uygulamaları, son yıllarda pazarlama literatüründe çalışılan önemli konulardan biridir.

Bu çalışmada çok kriterli karar verme tekniklerinden olan TOPSIS yöntemi ve gözetimsiz veri madenciliği yöntemlerinden olan kümeleme analizi kullanılarak telekomünikasyon sektöründe pazarlama stratejileri için müşteri seçimine yönelik bir yaklaşım önerilmektedir. Bu bağlamda, telekomünikasyon sektöründe faaliyet gösteren bir işletmenin başka firmalara geçen müşteri verisi kullanılarak, müşteriler TOPSIS yöntemi ile önem derecelerine göre sıralanmış ve k-medoids kümeleme analizi ile müşteriler önem derecelerine göre segmentlere ayrılmıştır. Analiz sonuçlarında, firmanın üçüncü segmentte (altın segment) yer alan müşterilerine geri kazanma faaliyetleri düzenlenmesi gerektiği belirlenmiştir. Söz konusu segmentte yer alan müşteri grubunun uluslararası görüşme yüzdesinin diğer gruplarda bulunan müşterilerle aynı veya daha düşük oranda gerçekleştiği dikkat çekmektedir. Bu da, müşteri geri kazanma faaliyetlerinde müşterilerin uluslararası görüşme hizmetine dayalı pazarlama stratejilerinin üretilmesinin yerine, konuşma ve veri kullanım hizmetlerine yönelik stratejilerin oluşturulması gerekliliğini ortaya koymaktadır.

Müşterilerin hizmet ve ürün tercihlerinin analitik bir bakış açısıyla ele alınarak, kârlı müşteri gruplarının firmaya sağladığı değer analiz edilmesi ve söz konusu gruplara ilişkin yeni kanal ve ürün geliştirme stratejilerinin belirlenmesi, müşteri farklılıklarının göz önüne alınarak segmentasyon yapılmasıyla mümkün olup, bu durum ayrıca müşteri kaybının engellenmesi açısından da önem kazanmaktadır. İşletme kaynaklarının tüm müşterilere aynı şekilde dağıtılması yerine, müşteri verisine dayalı bir şekilde kaynakların konumlandırılması, yeni müşterilerin kazanılmasının mı yoksa kaybedilen müşterilere yönelik pazarlama stratejilerinin geliştirilmesinin mi avantajlı olacağı sorusunu cevaplamaktadır. Bu açıdan, işletmenin kârına katkı sağlayan müşteri gruplarının nitelik ve niceliklerine göre segmente edilmesi konusu karşımıza çıkmaktadır. Yapısal ve yapısal olmayan müşteri verilerine dayalı segmentasyon için doğru analitik araçların kullanımı gerekli olup, özellikle veri madenciliğindeki kümeleme algoritmalarının bu alandaki kullanımının yaygınlaştığı görülmektedir. Yapılacak çalışmalarda, müşteri verisine uygun şekilde seçilecek modeller veya hiyerarşik, hiyerarşik olmayan veya hibrit kümeleme yaklaşımları gibi tekniklerle uygulama sonucundaki performansları da göz önünde bulundurularak, müşteri veya pazar segmentasyonu analizinin gerçekleştirilmesi; segmentasyon için kullanılacak olan değişkenlerin seçimi ile ağırlıklandırılması hususunun uzman kişilerin görüşleri alınarak yapılması önerilmektedir.

KAYNAKÇA

Aggarwal, C. C. (2015) “Data Mining: The Textbook”, Switzerland: Springer.

Arora, P. ve Varshney, D.S. (2016) “Analysis of K-Means and K-Medoids Algorithm For Big Data”, Procedia Computer Science, 78: 507-512.

- Beikkhakhian, Y., Javanmardi, M., Karbasian, M. ve Khayambashi, B. (2015) “The Application of ISM Model in Evaluating Agile Suppliers Selection Criteria and Ranking Suppliers Using Fuzzy TOPSIS-AHP Methods”, *Expert Systems with Applications*, 42 (15–16): 6224-6236.
- Berget, I. (2018) “Statistical Approaches to Consumer Segmentation”, *Methods in Consumer Research, New Approaches to Classic Methods*, 1: 353–382.
- Bianchi, F.M., Rizzi, A., Sadeghian, A. ve Moiso, C. (2016) “Identifying User Habits Through Data Mining on Call Data Records”, *Engineering Applications of Artificial Intelligence*, 54: 49-61.
- Brito, P.Q., Soares, C., Almeida, S., Monte, A. ve Byvoet, M. (2015) “Customer Segmentation in a Large Database of an Online Customized Fashion Business”, *Robotics and Computer-Integrated Manufacturing*, 36: 93-100.
- Brusco, M.J., Steinley, D., Cradit, J.D. ve Singh, R. (2012) “Emergent Clustering Methods for Empirical OM Research”, *Journal of Operations Management*, 30(6): 454-466.
- Craig A. B. ve Howard J. S. (2018) “U.S. Consumer Preferences for Telephone and Internet Services, Evidence from the RAND American Life Panel”, RAND Corporation, <https://www.rand.org/>, (11.03.2018).
- Deloitte (2015) “Opportunities in Telecom Sector: Arising from Big Data”, (<https://www2.deloitte.com>), (23.04.2018).
- Duncan, W. J. (1978) “Essentials of Management, 2nd Edition”, Hinsdale, The Dryden Press, Illinois
- Fu, X., Chen, X., Shi, Y.T., Bose, I. ve Cai, S. (2017) “User Segmentation for Retention Management in Online Social Games”, *Decision Support Systems*, 101: 51-68.
- Ghneimat, R. ve Jaser, E. (2015) “Classification of Mobile Customers Behavior and Usage Patterns using Self-Organizing Neural Networks”, *International Journal of Interactive Mobile Technologies*, 9(4): 4-11.
- Halkidi, M., Batistakis, Y. ve Vazirgiannis, M. (2001) “On Clustering Validation Techniques”, *Procedia - Social and Behavioral Sciences*, 189: 275-284.
- Han, J., Pei, J., & Kamber, M. (2011) “Data Mining: Concepts and Techniques”, Elsevier.
- Han, H. ve Trimi, S. (2018) “A fuzzy TOPSIS Method for Performance Evaluation of Reverse Logistics in Social Commerce Platforms”, *Expert Systems with Applications*, 103: 133-145.
- Harikumar, S. ve Surya PV. (2015) “K-Medoid Clustering for Heterogeneous Data Sets”, *Procedia Computer Science*, 70: 226-237.
- Hwang, C. L. ve Yoon, K. (1981) “Multiple Attribute Decision Making: A State of the Art Survey”, Springer-Verlag, New York.

- Jain, A. K. ve Dubes, R. C. (1988) “Algorithms for Clustering Data”, Prentice-Hall, Inc. *Journal of Intelligent Information Systems*, 17(2-3): 107-145.
- Kang, D., Jang, W. ve Park, Y. (2016) “Evaluation of E-Commerce Websites Using Fuzzy Hierarchical TOPSIS Based on E-S-QUAL”, *Applied Soft Computing*, 42: 53-65.
- Kaufman, L. ve Rousseeuw, P. J. (1987) “Clustering by Means of Medoids, Statistical Data Analysis Based on The L1-Norm and Related Methods”, North-Holland: 405–416.
- Kaufman, L. ve Rousseeuw, P. J. (1990) “Finding Groups in Data: An Introduction to Cluster Analysis”, USA: Wiley.
- Kim, G., Park, C.S. ve Yoon, K.P. (1997) “Identifying Investment Opportunities for Advanced Manufacturing Systems with Comparative-Integrated Performance Measurement”, *International Journal of Production Economics*, 50: 23-33.
- Kusumawardani, R.P. ve Agintiara, M. (2015) “Application of Fuzzy AHP-TOPSIS Method for Decision Making in Human Resource Manager Selection Process”, *Procedia Computer Science*, 72: 638-646.
- Laan, M.V., Pollard, K. ve Bryan, J. (2003) “A New Partitioning around Medoids Algorithm”, *Journal of Statistical Computation and Simulation*, 73(8), 575-584.
- Lord, E., Willems, M., Lapointe, F.J. ve Makarenkov, V. (2017) “Using the Stability of Objects to Determine the Number of Clusters in Datasets”, *Information Sciences*, 393: 29-46.
- Lucasius, C.B., Dane, A.D., Kateman, G. (1993) “On k-Medoid Clustering of Large Data Sets with the Aid of a Genetic Algorithm: Background, Feasibility and Comparison”, *Analytica Chimica Acta*, 282 (3): 647-669.
- Mavi, R.K., Goh, M. ve Mavi, N.K. (2016) “Supplier Selection with Shannon Entropy and Fuzzy TOPSIS in the Context of Supply Chain Risk Management”, *Procedia - Social and Behavioral Sciences*, 235: 216-225.
- Park, H.S. ve Jun, C.H. (2009) “A Simple and Fast Algorithm for K-Medoids Clustering”, *Expert Systems with Applications*, 36(2): 3336-3341.
- Reynolds, A. P., Richards, G. ve Rayward-Smith, V. J. (2004) “The Application of K-Medoids and Pam to the Clustering of Rules”, In *International Conference on Intelligent Data Engineering and Automated Learning*, Springer, Berlin, Heidelberg.
- Saglam, B., Salman, F., Sayin, S. ve Trkay, M. (2006) “A mixed-Integer Programming Approach to the Clustering Problem with an Application in Customer Segmentation”, *Eur. J. Oper. Res.*, 173(3), 866–879 .

- Sekhar, C., Patwardhan, M. ve Vyas, V. (2015) “A Delphi-AHP-TOPSIS Based Framework for the Prioritization of Intellectual Capital Indicators: A SMEs Perspective”, *Procedia - Social and Behavioral Sciences*, 189: 275-284.
- Sirisawat, P. ve Kiatcharoenpol, T. (2018) “Fuzzy AHP-TOPSIS approaches to Prioritizing Solutions for Reverse Logistics Barriers”, *Computers & Industrial Engineering*, 117: 303-318.
- Teichert, T., Shehu, E. Y. ve Von Wartburg, I. (2008) “Customer Segmentation Revisited: the Case of the Airline Industry”, *Transportation Research Part A*, 42: 227-242.
- Wei, C.P., Lee, Y.H. ve Hsu, C.M. (2003) “Empirical Comparison of Fast Partitioning-Based Clustering Algorithms for Large Data Sets”, *Expert Systems with Applications*, 24: 351-363.
- Xu, R. ve Wunsch, D.C. (2009) “Clustering”, <http://site.ebrary.com/id/10257659>. (30.03.2018).
- Zaki, M. J., Meira, W. (2014) “Data Mining and Analysis: Fundamental Concepts and Algorithms”, UK: Cambridge University Press.
- Zyoud, S.H. ve Hanusch, D.F. (2017) “A Bibliometric-Based Survey on AHP and TOPSIS Techniques”, *Expert Systems with Applications*, 78: 158-181.

EK-1 ANALİZLERDE KULLANILAN R KODLARI

```
telekom <- read.csv("new.csv",header=FALSE,sep=";")  
  
head(telekom)  
  
d <- as.matrix(telekom,nrow=1145,ncol=17,byrow = TRUE)  
  
w <- c(1, 1, 1,1, 1, 1,1, 1, 1,1, 1, 1,1, 1, 1,1,1)  
  
i <- i <- c("+", "+", "+","+ ", "+", "+", "+", "+", "+", "+", "+", "+", "+", "+", "+", "+")  
  
library(topsis)  
  
sonuc1 <- topsis(d,w,i)  
  
write.csv(sonuc1,"sonuc1.csv")  
  
skorlar <- read.csv("sonuc1",header=FALSE,sep=";")  
  
library(pam)  
  
library(cluster)  
  
library(factoextra)  
  
fviz_nbclust(skorlar, pam, method = "silhouette")+  
theme_classic()  
  
pam <- pam(skorlar,4 ,diss=FALSE,metric="euclidean",stand=FALSE)  
  
pam$clustering
```