

Trafik Kaza Verilerinin Kümelenme Analizi Yöntemi ile Modellenmesi

Yetiş Şazi MURAT*
Alper ŞEKERLER**

ÖZ

Son yıllarda, ülkemizde artan araç sahipliği ve trafik yoğunluğuna paralellik gösteren trafik kazaları acil çözüm gerektiren önemli bir problem haline gelmiştir. Bu problemi çözmek için kullanılan en yaygın yöntem ise kara noktaların belirlenmesi ve iyileştirilmesidir. Geleneksel kara nokta belirleme yöntemi, her bir kazanın yerinin harita üstünde işaretlenmesi ve en çok işaretin bulunduğu yerlerin “kara nokta” olarak etiketlenmesini kapsamaktadır. Bu çalışmada ise, Denizli ili 2004, 2005 ve 2006 yıllarına ait trafik kaza verileri, bilgisayar programları kullanılarak klasik ve bulanık kümelenme yöntemleriyle analiz edilmiş, elde edilen küme merkezlerine yakın bölgelerdeki trafik kazalarının daha yoğun olduğu noktalar kara nokta olarak belirlenmiştir. Belirlenen kara noktalar detaylı biçimde ele alınarak kazaya neden olan unsurlar incelenmiştir. Elde edilen sonuçlar yorumlanmış ve kara noktalar için çeşitli çözüm önerileri getirilmiştir.

Anahtar Kelimeler: Kümelenme analizi, modelleme, kara noktalar, trafik kazaları

ABSTRACT

Modelling Traffic Accident Data by Cluster Analysis Approach

In recent years, traffic accidents have become an urgent problem due to increasing car ownership and traffic density. One of the most common methods for this problem is determination and analysis of “black spots”. The conventional black spot identification method includes marking the location of each accident with a pin and investigation of black spots considering density of the pins on a map. In this study, the traffic accidents data of Denizli city for the years of 2004, 2005 and 2006 have been analyzed using the k-means and the fuzzy clustering methods. The spots that are densely located around the cluster centers are determined as “black spots” and are analyzed. The results of the analysis are evaluated regarding all features of the black spots and recommendations for improving traffic safety are reported.

Keywords: Cluster analysis, modelling, black spots, traffic accidents

Not: Bu yazı

- Yayın Kurulu'na 03.09.2008 günü ulaşmıştır.
- 30 Eylül 2009 gününe kadar tartışmaya açıktır.

* Pamukkale Üniversitesi, İnşaat Müh. Bölümü, Denizli - ysmurat@pau.edu.tr

** Pamukkale Üniversitesi, Fen Bilimleri Enstitüsü, Denizli - rep_la@hotmail.com

1. GİRİŞ

Gelişmekte olan dünyada özellikle son elli yılda nüfusun hızla artması nedeniyle trafik de kötü yönde bir değişim geçirmiştir. Trafikin artmasına bağlı olarak trafik sorunları da artmaya başlamıştır. Trafik kazaları ve trafik güvenliği ile ilgili sorunlar bu alanda başlıca sorunlardan bazılarıdır.

Trafik kazaları rasgele oluşmalarına rağmen, sıklığı birçok faktör tarafından etkilenmektedir. Kazalar istatistiksel bir değer olarak ele alındığında, hiç kimse oluşabilecek bir trafik kazasının yerini ve zamanını tahmin edemez. Kazaya etki eden faktörler ise başlıca insan, taşıt, yol ve çevresi başlıkları altında toplanabilir. Trafik kazalarından görülecek zararın azaltılması, alınan önlemlerin etkinliğine bağlıdır. Başlıca önlem alanları;

- *Trafik Mühendisliği Faaliyetleri:* Güvenli ve standartlara uygun yol ve kavşak tasarımı, bölünmüş yol yapımı, otopark, yaya kaldırımı, üst ve alt geçit inşaatı, trafik kontrol merkezleri ile trafik düzenlenmesi vb.
- *Eğitim Faaliyetleri:* Trafik güvenliğinin sağlanması amacıyla başta yolu kullanan sürücü, yolcu ve yayalar olmak üzere, trafikle yetkili yetkisiz yani halkın eğitilmesi, bilgilendirilmesi, motive edilmesi vb.
- *Denetim ve Yasal Düzenleme Faaliyetleri:* Kural ihlallerini önlemeye yönelik dürüst, sık ve teknolojiye dayalı olarak trafik kontrolleri ile trafik güvenliğini geliştirici ve uygulanabilir yasal düzenlemelerin yapılması vb.
- *İlk Yardım ve Acil Sağlık Hizmetleri:* Kaza sonrası mağdurlardan hafif yaralananın ağır hale gelmemesi, ağır yaralının ölmesini engellemek için süratli ve kaliteli hizmet veren uzman personel ve donanımlı ambulansın kaza yerinde görev yapması vb. dir.

Trafik güvenliği ile ilgili olarak, Karpat ve Yılmaz [1], Türkiye'deki trafik kaza verilerini incelemişler ve trafik kaza verilerini kümelenme analizi yöntemi ile iller bazında oluş şekillerine göre kümelendirmişlerdir. Analizler sonucunda, Adana, Ankara, İstanbul ve İzmir'in taşıt sayısı, kaza sayısı, ölü sayısı ve yaralı sayısı bakımından bir ilden oluşan farklı kümelerde yer aldığı, ayrıca Antalya, Bursa ve Konya'nın ortak bir kümede toplandığı açıkça görülmüştür. Geurts ve Wets[2], çalışmalarında kara nokta ve kara bölgelerin analizinde kullanılan yöntem ve teknikleri literatür olarak gözden geçirmişler, kara noktaları tanımlama ve sıralamada kullanılan birkaç alternatif yöntem üzerinde durmuşlardır. Sonuç olarak Belçika, Danimarka ve Avustralya' da kara noktaları analiz etmede kullanılan bu yöntem ve teknikler hakkında genel bir bakış sunmuşlardır. Kara nokta iyileştirme programı, kazaların meydana gelişlerini duyarlı olarak ele alabilen bir güvenlik geliştirme programı tipidir. Aslında, bu tarz programların uygulaması, gelişen ülkelerde mevcut olmayan veya limitli olan, ilgili kaza verilerine ihtiyaç duymaktadır. Bunun için Kowtanapanich vd.[3], çalışmalarında bu engelin üstesinden gelen destekleyici bir yaklaşımı (halk katılımı yaklaşımı) ele almışlardır. Amaçlanan, Kaza Halk Katılım Programı çerçevesi sayesinde, kara noktaların tanımlanmasına katkıda bulunmada bir halk katılımı yaklaşımının nasıl kullanılabilceğini ispatlamaktır. Kardeşin ve Saplıoğlu [4], coğrafi bilgi sistemi yardımı ile Isparta kenti, kentiçi trafik kazalarını analiz etmişler, çalışma sonucunda son beş yılda Isparta' da kavşaklar ile cadde ve sokaklarda oluşmuş

kaza sayılarını tespit ederek, kavşaklarda beş yılda kayıtlara geçen 1016 kazanın kara noktalarını belirlemişlerdir.

Kazaların ayrıık olarak analizinin yanında yoğunlaştığı bölgelerin belirlenmesi ve analizi de trafik güvenliği açısından oldukça önemlidir. Fakat kazaların yoğunlaştığı bölgeleri belirlemek ve bir merkez tespit etmek kolay değildir. Yoğunluk değeri veya yoğunlaştığı kesimler sayısal olarak tespit edilse dahi konumsal yerleşim (yerleştirme) ve yaklaşık bir merkezin oluşturulması güçtür ve pek çok deneme yapmayı gerektirir. İşte bu sorunların üstesinden gelmek ve kara nokta merkezlerini belirlemek amacıyla kümelenme analizi yaklaşımı kullanılmıştır. Ayrıca kara noktaların yalnızca konumsal olarak tespit edilmesi yeterli değildir. Bu tespitin ardından kara nokta merkezlerinde kazaya neden olan unsurlar da irdelenmelidir.

Bu çalışmanın genel amacı kara yollarında trafik kazalarının yoğun görüldüğü noktaların (kara noktaların) yerlerinin ve sayılarının kümelenme analizi ile tespit edilerek, bu noktaların iyileştirilmesi ve azaltılması yönünde çeşitli çözüm önerilerinin getirilmesidir.

2. KÜMELENME ANALİZİ

Kümelenme analizi, bir araştırmada incelenen birimleri aralarındaki benzerliklerine göre belirli gruplar içinde toplayarak sınıflandırma yapmayı, birimlerin ortak özelliklerini ortaya koymayı ve bu sınıflar ile ilgili genel tanımlamalar yapmayı sağlayan bir yöntemdir [5]. Analiz sonucu elde edilen kümeler yüksek düzeyde küme içi homojenlik ve yüksek düzeyde kümeler arası heterojenlik gösterirler [6].

Kümelenme analizi; birimleri p değişkene göre hesaplanan ve benzerlik ölçüsü olarak kullanılan bazı ölçüler kullanılarak homojen gruplara bölmek amacıyla kullanılır. Bu amaçlar dört grupta toplanabilir;

- 1) n sayıda birimi, nesneyi, oluşumu, p değişkene göre saptanan özelliklerine göre olabildiğince kendi içinde türdeş (homojen) ve kendi aralarında farklı (heterojen) alt gruplara (küme) ayırmak.
- 2) p sayıda değişkeni, n sayıda birimde saptanan değerlere göre ortak özellikleri açıkladığı varsayılan alt kümelere ayırmak ve ortak faktör yapıları ortaya koymak.
- 3) Hem birimleri hem değişkenleri birlikte ele alarak, ortak n birimi p değişkene göre ortak özellikli alt kümelere ayırmak.
- 4) Birimleri, p değişkene göre saptanan değerler için, izledikleri biyolojik ve tipolojik sınıflamayı ortaya koymak (taksonomik sınıflandırma yapmak).

Kümelenme işleminin uygulandığı veri setindeki her bir veriye nesne adı verilir ve kümelenme analizi, bu nesnelere benzerlik esasına göre aynı kümelere toplar. Bu nesnelere iki boyutlu düzlem üzerinde noktalarla gösterilir. Benzerlikler, nesne çiftleri arasındaki uzaklığın ölçüsüdür. X veri matrisinde yer alan n birimin p değişkene göre uzaklıkları, uzaklık matrisi adı verilen D matrisi ile gösterilir. D matrisinin elemanları d_{ij} biçiminde gösterilir. Birimlerin birbirleri ile olan benzerlik düzeyleri, benzerlik matrisi Sim, elemanları da sim_{ij} biçiminde gösterilir. Birimlerin benzerlikleri $sim_{ij} = 100 (1 - d_{ij} / \max d_{ij})$ biçiminde hesaplanır.

Trafik Kaza Verilerinin Kümelenme Analizi Yöntemi ile Modellenmesi

Kümelenme analizinin uygulama aşamaları;

Veri matrisinin belirlenmesi: Birim ya da değişkenlerin doğal gruplamaları hakkında kesin bilgilerin bulunmadığı popülasyonlardan alınan n sayıda birimin p sayıda değişkene ilişkin gözlemlerin elde edilmesidir.

Benzerlik ya da farklılık matrisinin belirlenmesi: Birimlerin/değişkenlerin birbirleri ile olan benzerliklerini ya da farklılıklarını gösteren uygun bir benzerlik ölçüsü ile birimlerin/değişkenlerin birbirlerine uzaklıklarının hesaplanmasıdır.

Kümelere ayırma: Uygun kümelenme yöntemi yardımı ile benzerlik/farklılık matrislerine göre birimlerin/değişkenlerin uygun sayıda kümelere ayrılmasıdır.

Yorumlama: Elde edilen kümelerin yorumlanması ve bu kümelenme yapısına dayalı olarak kurulan hipotezlerin doğrulanması için gerekli analitik yöntemlerin uygulanmasıdır.

Yaygın olarak kullanılan kümelenme yöntemleri birimler arasındaki uzaklıklara dayanan benzerlik ya da benzemezlik matrisine göre işlem yaptıklarından, farklı kümelenme yöntemleri farklı uzaklık ölçülerine göre farklı sonuçlar verebilmektedir. Ayırmaya dayanan kümelenme yöntemleri her veri setinin her bir birimini bir ve yalnızca bir kümeye ayırır. Böylelikle aşamalı ya da aşamalı olmayan kümelenme yöntemleri her bir birim için kesin karar alırlar ve bir kümeye atarlar. Sonuçları itibariyle yaklaşık aynı sonuçları veren kümelenme algoritmalarında bazı birimlerin farklı kümelere yer aldığı gözlenebilmektedir [7]. Bir diğer ifadeyle birimlerin küme üyeliklerinde bir kararsızlık ortaya çıkmaktadır. Bu tip durumlarda birimlerin küme üyeliklerinde bir bulanıklık söz konusu olmaktadır. Bulanık Kümelenme (Fuzzy Clustering) Yöntemi bu tip durumları tanımlamak için daha iyi bir yöntemdir. Çalışma kapsamında K-ortalamlar ve Bulanık-C ortalamlar yöntemleri kullanılmıştır.

2.1 K-Ortalamlar Kümelenme Yöntemi

K-ortalamlar yöntemi, endüstriyel ve bilimsel anlamda kullanılan en popüler kümelenme yöntemlerinden birisidir. Adı, küme merkezi (centroid) olarak adlandırılan, k tane kümenin her birini temsil eden noktaların ortalamasından gelir. K-ortalamlar kümelenme analizi öklit uzaklığını kullanır. Araştırmacı arzulanan küme sayısını baştan belirlemek zorundadır. K-ortalamlar algoritması aşağıdaki amaç fonksiyonunu minimize etmeye çalışır;

$$J = \sum_{j=1}^k \sum_{i=1}^{n_j} \|x_{ij} - c_j\|^2 \quad (2.1)$$

Burada, $\|x_{ij} - c_j\|^2$, x_{ij} verisi ile c_j küme merkezi arasındaki mesafedir. J, n verinin kendi küme merkezlerinden olan toplam uzaklığıdır.

K-ortalamlar metodu, ilk önce n adet nesneden k adet nesne seçer ve bu nesnelerin her biri, bir kümenin merkezini veya orta noktasını temsil eder. Geriye kalan nesnelere her biri kendisine en yakın olan küme merkezine göre kümelere dağılırlar. Ardından her küme

için ortalama hesaplanır ve hesaplanan bu değer o kümenin yeni merkezi olur. Bu işlem tüm nesnelere kümelere yerleşinceye kadar devam eder [8].

2.2 Bulanık C-Ortalamlar Yöntemi

Klasik kümelendirme yönteminde, bir eleman bir kümeye ya aittir (üyelik= 1) ya da ait değildir (üyelik= 0). Gerçekte bir eleman bir kümeye ne tam aittir ne de değildir. Yani bu elemanın o küme için bir aittir derecesi (üyelik değeri) olmalıdır. Bu üyelik değeri 0 ile 1 arasında sonsuz değer alabilmektedir ve bu üyelik değerlerinin toplamı daima 1'e eşittir. Böylelikle eleman **en yüksek üyelik katsayısına** sahip olduğu kümeye atanır. Birbirine çok benzeyen birimler aynı kümede yüksek üyelik ilişkisine göre yer alırlar. Bulanık algoritmalarda, bir eleman kümelendirirken veya sınıflandırılırken elemanın sınıfını belirlemenin yanında o sınıfa ne kadar ait olduğuna dair bir bilgi de verildiğinden bulanık algoritmalar, klasik algoritmalara oranla daha fazla bilgi içerirler [9].

Bulanık kümelendirme yönteminde, ilk olarak (2.2) eşitliğine göre U üyelik matrisine rasgele değer atanır.

$$\sum_{i=1}^c u_{ij} = 1, \forall_j = 1, \dots, n \quad (2.2)$$

Bulanık c-ortalamlar metodunda kullanılan farklılık fonksiyonu,

$$J(U, c_1, c_2, \dots, c_c) = \sum_{i=1}^c J_i = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2 \quad (2.3)$$

u_{ij} : 0 ile 1 arasındadır,

c_i : i. kümenin merkezi,

d_{ij} : i. merkez c_i ile j. veri noktası arasındaki öklit mesafesi,

m : ağırlıklandırma faktörü (minimum değeri 1 alınır, genel olarak 2 değeri kullanılır).

Farklılık fonksiyonunu minimum yapmak için iki koşul vardır. Bu koşullarda (2.4) ve (2.5) eşitliklerinde verilmiştir.

$$c_i = \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m} \quad (2.4)$$

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}} \right)^{2/(m-1)}} \quad (2.5)$$

Bulanık c-ortalamalar algoritması Bezdek [10] tarafından şu şekilde verilmiştir;

1. (2.2) denklemini sağlayacak şekilde U üyelik matrisine rasgele değerler ata,
2. (2.4) denklemini kullanılarak c_i merkezleri hesapla,
3. (2.3) denklemini kullanarak merkezler ile veri noktaları arasındaki farklılığı hesapla. Önceki iterasyon üzerine artışı eşik değerin altındaysa dur,
4. (2.5) denklemini kullanarak yeni bir U matrisi hesapla. Adım 2'ye git.

2.2.1 Doğrulama

Bulanık kümelene medeki ana problemlerden birisi küme sayısının baştan belirlenmesi gerekliliğidir. Farklı başlangıç küme sayısı seçimleri, farklı kümelene melerin ortaya çıkmasına sebep olur. Onun için, kümelene me analizinden sonra her bir bulanık bölmenin doğrulamasının yapılması gereklidir. Bu çalışmada birkaç kümelene me indeksi kullanılarak optimum başlangıç küme sayısı tespit edilmiş ve bu küme sayısına göre veri seti bulanık kümelene meye tabi tutulmuştur. Bu indeksler;

- Ayırma Katsayısı (Partition Coefficient, PC)
- Sınıflandırma Entropisi (Classification Entropy, CE)
- Ayırma İndeksi (Partition Index, SC)
- Ayırıştırma İndeksi (Separation Index, S)
- Xie ve Beni İndeksi (Xie and Beni's Index, XB)
- Dunn İndeksi (Dunn's Index, DI)

Ayırma katsayısı (PC), iki bulanık kümenin üst üste gelme miktarını ölçer. Verilerin özelliklerine direk bağlantısının eksik olması bu indeksin dezavantajıdır. Optimum küme sayısı, bu indeksin maksimum değerine karşı gelen küme sayısıdır. İndeks aralığı $[1/c, 1]$.

$$PC(c) = \frac{1}{N} \sum_{i=1}^c \sum_{k=1}^N (u_{ik})^2 \quad (2.6)$$

Sınıflandırma entropisi (CE), küme ayrımlarının bulanıklığını ölçer. Optimum küme sayısı, bu indeksin minimum değerine karşı gelen küme sayısıdır. İndeks aralığı $[0, \log a(c)]$.

$$CE(c) = -\frac{1}{N} \sum_{i=1}^c \sum_{k=1}^N u_{ik} \log_a a u_{ik} \quad (2.7)$$

Ayrırma indeksi (SC), toplam sıklık ve küme ayırma oranıdır. SC kullanışlı bir indekstir ve bu indeksin düşük değeri, daha iyi bir ayırmanın ispatıdır.

$$SC(c) = \sum_{i=1}^c \frac{\sum_{k=1}^N (u_{ik})^m \|x_k - v_i\|^2}{N_i \sum_{k=1}^c (u_{ik})^m \|v_k - v_i\|^2} \quad (2.8)$$

Ayrışma indeksi (S), geçerli bölünme için minimum uzaklık ayrışmasını kullanır.

$$S(c) = \frac{\sum_{i=1}^c \sum_{k=1}^N (u_{ik})^2 \|x_k - v_i\|^2}{N \min_{i,k} \|v_k - v_i\|^2} \quad (2.9)$$

Xie ve Beni İndeksi (XB), kümeler arasındaki toplam varyasyon oranının ve kümelerin ayrışma oranının miktarını belirlemeyi amaçlar. Optimum küme sayısı, bu indeksin minimum değerine karşı gelen küme sayısıdır.

$$XB(c) = \frac{\sum_{i=1}^c \sum_{k=1}^N (u_{ik})^m \|x_k - v_i\|^2}{N \min_{i,k} \|x_k - v_i\|^2} \quad (2.10)$$

Dunn indeksi (DI), sıklık teşhisini ve ayrık küme kullanımını önerir.

$$DI(c) = \min_{i \in C} \left\{ \min_{k \in C, i \neq k} \left\{ \frac{\min_{x \in C_i, y \in C_k} d(x, y)}{\max_{k \in C} \left\{ \max_{x, y \in C} d(x, y) \right\}} \right\} \right\} \quad (2.11)$$

3. TRAFİK KAZA VERİLERİNİN ANALİZİ

3.1 Konum (x, y) Verisinin K-Ortalamalar ile Kümelenme Analizi

Konum verisi, kazanın meydana geldiği noktanın, MAPINFO yazılımında bulunan sokak tanımlama sisteminin kullanılmasıyla elde edilen veridir. Bütün detayları içeren trafik kaza verisi, Coğrafi Bilgi Sistemi (CBS) veritabanında kaydedilmiştir. Coğrafi Bilgi Sistemi (CBS), ülke veya yeryüzü referans sistemine bağlı yer-referanslı (geo-referenced data) verileri girme, arşivleme, analiz etme ve görüntüyle çıktı elde etmek için geliştirilmiş bilgisayar tabanlı bir sistemdir [11]. Şekil 1 Denizli şehri GIS haritasındaki örnek yönlendirilmiş veriyi göstermektedir [12].

Trafik Kaza Verilerinin Kümelenme Analizi Yöntemi ile Modellenmesi



Şekil 1 GIS' de Denizli haritasındaki örnek yerleştirilmiş trafik kaza verisi

Konum verisinin k-ortalamlar kümelenme analizinde optimum küme sayısı bilinmediği için başlangıç küme sayısı tahmini olarak 15 seçilmiştir. Böylelikle elde edilen minimum iterasyonlar incelendiğinde, optimum küme sayısı olarak, değişim yüzdesinde belirgin bir düşüşün görülmediği küme sayısı olan 7 seçilebilir. Bu seçilen optimum küme sayısına göre konum verisi tekrar analiz edilerek Tablo 2'deki değerler elde edilir.

Tablo 1 Konum verisinin k-ortalamlar kümelenme analizi sonucu minimum iterasyonu

Minimum İterasyon			
İterasyon No	Küme No	Değişim Yüzdesi	Yüzdenin Çizgi Grafiği
2	2	65,36	
4	3	43,46	
9	4	31,43	
10	5	24,45	
13	6	18,81	
16	7	15,56	
20	8	13,56	
22	9	12,2	
27	10	11,02	
29	11	9,54	
31	12	8,97	
36	13	7,92	
38	14	7,22	
40	15	6,82	

Tablo 2 Konum verisinin k-ortalamalar kümelene analizi sonucu iterasyonu

İterasyon			
İterasyon No	Küme No	Değişim Yüzdesi	Yüzdenin Çizgisel Grafiği
1	2	74,33	
2	2	65,36	
3	2	65,36	
4	3	43,46	
5	3	43,59	
6	3	43,59	
7	4	32,6	
8	4	32,6	
9	4	32,48	
10	5	24,45	
11	5	25,16	
12	5	25,82	
13	6	18,81	
14	6	18,81	
15	6	18,81	
16	7	15,56	
17	7	15,56	
18	7	15,56	

Kullanılan bilgisayar programında her bir küme, rastgele başlangıç kısmında girilen sayı kadar iterasyona tabi tutulur. İterasyon tablosu, bu kümelere ait iterasyonları ve değişim yüzdelerini verir. Elde edilen küme merkezleri ise Tablo 3' deki gibidir.

Tablo 3 Konum verisinin k-ortalamalar kümelene analizi sonucu küme merkezleri

Küme No	Değişkenler		Veri Sayısı
	x	y	
1	29,084884	37,787184	2480
2	29,041337	37,772047	302
3	29,102469	37,798443	973
4	29,071570	37,777544	835
5	29,096437	37,780394	1589
6	29,094854	37,759552	1274
7	29,085393	37,772304	1645

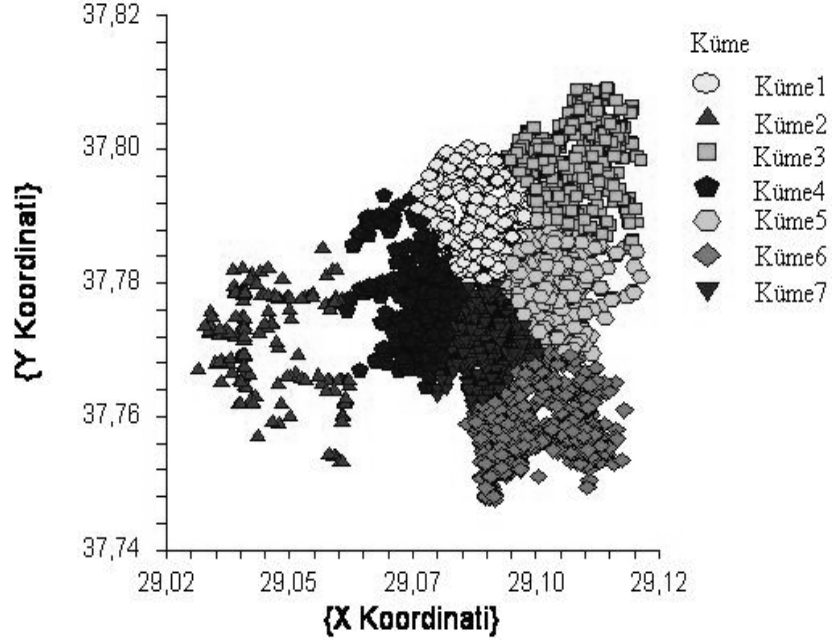
Trafik Kaza Verilerinin Kümelenme Analizi Yöntemi ile Modellenmesi

Elde edilen tüm konum verisine ait uzaklıklar matrisinin boyutu çok büyük olduğu için, Tablo 4’ de sadece 10 elemana ait uzaklık matrisi örnek olarak gösterilmiştir.

Tablo 4 Konum verisinin k-ortalamlar kümelenme analizi sonucu uzaklıklar matrisi

Uzaklıklar Matrisi								
Sıra	Küme	Uzaklık1	Uzaklık2	Uzaklık3	Uzaklık4	Uzaklık5	Uzaklık6	Uzaklık7
1	6	2,0881	3,6918	3,211	1,7961	1,6449	0,5748	0,8612
2	5	1,0082	4,4046	1,3439	2,0275	0,2272	1,9605	1,2735
3	3	2,0863	5,6177	0,4475	3,3893	1,9861	3,6739	2,9819
4	7	1,8124	2,7705	3,2993	0,985	1,8962	1,5175	0,8113
5	1	0,0738	3,5649	1,6572	1,3025	1,1259	2,4909	1,2961
6	1	0,2271	3,7723	1,5167	1,4466	0,8389	2,2988	1,1852
7	3	1,3991	4,9876	0,4819	2,6653	1,1069	2,8068	2,1237
8	2	4,0068	0,4384	5,6096	2,8114	4,7441	4,682	3,8387
9	5	1,0936	3,9922	1,9838	1,6461	0,438	1,401	0,6651
10	1	0,1798	3,5293	1,7573	1,2005	0,9716	2,2525	1,0577

Konum verisinin kümelere dağılımı ise Şekil 2’deki gibi elde edilmiştir.



Şekil 2 Konum verisinin k-ortalamlar küme analizi sonucu küme grafiği

Bu analizde küme merkezlerine göre elde edilen kara noktalar ise Tablo 5’ de gösterilmiştir.

Tablo 5 Konum verisinin k-ortalamlar kümelenme analizi sonucu belirlenen kara noktalar

Kavşak veya Cadde (Kara Nokta)	Küme No	Kaza Sayısı
İstasyon Kavşığı	1	2480
Çınar Mey	7	1645
Deliktaş Kavşığı (Emniyet Müd.)	5	1589
Fatih Kavşığı	6	1274
Ankara Bulvarı (Sevindik Alt Geçidi)	3	973
Özay Gönüm Meydanı	4	835
Yeni Adliye	2	302

3.2 Konum (x, y) Verisinin Bulanık Kümelenme Analizi

Daha önce de bahsedildiği üzere bulanık kümelenme metodunda bir elemanın bir kümeye aitlik derecesi (üyelik değeri) vardır. Bu üyelik değeri 0 ile 1 arasında sonsuz değer alabilmektedir ve bu üyelik değerlerinin toplamı daima 1’e eşittir. Bundan dolayı eleman **en yüksek üyelik değerine** sahip olduğu kümeye atanır.

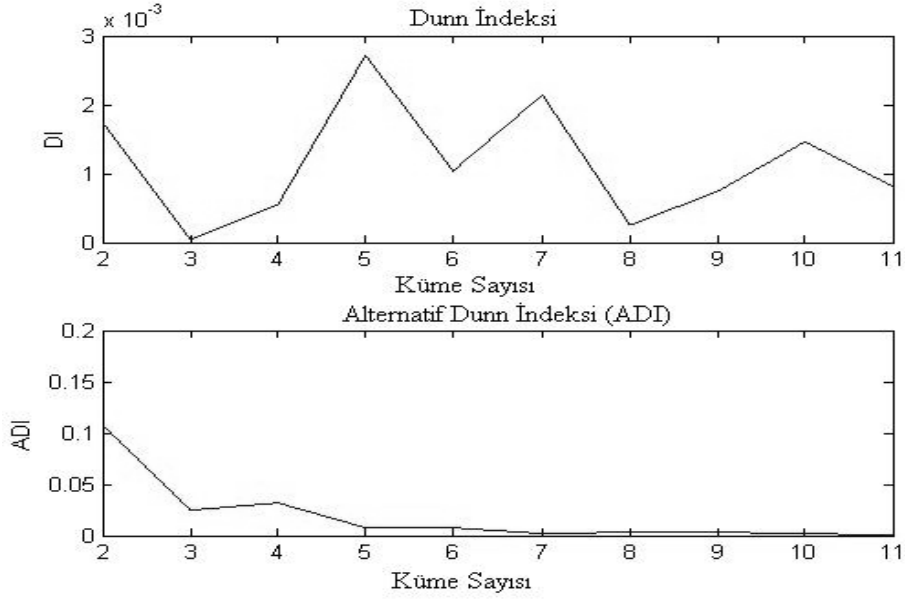
Konum verisinin bulanık kümelenme analizi için ilk olarak NCSS programının deneme sürümü kullanılmış, fakat veri boyutunun büyüklüğü nedeniyle analizin çok uzun sürmesinden ve kümelenmenin grafiğinin çizilememesinden dolayı NCSS programının kullanımından vazgeçilerek analize MATLAB programı ile devam edilmiştir.

Burada da optimum küme sayısı bilinmediğinden dolayı, veriler iki kere analize tabi tutulmuştur. İlk analizde, elde edilen *ayırma katsayısı (partition coefficient, PC)*, *sınıflandırma entropisi (classification entropy, CE)*, *ayırma indeksi (partition index, SC)*, *ayırıştırma indeksi (separation index, S)*, *Xie ve Beni indeksi (Xie and Beni’s Index, XB)*, *Dunn indeksi (Dunn’s index, DI)* değerlerine bakılarak optimum küme sayısına karar verilmiştir. Fakat ayırma katsayısının maksimum, sınıflandırma entropisinin minimum değerinin görüldüğü küme sayısının genelde iki veya üç olmasının mantıksız olmasından dolayı ve ayırma katsayısının, verilerin özelliklerine direk bağlantısının eksik olmasının bir dezavantaj olmasından dolayı, bu iki katsayının değerlendirilmesinden vazgeçilmiştir.

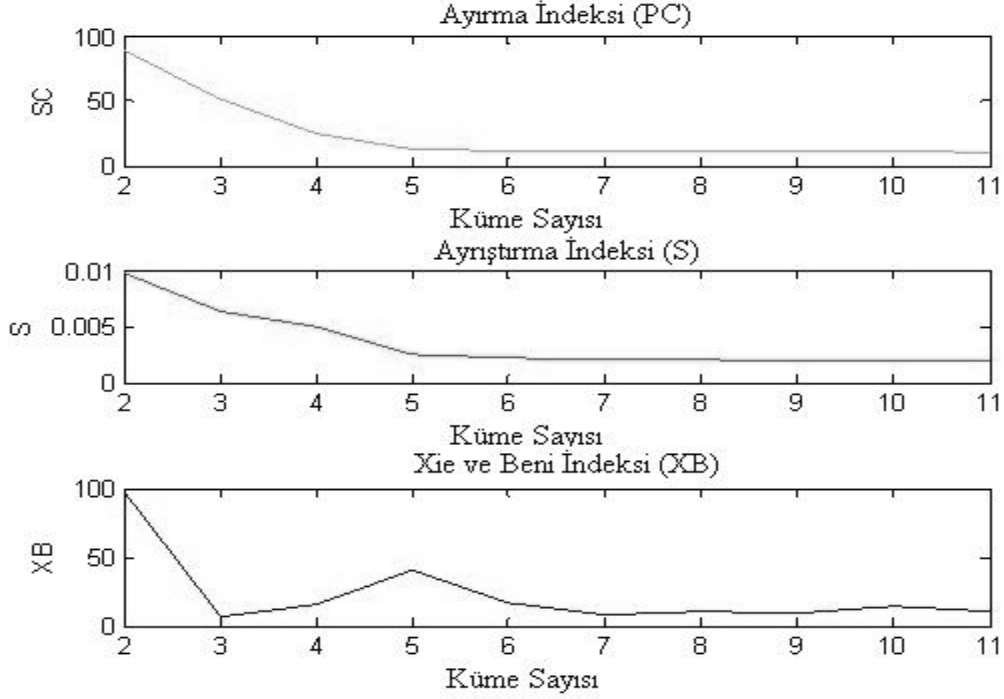
Optimum küme sayısı ise SC, S ve XB’ in minimum değerlerine denk gelen 11 olarak seçilmiştir. Tablo 6, Şekil 3 ve Şekil 4’te doğrulama katsayılarının küme sayısına göre değişimleri verilmiştir.

Tablo 6 Doğrulama katsayıları

Küme No	SC	S	XB	DI
2	88,6370	0,0097	96,548	0,0017342
3	51,2240	0,0064	10,879	2,24E-05
4	24,6720	0,0050	16,086	0,00054551
5	12,5210	0,0025	40,156	0,0027215
6	11,7100	0,0022	17,276	0,0010383
7	11,6490	0,0020	14,228	0,002147
8	11,1400	0,0021	10,769	0,00025443
9	10,9860	0,0020	11,737	0,00073721
10	10,8830	0,0020	14,86	0,0014641
11	10,6270	0,0019	10,762	0,00079156



Şekil 3 Dunn indeksi (DI) ve alternatif Dunn indeksi (ADI)



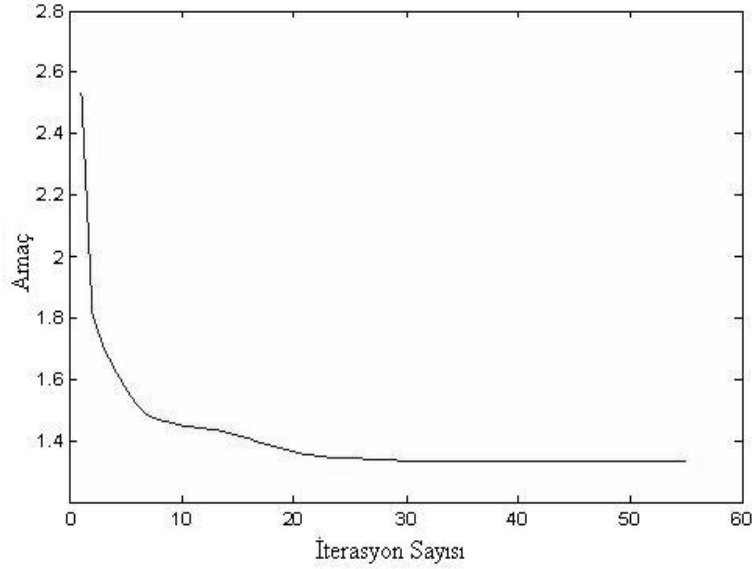
Şekil 4 Ayrırma indeksi (PC), ayrıştırma indeksi (S), Xie ve Beni indeksi (XB)

Amaç fonksiyonunun iterasyon adımları Tablo 7' deki gibidir;

Tablo 7 Amaç fonksiyonu iterasyon adımları

İt. No	Amaç Fonk.	İt. No	Amaç Fonk.	İt. No	Amaç Fonk.	İt. No	Amaç Fonk.	İt. No	Amaç Fonk.
1	2,5314	12	1,4416	23	1,3477	34	1,3339	45	1,3333
2	1,8145	13	1,4358	24	1,3447	35	1,3337	46	1,3333
3	1,708	14	1,4278	25	1,3424	36	1,3336	47	1,3333
4	1,6376	15	1,4176	26	1,3406	37	1,3335	48	1,3333
5	1,5686	16	1,4064	27	1,339	38	1,3334	49	1,3333
6	1,5145	17	1,3956	28	1,3376	39	1,3334	50	1,3333
7	1,484	18	1,3855	29	1,3365	40	1,3334	51	1,3333
8	1,4675	19	1,3759	30	1,3356	41	1,3334	52	1,3333
9	1,4578	20	1,3665	31	1,3349	42	1,3334	53	1,3333
10	1,4513	21	1,3583	32	1,3345	43	1,3333	54	1,3333
11	1,4464	22	1,352	33	1,3341	44	1,3333	55	1,3333

Trafik Kaza Verilerinin Kümelenme Analizi Yöntemi ile Modellenmesi



Şekil 5 Amaç fonksiyonunun değişim grafiği

Tablo 7 ve Şekil 5'den görüldüğü üzere 30. iterasyondan sonra amaç fonksiyonunda belirgin bir değişim elde edilememiştir.

Konum veri setinin çok büyük olmasından dolayı her bir elemanın üyelik derecesini göstermek yerine, örnek olarak 5 elemanın üyelik dereceleri Tablo 8' de gösterilmiştir. Elemanlar yüksek üyelik değerine sahip oldukları kümeye aittirler. Dolgulu hücreler hangi elemanın hangi kümeye ait olduğunu göstermektedir.

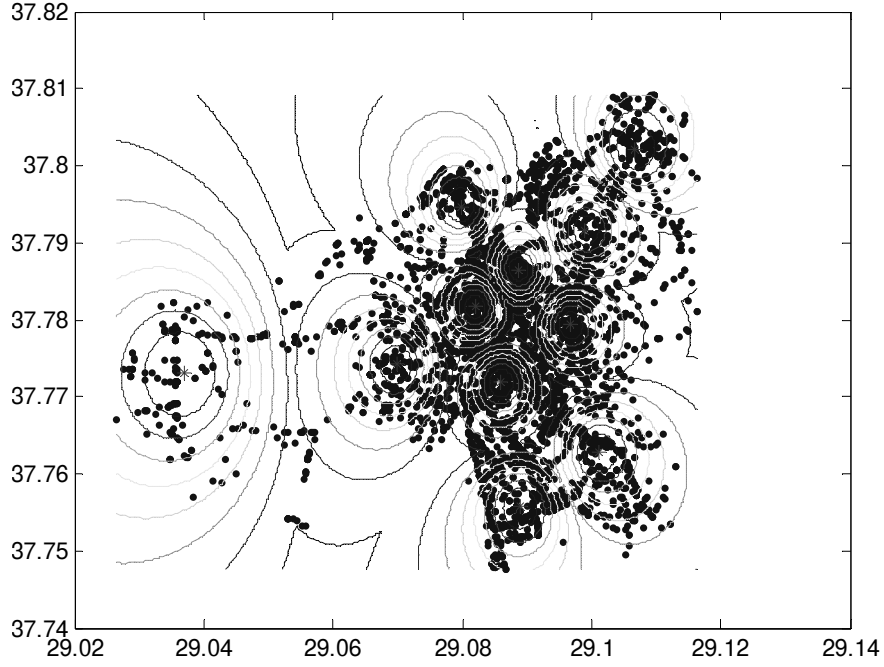
Tablo 8 Eleman üyelik dereceleri

Küme No	Eleman No				
	1	2	3	4	5
1	0,047315	0,037842	0,00019467	0,11425	0,22417
2	0,041651	0,010648	9,3789e-005	0,27779	0,024679
3	0,03314	0,10387	0,00035198	0,053878	0,44852
4	0,20156	0,035137	0,00014765	0,25527	0,037229
5	0,053094	0,62485	0,00031696	0,051701	0,042116
6	0,010114	0,019192	0,99724	0,013447	0,013664
7	0,017486	0,019957	0,00026063	0,036372	0,14466
8	0,0071925	0,0023314	3,4913e-005	0,017696	0,0038993
9	0,019971	0,11484	0,001154	0,025915	0,041368
10	0,45523	0,010736	8,0098e-005	0,10625	0,0091146
11	0,11325	0,020602	0,00012466	0,047427	0,010577

Elde edilen kmelerin merkezleri ise Tablo 9' daki gibidir.

Tablo 9 Konum verisinin bulanık kmelenme analizi sonucu kme merkezleri

Kme No	Kme Merkezleri	
	x	y
1	29,082	37,782
2	29,07	37,774
3	29,089	37,787
4	29,086	37,772
5	29,097	37,779
6	29,107	37,802
7	29,08	37,795
8	29,037	37,773
9	29,099	37,792
10	29,089	37,756
11	29,101	37,763



řekil 6 Konum verisinin bulanık kmelenme analizi sonucu kme grafiđi

Trafik Kaza Verilerinin Kümelenme Analizi Yöntemi ile Modellenmesi

Denizli kentine ait trafik kaza verisinin bulanık kümelenme analizi sonucundaki küme merkezlerine göre elde edilen kara noktalar ve kaza sayıları ise Tablo 10' daki gibidir.

Tablo 10 Konum verisinin bulanık kümelenme sonucu belirlenen kara noktalar

Kavşak veya Cadde (Kara Nokta)	Küme No	Kaza Sayısı
İstasyon Kavşağı	3	1389
Çınar Meydanı	4	1349
Vilayet Önu Kavşağı	1	1303
Deliktaş Kavşağı	5	1253
Demokrasi Meydanı	10	662
Özay Gönlüm Meydanı	2	648
Fatih Kavşağı	11	631
25. Sokak	9	605
İzmir Bulvarı (Zafer Katrancı Köprülü Kav.)	7	509
Sevindik	6	501
Yeni Adliye	8	248
	TOPLAM	9098

4. SONUÇLAR

Yapılan çalışmada, Denizli kenti 2004, 2005 ve 2006 yıllarına ait trafik kaza verisinin kümelenecek kara noktaların belirlenmesi amaçlanmıştır. Bu amaçla, trafik kazalarının meydana geldiği noktaların koordinat verisi k-ortalamlar ve bulanık kümelenme analizine tabi tutularak, küme merkezlerine denk gelen kara noktalar belirlenmiştir.

Çalışmadaki analizler sonucunda K-ortalamlar yöntemi ile 7, Bulanık-C ortalamlar yöntemi ile 11 adet kara nokta tespit edilmiştir. Buna göre Bulanık-C ortalamlar yöntemi ile daha duyarlı sonuç elde edildiği ifade edilebilir. K-ortalamlar yönteminin özelliğinden dolayı bazı noktalar yakalanamamıştır. Oysa Bulanık-C ortalamlar yöntemi ile 4 kara nokta daha ortaya çıkarılmıştır. Bu noktalar Denizli kentindeki trafik hacmi ve kazalar açısından belirli özelliklere sahiptir.

İlk kara nokta olarak belirlenen, Gazi Mustafa Kemal Bulvarı ile İzmir Bulvarının kesiştiği İstasyon Kavşağı ve civarı, trafik yoğunluğunun fazla olması ve otobüs terminalinin yakınlığı açısından önem kazanmaktadır. Burada kazaların büyük çoğunluğu ışık ihlali ve sürücülerin agresif davranışları sebebiyle meydana gelmektedir. Bu konuda toplumumuzun trafik güvenliği lehinde daha duyarlı olması gerekmektedir. Bunun için sürücü, yolcu, yaya, yetkili veya yetkisiz herkesin eğitilmesi, bilgilendirilmesi ve motive edilmesi çok büyük önem arz etmektedir. Bunun haricinde bu konudaki denetim ve yasal düzenleme faaliyetlerinin geliştirilmesinin, kazaların azalmasında çok büyük bir etkisi olacaktır.

Ayrıca kavşağın eşdüzey kesiminin yeniden düzenlenmesi ile bazı kazaları önlemek mümkündür.

Bu çalışmada sonucunda, Denizli kentinin diğer bir önemli kavşağı olan ve trafik yoğunluğunun çok fazla olduğu Çınar Meydanı, ikinci kara nokta olarak belirlenmiştir. Bu kavşak, şehrin günlük trafiğinin büyük çoğunluğuna hizmet etmektedir. Bu kavşakta kaza sayısının fazla olması, sinyal zamanlamasındaki problemlerden, kavşak geometrik tasarımından, yaya trafiğinin çok fazla olmasından, bazı sürücülerin agresif araç sürmelerinden ve dikkatsizlikten dolayı kaynaklanabilir. Özellikle yaya yoğun bir kavşak olmasına rağmen sinyalizasyonda yaya akımları ile taşıtların filtre geçiş biçiminde düzenlenmesi doğru değildir. Burada, kavşak sinyalizasyonun tekrar tasarlanması, teknolojiyen mümkün olduğunca faydalanılarak kontrollerin yapılması ve gerekli cezai uygulamalarının yerine getirilmesi ile ilerde olabilecek kaza sayısında ciddi bir azalma sağlanabilir.

K-ortalamar ve Bulanık-C ortalamar yöntemleri ile ortak olarak belirlenen Deliktaş, Fatih ve Sevindik kavşakları Denizli kentini İzmir, Antalya ve Ankara kentlerine bağlayan karayolları üzerinde bulunmaktadır. Bu kavşaklar transit özellikli trafik ile kentiçi trafik akımları tarafından ortak olarak kullanılmaktadır. 2007 yılında bu kavşakların geometrileri değiştirilmiş ve transit akımları ayırmak amacıyla farklı düzeyli biçime dönüştürülmüştür. Çalışma kapsamında yapılan analizlerde 2004-2006 yıllarına ait veriler ve kavşak geometrileri dikkate alınmıştır. Kavşak geometrilerindeki değişimin kara nokta özelliğine etkisini incelemek amacıyla 2007 yılına ait trafik kaza verileri incelenmiş ve söz konusu kavşaklardaki kaza sayılarının arttığı tespit edilmiştir. Dolayısıyla kara nokta özelliklerini korudukları anlaşılmıştır. Bu kavşaklardaki kaza sayılarını azaltmak amacıyla özellikle akımların örülme alanlarının doğru biçimde tasarlanması, yeterli örülme mesafesinin bırakılması, ilgili işaretleme ve aydınlatmanın yapılması gibi önlemler alınabilir. Ayrıca kavşakların eşdüzey kesimlerinde yaya geçişleri için uygun tasarımın yapılması, nizami olmayan geçiş alanlarının ortadan kaldırılması kaza sayılarını azaltabilecektir. Bu kavşakların yakınında (önce ve sonrasında) yer alan akaryakıt istasyonu gibi bazı tesislerin bağlantı yollarının yeterince güvenli yapılması faydalı olabilecek diğer önlemler arasındadır.

Demokrasi Meydanı kavşağı ise kümelenme analizi sonucunda tespit edilen diğer bir kara noktadır. Bu kavşak sinyal denetimli eşdüzey yuvarlak ada biçimindedir. Gün içinde zirve saatler dışında trafik yoğunluğunun oldukça az olduğu söylenebilir. Buna rağmen gün boyunca tek bir sinyal devresi ve faz planı ile denetim sağlanmaktadır. Bu kavşakta bazı sürücüler tarafından kırmızı sinyalin ihlal edildiği gözlenmiştir. Trafik kazalarının önemli bir kısmı bu ihlal nedeni ile oluşmaktadır. Kavşağın sinyal süreleri ve faz planlarının yeniden düzenlenmesi ile kaza sayılarının azalması olasıdır.

Semboller

- p : Değişken sayısı
n : Birim sayısı
X : Veri matrisi

Trafik Kaza Verilerinin Kümelene Analizi Yöntemi ile Modellenmesi

D : Uzaklıklar matrisi

d_{ij} : Uzaklık değeri

Sim : Benzerlik matrisi

sim_{ij} : Benzerlik matrisi elemanları

J : Tüm verilerin kendi küme merkezlerinden toplam uzaklığını veren amaç fonksiyonu

x_{ij} : j kümesindeki i nolu elemanın koordinatı

c_j : j küme merkezinin koordinatı

k : Küme sayısı

$J(U,c)$: Farklılık fonksiyonu

u_{ij} : j kümesindeki i nolu elemanın üyelik derecesi

c_i : i kümesinin merkezi

d_{ij} : i küme merkezi c_i ile j veri noktası arasındaki öklit mesafesi

m : Ağırlıklandırma faktörü

$PC(c)$: Ayırma katsayısı

u_{ik} : k kümesinin i elemanının üyelik derecesi

$CE(c)$: Sınıflandırma entropisi

$SC(c)$: Ayırma indeksi

$S(c)$: Ayrışma indeksi

$XB(c)$: Xie ve Beni indeksi

$DI(c)$: Dunn indeksi

Teşekkür

Bu çalışmada 105G090 numaralı TÜBİTAK projesi kapsamında Denizli İl Emniyet Müdürlüğü'nden elde edilen trafik kaza verileri kullanılmıştır. Yazarlar katkılarından dolayı TÜBİTAK'a ve Denizli İl Emniyet Müdürlüğü'ne şükranlarını sunarlar.

Kaynaklar

- [1] Karpat, G. ve Yılmaz, V. (1997), Türkiye' deki Trafik Kazaları Oluş Şekillerinin, Kazanın Olduğu Yerdeki Trafik, Aydınlatma ve Yol Durumuna Göre İller Bazında İncelenmesi, Uluslararası Trafik ve Yol Güvenliği Kongresi, Ankara, 1-9

- [2] Geurts, K. and Wets, G. (2003), Black Spot Analysis Methods: Literature Review, Steunpunt Verkeersveiligheid Bij Stijgende Mobiliteit, RA-2003-07, Diepenbeek, 30s
- [3] Kowtanapanich, W., Tanaboriboon, Y. and Chadbunchachai, W. (2005), Applying Public Participation Approach To Black Spot Identification Process, IATSS Res (Int Assoc Traffic Saf Sci), 30 (1): 73-85
- [4] Karařahin, M. ve Saphloęlu, M. (2005), Coęrafi Bilgi Sistemi Yardımı İle Isparta İli Kentiçi Trafik Kaza Analizi, Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 12 (3): 321-332
- [5] řahin, M. ve Hamarat, B. (2002), G10 – Avrupa Birlięi ve OECD Ülkelerinin Sosyo-Ekonomik Benzerliklerinin Fuzzy Kümelenme Analizi İle Belirlenmesi, ODTÜ Uluslararası Ekonomi Kongresi VI, Ankara
- [6] Sharma, S. (1996), Applied Multivariate Techniques, John Wiley & Sons, Inc., New York, 493s
- [7] Hamarat, B. (1998), Türkiye’de Saęlık Açısından Homojen İl gruplarının Belirlenmesine İliřkin İstatiksel Bir Yaklaşım, Y. Lisans Tezi, Anadolu Üniversitesi Fen Bilimleri Enstitüsü, Eskişehir, 75s
- [8] Han, J. and Kamber, M. (2000), Data Mining Concepts and Techniques, Morgan Kaufmann Publishers, San Francisco, 550s
- [9] Höppner, F., Klawonn, F., Rudolf, K. and Runkler, T. (1999), Fuzzy Cluster Analysis, Wiley Publishers, Wolfenbüttel, 300s
- [10] Bezdek, J. C. (1973), Fuzzy Mathematics in Pattern Classification, PhD Thesis, Applied Math. Center, Cornell University, New York
- [11] Yıldırım, H. (1994), ‘Coęrafi Bilgi Sistemleri’, MAM Uzak Teknolojileri Bölümü Yayını, MAM-Tn 04, Marmara Arařtırma Merkezi
- [12] řekerler, A. (2008), ‘‘Trafik Kaza Verilerinin Kümelenme Analizi Yöntemi ile İncelenmesi’’, Yüksek Lisans Tez Çalışması, Pamukkale Üniversitesi Fen Bilimleri Enstitüsü, Denizli, 113 s.