
TÜRKİYE'DE SEKTÖRLERİN ÖLÜMLÜ KAZA TÜRLERİNE GÖRE KÜMELENMESİ: İKİLİ KÜMELEME YÖNTEMİ

Ahmet KOCATÜRK¹, Bülent ALTUNKAYNAK², H. Hasan ÖRKÜCÜ³

Öz

Türkiye'deki sektörlerin tehlike sınıfına göre kümeleneği olması, işkollarının ve işyerinin tehlike sınıf ve derecesine ait prim oranlarını ve tehlike derecelerini belirlenmesinde etkilidir. Bu amaçla kullanılan geleneksel kümeleme yöntemleri tek bir boyutu dikkate aldığından daha sınırlı bilgiler içermektedir. Bu sektörleri, benzer olan ölümlü kaza türleri açısından bir arada ele alarak kümelemek daha ayrıntılı sonuçlar verecektir. Bu sonuçları elde etmek için geleneksel kümeleme yöntemlerinden geliştirilerek bulunan ikili kümeleme yöntemleri ele alınmıştır. Bu çalışmada ikili kümeleme yöntemlerinden biri olan Bimax algoritması kullanılmıştır. Çalışmada ele alınan matris şeklindeki veri setinde, satırlar sektörleri, sütunlar ise ölümlü kaza türlerini göstermektedir. Bu çalışmada sektörler ve kaza türleri birlikte kümelenecek riskli sektör grupları oluşturulmuştur. Elde edilen ikili kümelere göre inşaat sektörleri ve taşımacılık sektörleri birlikte kümelenemiştir. Bu sektörlerin birlikte kümeleneği, motorlu taşıtlarla ve yüksek bir yerden düşme sonucu meydana gelen kazaların olmasından kaynaklanmaktadır. Bu gerçek uygulama sayesinde sektörlerle ait risk sınıflandırmaları yapılarak ölümlü kazalara karşı önlem alınmasına yönelik tedbirler uygulanabilir.

Anahtar Kelimeler: Sektörler, Ölümlü Kazalar, İkili Kümeleme, Bimax Algoritması.

JEL Sınıflandırması: C18, C38, C89.

CLUSTERING BY SECTORS ACCORDING TO THE TYPE OF FATAL ACCIDENTS IN TURKEY: THE BICLUSTERING METHOD

Abstract

Sectors in Turkey are clustered according to hazard class is effective in determining the premium rates and hazard ratings for hazard class and the degree of business and workplace. Traditional clustering methods used for his purpose have more limited information because they take into account a single dimension. Clustering these sectors together in terms of similar types of fatal accidents will give detailed results. In order to obtain these results, binary clustering methods developed by traditional clustering methods are discussed. In this study, Bimax algorithm which is one of the biclustering methods is used. In the dataset in the form of matrix covered in the study, the rows indicate the sectors and the columns show the types of the fatal accidents. According to the obtained biclusters, construction and transportation sectors cluster together. The clustering of these sectors is due to the accidents that occurred with motor vehicles and falling from a high place. In this study, sectors and types of accidents were clustered together to form risky sector groups. With this actual practice, measures can be taken to take precautions against fatal accidents by making risk classifications for the sectors.

Keywords: Sectors, Fattal Accident, Biclustering, Bimax Algorithm.

JEL Classification: C18, C38, C89.

¹ Arş. Gör., Gazi Üniversitesi, Fen Fakültesi, İstatistik Bölümü, ahmetkocaturk@gazi.edu.tr

² Doç. Dr., Gazi Üniversitesi, Fen Fakültesi, İstatistik Bölümü, bulenta@gazi.edu.tr

³ Doç. Dr., Gazi Üniversitesi, Fen Fakültesi, İstatistik Bölümü, hhorkcu@gazi.edu.tr

1. Giriş

Günümüzde ölümlü kazalar, kazaların önlenabilir olması gerçeği de dikkate alındığında kabul edilemez olaylardır. Bu tür istenemeyen olaylarda bir taraftan çalışanlar hayatını kaybederken, bir taraftan da o kazanın meydana geldiği ülke itibarını kaybetmektedir. İş kazaları ve bu kazaların sonucunda oluşan ölüm sayıları Türkiye için büyük ölçüde kayıplar meydana getirirken, iş yapan kişiler için de ciddi bir iş güvenliği problemi oluşturmaktadır. Son yıllarda ölümlü kazaların sayısında artma eğilimi görülmektedir. Bu nedenle iş ile ilgili ölümlerin azaltılması çok önemli bir yere sahip olduğu (Ceylan, 2014:1-6)'da belirtilmiştir. Bu önlenmeler; daha güvenli çalışma koşullarının oluşturulabilmesi için ilgili planlama ve politikaların belirlenmesinde etkilidir.

Türkiye'deki sektörlerin az tehlikeli, tehlikeli ve çok tehlikeli olarak sınıflandırılması, işyeri hekimi ve iş güvenliği uzmanı başta olmak üzere, işyerinin tabii olduğu tehlike sınıfına göre belirlenmektedir. İşyerlerinin iş sağlığı ve güvenliği açısından yer aldığı tehlike sınıfları "İşyeri Tehlike Sınıfları Listesinde" belirtilmiştir. Ölümlü kaza türleri bir arada ele alınarak sektörlerin kümelenmesi ile birlikte riskli sektör grupları oluşturularak tehlike sınıfları listesine daha etkin bir çözümün bulunması birçok nedenden dolayı faydalıdır.

Sektörlerdeki risklerin belirlenmesine yönelik en etkin tekniklerin belirlenmesi, kullanılması ve belirlenen risklerin sınıflandırılması, bu aşamanın amaçları olarak belirlenmiştir. Risklerin belirlenmesinde farklı bilgi toplama ve değerlendirme teknikleri kullanmak mümkündür. Sektörler için işe yarar sonuçlar elde edilmesinde çok sayıda ve çeşitli teknikler vardır. Önemli olan nokta, işe yarayan ve kullanılabilir tekniklerin seçilmesidir. Tekniklerinin seçilmesinden sonra risklerin, kaynakları ile birlikte sınıflandırılması yapılmaktadır. Bu çerçevede, tehlike sınıflandırması belirli bir hiyerarşik yapı içerisinde yapılmalıdır. Sınıflandırma işlemlerinde veri madenciliği teknikleri yaygın olarak kullanılmaktadır.

Sektörlerin kaza türleri ile birlikte kümelenmesini sağlayacak yöntem klasik kümeleme yöntemlerinden farklı olarak ikili kümeleme (biclustering) yöntemidir. Çünkü sektörleri bir boyutta kümelemek yerine, daha tutarlı sonuçlar elde etmek için ölümlü kaza türlerini de dikkate alınması gerekir. Bu şekilde yapılan yöntem ile sonuçlardan daha ayrıntılı bilgiler elde edilir. Amaç, birbiriyle ilişkili kaza türlerini dikkate alarak sektörleri kümelemektir. Bu sayede kümelenen sektörlerin tehlike sınıflarını belirlerken, aynı zamanda hangi sektör kümelerinin hangi kaza kümeleriyle ilişkili olduğu da bulunmuş olur.

Bu çalışmada ele alınan veri seti, Türkiye'de 2011 yılında gerçekleşen, ölümlü kazalar için 1, diğer durumlar için 0 değerlerinden oluşan ikili matris yapısındadır. Matris satırları sektörleri, sütunları ise ölümlü kaza türlerini göstermektedir. Çalışmanın temel amacı sektörler ve kaza türlerini aynı anda kümeleyerek veri matrisinin alt matrislere ayrıştırılmasını sağlamaktır. İkili kümeleme algoritmaları içerisinde en etkin algoritma olarak Bimax yöntemi çalışmanın ikinci bölümünde ayrıntılı bir şekilde verilmiştir. Üçüncü bölümde verilerin analizi yapılmıştır. Dördüncü bölümde ise bu analizin sonuçları ikili kümeleme grafikleri ile verilmiştir. Son bölümde ise elde edilen sonuçlar yorumlanarak tartışılmıştır.

2. Yöntem

Kümeleme yöntemlerinde tek yönlü kümeleme yapıldığı için bazı amaçları karşılamamaktadır. Örneğin gen ifade verisi analizinde belirli durumlara göre gen ifade değerlerini kullanarak genleri, belirli genlere bağlı olarak durumları kümelemektir. Bu yüzden kümeleme yöntemi (sadece genleri veya sadece durumları kümeleme) gen ifade veri analizinde belirtilen amaçlara karşılık vermemektedir.

Bir veri seti ile çalışırken, sadece örneklerin analizini değil, aynı zamanda onların bileşenlerinin (genelde özellikleri ya da nitelikleri) de analizini yapmak mümkündür. Örnekleri bir küme olarak birbiri ile ilişkilendirmek için belirli bir özellikte alt kümesine göre kümelemek doğal bir durumdur. Buna göre niteliklerin kümesiyle ilişkili olan alt kümeler, örnek kümesine bağlanır. Bu alt küme "ikili

küme" olarak adlandırılır ve ikili kümeler içerisinde veri setini bölme problemine de "ikili kümeleme problemi" adı verilir (Busygın vd., 2008:2964-2987).

Kümeleme yönteminin gen ifade veri analizindeki yetersizliği sebebiyle ikili kümeleme (biclustering) yöntemi geliştirilmiştir. İkili kümeleme yöntemi, diğer kümeleme yöntemlerinden türetilmiştir. Kümeleme yönteminde bir matrisin sadece satır veya sütunları kümelendirilirken, ikili kümeleme yöntemi ile satır ve sütunlar eş zamanlı kümelendirilmektedir. İlk olarak Cheng ve Church araştırmacıları tarafından tanımlandığı bilinse de bu yöntemi ilk (Hartigan, 1972:123-129)'da kullanmıştır. İkili kümeleme yöntemi ile küçük matris grupları üretilmekte ve gruplarda hangi gen grubunun hangi durum kümesinde veya hangi durum kümesinin hangi gen grubunda aktif olduğu bilgisi elde edilebilmektedir.

İkili kümeleme yöntemi farklı olarak veri matrisinin iki boyutunda kümeleme yapan bir yöntemdir. Diğer bir deyişle kümeleme küresel bir model olarak, ikili kümeleme ise yerel bir model olarak işlem yapmaktadır (Madeira vd., 2004:24-45).

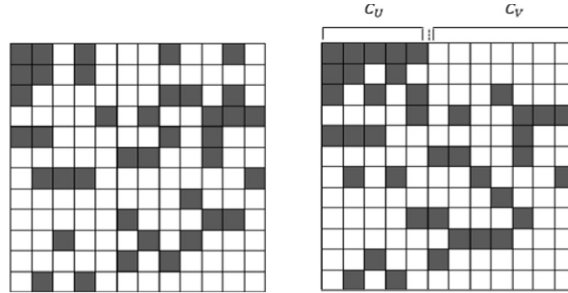
2.1. Bimax Bicluster Yöntemi

İkili veri setleri (binary), bir grup nesnelere ve onların mümkün özellikleri arasındaki ilişkileri hakkında verileri depolamak için yoğun ve basit bir yol sunar. Bimax, çok hızlı bir yanıt sağlayan basit bölmeli ve başarılı bir yaklaşım kullanır (Rodriguez-Baena vd., 2011: 2738-2745). (Prelic vd., 2006: 1122-1129) tarafından temeli oluşturulan Bimax- İkili Kümeleme yöntemi mantıksal bir matrisin alt gruplara bölmesini araştırır. Bimax yöntemi ikili kümeleme algoritmalarının karşılaştırılmasında bir başlangıç noktası olarak kullanılmak üzere geliştirilmiştir. Bimax referans metot olarak geliştirildiğinden amaç, tamamen ikili matristeki 1'li tüm çiftlikleri bulmaktır.

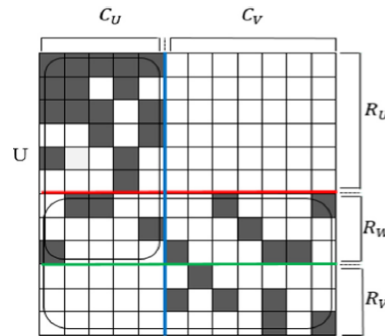
2.2. Bimax Bicluster Algoritması

Adım 1: Orijinal E matrisinin satırlarında karışık bir şekilde 1 ve 0'ların bulunduğu elemanları iki ayrı sütun kümesi olarak (CU ve CV) ikiye bölünür. Şekil 1'de örnek olarak ilk satırda 1 değeri alanlar CU kümesi olarak, 2 değeri alanlar ise CV kümesi olarak ayrılmıştır.

Şekil 1: Bimax Algoritmasının Adım 1 Gösterimi



Şekil 2: Bimax Algoritmasının Adım 2 ve Adım 3 Gösterimi



Kaynak: Prelic vd., 2006:1124.

Adım 2: E matrisi, m satır elemanlı üç alt matrise bölünür. Buradaki RU matrisi sadece CU matrisindeki sütun eleman değerleri 1 olanları içerir. RW matrisi hem CU hem de CV matrisindeki sütun eleman değerleri 1 olanları içerir. RV matrisi ise, sadece CV matrisindeki sütun eleman değerleri 1 olanları içerir. Yeniden düzenlendikten sonra matris Şekil 2'deki gibi görülür.

Adım 3: U ve V matris yapıları belirlenir ve RU ve CV alt matrisleri silinir.

Adım 4: Süreç, eleman değerleri sadece 1'lere ait alt matrisleri bulana kadar ve minimum boyutlu matris elde edilene kadar Adım 1'den Adım 3'e kadar yinelemeli olarak devam eder.

Adım 5: İkili küme olarak sadece 1'lerin bulunduğu en büyük matris kaydedilir ve bu ikili küme matris elemanlara E matrisinden silinerek yeniden ikili küme bulmaya çalışarak devam eder.

3. Verilerin Analizi

Çalışmada istatistiksel analizler için R-project 3.2.5 yazılımı kullanılmıştır. Bu programda ikili kümeleme yöntemlerini kullanabilmek için öncelikle "biclust" paketi yüklenmiştir. Bu paket içerisinde iki değerli matrisler için referans yöntem olan "Bimax Yöntemi" kullanılmıştır.

3.1. Değişkenler

Satır elemanları: NACE koduna sahip Sektörler (Gerçek veri toplam 234)

Sütun elemanları: Kaza türleri (Gerçek veri toplam 44)

Matris elemanları: Türkiye'deki 2011 yılı içerisinde işyerlerindeki hangi kaza türlerinde kaç kişinin öldüğü gösterilmiştir.

3.2. Demografik Özellikler

234 sektörde gerçekleşen ölüm sayısı toplamı 1100'dür. Veri matrisinde satır elemanları (sektörler) içerisinde 1'den fazla ölümlü kazaların meydana geldiği değerler ele alındığında 135 sektör kalır. Yani 99 sektörde 1'den fazla ölümlü kaza meydana gelmediği için analize dâhil edilmemiştir. Benzer şekilde toplamda 44 kaza türünde 1'den fazla ölümlü kazanın meydana geldiği değişken sayısı da 32'dir. 12 kaza türünde 1'den fazla ölüm gerçekleşmediği için analizden çıkarılmıştır. Sonuç olarak analize dâhil edeceğimiz değişkenler satırlarda (sektörlerde) 117, sütunlarda (ölümlü kaza türlerinde) 32'dir. Satır elemanlarını oluşturan 117 sektör Ek 1'de, 32 ölümlü kaza türleri de Ek 2'de frekans değerleri ile birlikte verilmiştir.

117x32'lik frekans değerli matriste toplamda gerçekleşen ölüm sayısı 949'dur. Sektörler içerisinde 208 frekans değeri alan ve dikkat çeken en yüksek değere sahip sektör F4120 NACE koduna sahip "İkamet amaçlı olan veya ikamet amaçlı olmayan binaların inşaatı" dir. Buradaki satır elemanı tüm veri setinin %22'sini oluşturmaktadır. Aynı zamanda bu sektörde 24 farklı ölümlü kaza meydana gelmiştir. Bu sektör içerisinde en fazla ölümün gerçekleştiği kaza türü ise 301 kodlu "yüksek bir yerden derin bir yere düşme"dir. Bu aynı zamanda toplam veride 214 frekans değerine sahip kaza türüdür. Bu kaza türü de tüm verinin %22,5'ini oluşturmaktadır.

3.3. Deneysel Çalışma

Bu bölümde, buraya kadar anlatılan teorik çalışmalara ilişkin sayısal örneklerle yer verilmiştir. Analizde kullanılan 234x44'lük matris, 2011 yılına ait iş yeri ve kaza türlerinde ölen kişileri göstermektedir. Ancak kullanılacak yöntemde matrisin iki değerli (binary) olması gerektiğinden dönüşüm uygulanmıştır. Herhangi iş yerindeki kaza türlerinden ölen kişi sayısı 1'den fazla olan matris elemanları değerleri 1 olarak kodlanmış olup, ölüm olmayanlarda ise 0 olmuştur. Yeniden düzenlenen matriste iş yerlerinde 1'den fazla ölüm olan 117 değişken bulunmaktadır. Yani matristeki satır sayısı 117'dir. Benzer şekilde kaza türlerinde 1'den fazla ölüm gerçekleşen değişkenler ise 32'dir. Sonuç olarak yöntemi uygulayacağımız matris 117x32'lik iki değerli matristir.

Veri matrisinde bulunan toplam 3744 elemandan 456'sı 1 değeri alırken, 3288 eleman 0 değeri almıştır. Toplamda veri setinin %12'sini 1 değerleri oluşturur.

Düzenlenen veri matrisinde ikili küme alt matrislerinin satır ve sütun sayılarının minimum değerlerini toplam değerlerinin %5'inden az olmaması koşulu ile belirlenirse;

- Sektörlerin bulunduğu satır sayısı 117 olduğundan %5'i yaklaşık olarak 6'dır. Yani ikili kümelerde iş yerlerini sınıflandırırken en az 6 elemana ait olması gerekir.

- Kaza türlerinin minimum sayısı bu koşula göre 2 olmasına rağmen anlamlı sonuçlar elde etmek için 3'ten fazla olması gerekmektedir (B.Wang et al. 2016).

Parametre değerleri belirlendikten sonra Bimax algoritması kullanılarak ikili kümeleme işlemindeki program kodu aşağıdaki gibidir:

- sonuc<-biclust (x, method = BCBimax(), minr=6, minc=4, number=100)

4. Bulgular

Bu komut sağladıktan sonra bulunan ikili küme sayıları Tablo 1'de gösterilmiştir:

Tablo 1: İkili Kümelere Ait Satır ve Sütun Elemanları

İkili Kümeler	Matris Büyüklüğü	Satır Elemanları	Sütun Elemanları
BC 1	8x4	62 63 65 72 77 93 94 117	1 3 6 32
BC 2	6x4	62 63 65 67 94 102	1 3 9 32
BC 3	10x4	42 47 62 63 65 72 77 93 104 114	1 6 25 32
BC 4	6x5	47 62 63 65 72 104	1 6 11 25 32
BC 5	6x4	62 72 77 88 93 94	1 3 6 17
BC 6	6x4	62 67 72 77 93 94	1 3 17 32
BC 7	6x4	42 62 72 77 93 94	1 6 17 32
BC 8	6x4	47 62 63 65 94 112	1 6 9 32
BC 9	6x4	42 62 67 72 77 93	1 17 25 32
BC 10	7x4	62 63 65 67 72 77 93	1 3 25 32
BC 11	7x4	19 62 63 65 72 77 93	1 3 6 25
BC 12	6x5	62 63 65 72 77 93	1 3 6 25 32

Yukarıdaki tabloda görüldüğü gibi toplam 12 ikili küme bulunmuştur. Bu kümelerin belirtildiği satır ve sütun elemanları iş yerlerindeki ve kaza türlerindeki elemanların sıra sayısını göstermektedir. Örneğin ilk sıradaki ikili küme BC1'e ait "F4120, F4211, F4221, F4323, F4399, H4941, H4942 ve R9605" NACE kodlarına sahip işyerleridir. Bu iş yeri sınıflaması şunlardır:

- İkamet amaçlı olan veya ikamet amaçlı olmayan binaların inşaatı
- Kara yolları ve otoyolların inşaatı
- Akışkanlar için hizmet projelerinin inşaatı
- Başka yerde sınıflandırılmamış diğer özel inşaat faaliyetleri
- Kara yolu ile yük taşımacılığı
- Ev ve iş yerlerine verilen taşımacılık hizmetleri

Bu işyerlerindeki toplam frekans değeri 96'dır ve bütün matrise oranı ise yaklaşık %21'dir. Yani BC 1 alt matrisi tüm verinin %21'ini oluşturmaktadır.

BC 1'e ait kaza türlerine bakılırsa:

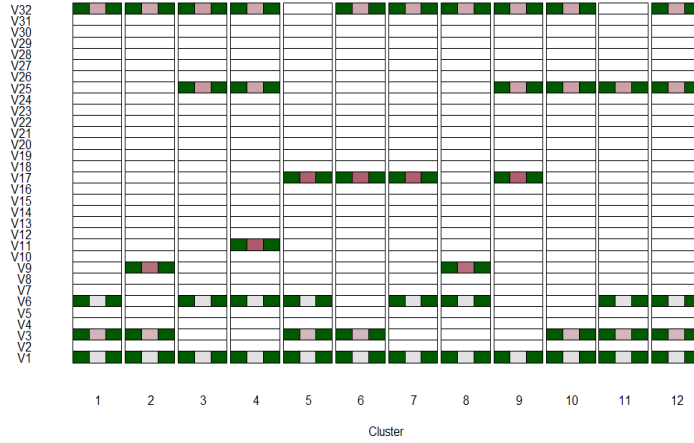
- Motorlu Tasıtın Diğer Bir Vasıtaya Veya Bir İnsana Çarpması
- Motorlu Tasıtın Devrilmesi Sonucu Yaralanmalar
- Yüksek Bir Yerden Derin Bir Yere Düşme
- Diğer Nedenler

Bu kaza türlerinin toplam frekans değeri 199 olup, bütün matrise oranı ise yaklaşık %44'dür. Yani bu kaza türleri tüm verinin %44'ünü oluşturmaktadır.

Veri setindeki %13'ünü oluşturan 1.sıradaki 101 kodlu kaza türü motorlu taşıtların diğer bir vasıtaya veya bir insana çarpması sonucu oluşan ölümleri kapsamaktadır. Bu kaza türü tüm ikili kümelerde bulunmaktadır. 12 ikili kümenin 7'sinde ortak olan bir diğer kaza türü ise 3.sırada bulunan 103 kodlu motorlu taşıtların devrilmesi sonucu yaralanmalardan kaynaklanan ölümlerdir. Bu kaza türü de tüm verinin yaklaşık %10'unu oluşturmaktadır. Kazaların %13'ünü oluşturan ve ikili kümelerin 8'inde ortak olarak bulunan kaza türü ise 6.sırada bulunan 301 kodlu yüksek bir yerden derin bir yere düşmeden kaynaklı ölümlerin olduğudur. 25.sırada bulunan 900 kodlu kaza türü ise kazaların yaklaşık %8'ini oluşturmaktadır. 32.sırada bulunan ve 10 tane ikili kümede ortak olan diğer nedenlerden kaynaklanan kaza türü olarak geçen durum ise kazaların %9'unu oluşturmaktadır.

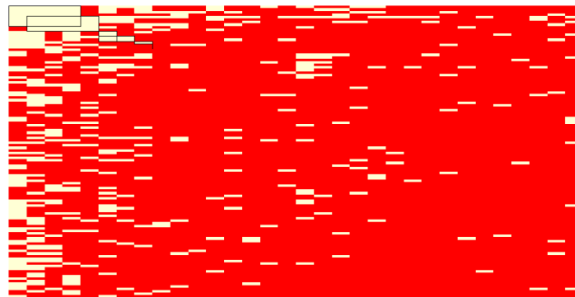
Kaza türlerinin hangi ikili kümeyle ait olduğunu kolaylıkla yorumlamak için programda elde ettiğimiz ikili küme üye grafiğine bakılır.

Şekil 3: Kaza Türleri İçin İkili Küme Üye Grafiği



Yukarıda ikili kümelerin kaza türlerine göre sınıflandırılması gösterilmiştir. Ayrıca her kümeyle ait sütun elemanları ortasında renklerle ölçeklendirmeler yapılmıştır. Bu ölçekler renklendirmenin koyuluğuna göre o sütuna ait tüm veriye göre ortalamasını belirtmiştir. İlk kümeyle göre yorumlayacak olursak, 1. ve 6. kaza türü 3. kaza türüne göre ortalamasının daha çok uzağında kalmıştır. Diğer kümelerdeki kaza türlerini de buna benzer olarak yorumlamamız mümkündür.

Şekil 4: İkili Kümeler İçin Isı Grafiği



Şekil 4'te ise elde ettiğimiz 12 ikili küme için ısı grafiği gösterilmiştir. Burada görüldüğü gibi BC 1 ve BC 2'de örtüşen 4 sektör ve tüm kaza türleri mevcut olup, diğer ikili kümelerde aynı kaza türleri olsa da sektörlerin birbiri ile örtüşmediği ikili kümeler de vardır.

Bulunan ilk üç ikili kümeler içerisinde yer alan sektörlerin tehlike sınıflarını Tablo 2’de görülebilir.

Tablo 2: İlk 3 İkili Kümelere Ait Sektörlerin Tehlike Sınıfları

NACE Kodları	İkili Kümeler	Tehlike Sınıfı
BC1		
F4120	İkamet amaçlı olan veya ikamet amaçlı olmayan binaların inşaatı	Çok Tehlikeli
F4211	Kara yolları ve otoyolların inşaatı	Çok Tehlikeli
F4221	Akışkanlar için hizmet projelerinin inşaatı	Çok Tehlikeli
F4399	İnşaatlardaki diğer bütünleyici ve tamamlayıcı işler	Çok Tehlikeli
H4941	Kara yolu ile yük taşımacılığı	Çok Tehlikeli
H4942	Ev ve iş yerlerine verilen taşımacılık hizmetleri	Çok Tehlikeli
BC2		
F4120	İkamet amaçlı olan veya ikamet amaçlı olmayan binaların inşaatı	Çok Tehlikeli
F4211	Kara yolları ve otoyolların inşaatı	Çok Tehlikeli
F4221	Akışkanlar için hizmet projelerinin inşaatı	Çok Tehlikeli
F4291	Su projeleri inşaatı	Çok Tehlikeli
H4942	Ev ve iş yerlerine verilen taşımacılık hizmetleri	Tehlikeli
I5610	Lokantalar ve seyyar yemek hizmeti faaliyetleri	Az Tehlikeli
BC3		
C2512	Metalden kapı ve pencere imalatı	Tehlikeli
C2599	Başka yerde sınıflandırılmamış diğer fabrikasyon metal ürünlerin imalatı	Tehlikeli
F4120	İkamet amaçlı olan veya ikamet amaçlı olmayan binaların inşaatı	Çok Tehlikeli
F4211	Kara yolları ve otoyolların inşaatı	Çok Tehlikeli
F4221	Akışkanlar için hizmet projelerinin inşaatı	Çok Tehlikeli
F4399	İnşaatlardaki diğer bütünleyici ve tamamlayıcı işler	Çok Tehlikeli
H4941	Kara yolu ile yük taşımacılığı	Çok Tehlikeli
M7022	İşletme ve diğer idari danışmanlık faaliyetleri	Az Tehlikeli
Q9200	Kumar ve müşterek bahis faaliyetleri	Az Tehlikeli

İkili kümelere bulunan sektörlerin tehlike sınıfları; BC1 için tüm sektörler çok tehlikeli, BC2 ve BC3 için genellikle çok tehlikeli olarak Tablo 2’de görülmektedir.

5. Sonuç

Türkiye’de gerçekleşen ölümlü kazalar, belirli sektörler tarafından tehlikeli sınıflandırılırken dikkate alınması gereken en önemli etkilerden biri belirli kaza türlerinin etkisinde olmasıdır. Bu yüzden sektörleri tehlikelerine göre sınıflarken, aynı zamanda kaza türlerinin de birbiri ile ilişkisi olduğuna bakılabilir. Geleneksel kümeleme yöntemleri ile yapılan çalışmalara kıyasla, ikili kümeleme yöntemi daha tutarlı ve etkin sonuçlar vermektedir. Çünkü sektörlerin tehlikeli olmasına etkili olacak ölümlü kaza türleri bu kümeleyi yaparken dikkate alınması gerekir.

Bu çalışmada ikili kümeleme yöntemlerinde uygulanan Bimax algoritması seçilmesinin nedeni geleneksel kümeleme yöntemlerine göre daha ayrıntılı ve gerçekçi sonuçlar sunmasıdır. Bu yüzden çalışmanın analiz kısmında sektörler ve kaza türleri birlikte kümelenecek tehlikeli sektör grupları oluşturulmuştur.

Tehlikeli sektör gruplarından genel olarak inşaat sektörü ve taşımacılık sektörlerinin alt bölümleri birlikte kümelenecek ilk ikili kümenin satır elemanlarını oluşturmuştur. Sektörleri birlikte kümelirken dikkate alınan kaza türleri ise motorlu taşıtın diğer bir vasıtaya veya bir insana çarpması, motorlu taşıtın devrilmesi ve yüksek bir yerden düşmesidir. İnşaat sektöründe veya taşımacılık sektörü birlikte kümelirken kaza türlerinden motorlu taşıt kazası ve yüksek bir yerden düşerek meydana gelen kaza türleri önemli bir şekilde etkilemektedir.

Sektörler içerisinde en fazla kaza türlerinden olan motorlu taşıtların diğer bir vasıtaya veya bir insana çarpması sonucu oluşan ölümler elde edilen tüm ikili kümelerin elemanı olarak bulunmuştur. Tehlike sınıfına girecek sektörler için en önemli kaza türü olarak belirlenebilir. Bu

kaza türleri dikkate alındığında tehlike sınıfında olabilecek sektörler bulunabilir. Ayrıca belirli bir sektörün tehlike sınıfına ait olacağını belirlerken tek bir kaza türü değil, birlikte kümelenen kaza türleri dikkate alınması daha anlamlı ve etkili sektör grupları oluşturacaktır.

Bu gerçek veriyle uygulanan analiz yöntemi sayesinde sektörler için risk sınıflandırılması yapılarak işyerlerindeki ölümlü kazalara karşı önlem alınabilir. Ayrıca bu önlemlere karşı sektörler için farklı prim uygulamaları geliştirilebilir ve kümelenen sektörler için daha tutarlı ve doğru bir şekilde yaptırımlar olabilir.

Kaynakça

- Busygin, S., Prokopyev, O. ve Pardalos, P. M. (2008). Biclustering in Data Mining. *Computers & Operations Research*, 35(9), 2964-2987.
- Ceylan, H. (2014). Türkiye’de İnşaat Sektöründe Meydana Gelen İş Kazalarının Analizi. *International Journal of Engineering Research and Development*, 6(1), 1-6.
- Cheng, Y. ve Church, G. M. (2000). Biclustering of Expression Data. *International Conference on Intelligent Systems for Molecular Biology*, 8, 1-7.
- Eisen, M. B., Spellman, P. T., Brown, P. O. ve Botstein, D. (1998). Cluster Analysis and Display of Genome-wide Expression Patterns. *Proceedings of the National Academy of Sciences*, 95(25), 14863-14868.
- Hartigan, J. A. (1972). Direct Clustering of A Data Matrix. *Journal of The American Statistical Association*, 67(337), 123-129.
- Kluger, Y., Basri, R., Chang, J. T. ve Gerstein, M. (2003). Spectral Biclustering of Microarray Data: Co-clustering Genes and Conditions. *Genome Research*, 13(4), 703-716.
- Lazzeroni, L. ve Owen, A. (2002). Plaid Models for Gene Expression Data. *Statistica Sinica*, 61-86.
- Madeira, S. C. ve Oliveira, A. L. (2004). Biclustering Algorithms for Biological Data Analysis: A Survey. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 1(1), 24-45.
- Murali, T. M. ve Kasif, S. (2003). Extracting Conserved Gene Expression Motifs from Gene Expression Data. *In Pacific Symposium on Biocomputing*, 8, 77-88.
- Prelić, A., Bleuler, S., Zimmermann, P., Wille, A., Bühlmann, P., Gruissem, W. ve Zitzler, E. (2006). A Systematic Comparison and Evaluation of Biclustering Methods for Gene Expression Data. *Bioinformatics*, 22(9), 1122-1129.
- Rodriguez-Baena, D. S., Perez-Pulido, A. J. ve Aguilar– Ruiz, J. S. (2011). A Biclustering Algorithm for Extracting Bit-Patterns from Binary Datasets. *Bioinformatics*, 27(19), 2738-2745.
- Turner, H., Bailey, T. ve Krzanowski, W. (2005). Improved Biclustering of Microarray Data Demonstrated through Systematic Performance Tests. *Computational Statistics & Data analysis*, 48(2), 235-254.

EKLER

Ek 1. Matristeki Sütun Değişkeni Olan Kaza Türleri İçin Frekans Değerleri ve Açıklamaları

Kaza Türü	Frekans	Açıklama	
1	101	59	MOTORLU TASITIN DİĞER BİR VASITAYA VEYA BİR İNSANA ÇARPMASI
2	102	7	MOTORLU TASITA BINİP İNERKEN MEYDANA GELEN KAZALAR
3	103	42	MOTORLU TASITIN DEVRİLMESİ SONUCU YARALANMALAR
4	104	13	DURAN BİR MOTORLU TASIT SEBEBİYLE VEYA ÜZERİNDE OLAN KAZALAR SINIFLANDIRILAMAYAN MOTORLU VEYA MOTORSUZ BİR TASITIN SEBEP
5	108	4	OLDUGU KAZALAR
6	301	61	YÜKSEK BİR YERDEN DERİN BİR YERE DÜŞME
7	302	8	KAZA NETİCESİ SUDA BOĞULMA VE SUYA DÜŞME
8	303	11	KİSİLERİN HEMZEMİN ORTAMDA DÜŞMESİ
9	400	23	MAKİNALARIN SEBEP OLDUGU KAZALAR
10	501	5	ATESLİ SİLAHLAR SONUCU ORTAYA ÇIKAN KAZALAR
11	502	18	YANICI MADDELERİN ATES ALMASI VE PATLAMASINDAN İLE Rİ GELEN KAZALAR
12	503	6	BASINÇ ALTINDAKİ BİR CİSMİN PATLAMASINDAN ÇIKAN KAZALAR
13	603	7	SICAK BİR MADDEDEN, SIVIDAN, GAZDAN, ALEVDEN MEYDANA GELEN KAZALAR
14	701	19	KAYAN VE ÇÖKEN (TOPRAK, KAYA, TAŞ, KAR) CİSİMLERİN ÇARPMASI VE DEVİRMESİ
15	702	9	ÇÖKMELER
16	703	9	TASIMA SİRASINDA TASINAN CİSİMLERİN DÜŞMESİ SONUCU OLUSAN KAZALAR
17	704	19	BASKA YERDE SINIFLANDIRILMAMIS DÜŞEN CİSİMLERİN ÇARPMASI , DEVRİLMESİ
18	801	8	VÜCUDUN VEYA BİR ORGANIN İKİ CİSİM ARASINDA KALIRAK SIKIŞMASI, EZİLMESİ
19	802	8	BİR CİSMİN ÇARPMASI SONUCUNDA ÇÖKEN, DEVRİLEN BİR CİSİM ALTINDA YARALANMA
20	803	3	DURAN CİSİMLERE ÇARPMASI (DAHA ÖNCEKİ DÜŞMELER SEBEBİYLE ÇARPISMALAR HARIÇ)
21	805	3	DÜŞEN CİSİMLERİN DİSİNDA HAREKET EDEN CİSİMLERİN ÇARPMASI
22	807	3	SABİT BİR MEKAN İLE HAREKET EDEN CİSİM ARASINDA SIKIŞMA
23	808	5	HAREKET EDEN CİSİMLERİN ARASINDA SIKIŞMAK
24	809	5	KESİCİ VE BATICI ALETİN SEBEP OLDUGU KAZA
25	900	36	ELEKTRİK AKIMINDAN İLERİ GELEN KAZALAR
26	1004	2	AGIR YORUCU HAREKETLER
27	1102	3	NEFES BORUSUNUN BİR GIDA MADDESİYLE TIKANMASI
28	1500	2	KAYNAK YAPARKEN MEYDANA GELEN KAZALAR
29	1601	5	KENDİ KENDİNE ÖLDÜRME VE YARALAMA
30	1602	6	BİR BASKASI TARAFINDAN ÖLDÜRÜLME VE YARALANMA
31	1801	10	ZARARLI MADDELERİN SOLUNUMU VEYA SİNDİRİM YOLUYLA TEMAS, İLİTİ VEYA MUKOZADAN EMİLME
32	1900	37	DİĞER NEDENLER

Ek 2. Matristeki Satır Değişkeni, NACE Kodlu İş Yeri Türleri İçin Frekans Değerleri ve Açıklamaları

NACE Kodu	Frekans	Açıklama	
1	A0130	2	Dikim için bitki yetiştirilmesi
2	A0220	2	Endüstriyel ve yakacak odun üretimi
3	B0510	7	Taş kömürü madenciliği
4	B0520	7	Linyit madenciliği
5	B0610	2	Ham petrol çıkarımı
6	B0721	4	Uranyum ve toryum cevherleri madenciliği

7	B0729	4	Diğer demir dışı metal cevherleri madenciliği
8	B0811	13	Süsleme ve yapı taşları ile kireç taşı, alçı taşı, tebeşir ve kayağantaşı (arduvaz-kayraktaşı) ocakçılığı
9	B0812	4	Çakıl ve kum ocaklarının faaliyetleri; kil ve kaolin çıkarımı
10	B0891	2	Kimyasal ve gübreleme amaçlı mineral madenciliği
11	B0899	5	Başka yerde sınıflandırılmamış diğer madencilik ve taş ocakçılığı
12	B0990	2	Madencilik ve taş ocakçılığını destekleyici diğer faaliyetler
13	C1039	2	Başka yerde sınıflandırılmamış meyve ve sebzelerin işlenmesi ve saklanması
14	C1041	3	Sıvı ve katı yağ imalatı
15	C1071	3	Ekmek, taze pastane ürünleri ve taze kek imalatı
16	C1081	2	Şeker imalatı
17	C1083	2	Kahve ve çayın işlenmesi
18	C1084	2	Baharat, sos, sirke ve diğer çeşni maddelerinin imalatı
19	C1320	4	Dokuma
20	C1330	6	Tekstil ürünlerinin bitirilmesi
21	C1391	2	Örgü (triko) veya tığ işi (kroşe) kumaşların imalatı
22	C1411	3	Deri giyim eşyası imalatı
23	C1413	4	Diğer dış giyim eşyaları imalatı
24	C1416 (C1419)	2	Diğer giyim eşyalarının ve giysi aksesuarlarının imalatı
25	C1610	2	Ağaçların biçilmesi ve planyalanması
26	C1621	2	Ahşap kaplama paneli ve ağaç esaslı panel imalatı
27	C1623	2	Diğer bina doğramacılığı ve marangozluk ürünlerinin imalatı
28	C1721	2	Oluklu kağıt ve mukavva imalatı ile kağıt ve mukavvadadan yapılan muhafazaların imalatı
29	C1920	2	Rafine edilmiş petrol ürünleri imalatı
30	C2016	2	Birincil formda plastik hammaddelerin imalatı
31	C2221	4	Plastik tabaka, levha, tüp ve profil imalatı
32	C2223	2	Plastik inşaat malzemesi imalatı
33	C2320	2	Ateşe dayanıklı (refrakter) ürünlerin imalatı
34	C2331	2	Seramik karo ve kaldırım taşları imalatı
35	C2332	2	Fırınlanmış kilden tuğla, karo ve inşaat malzemeleri imalatı
36	C2352	2	Kireç ve alçı imalatı
37	C2361	6	İnşaat amaçlı beton ürünlerin imalatı
38	C2370	5	Taş ve mermerin kesilmesi, şekil verilmesi ve bitirilmesi
39	C2410	3	Ana demir ve çelik ürünleri ile ferro alaşımların imalatı
40	C2420	3	Çelikten tüpler, borular, içi boş profiller ve benzeri bağlantı parçalarının imalatı
41	C2451	2	Demir döküm
42	C2512	5	Metalden kapı ve pencere imalatı
43	C2521	2	Merkezi ısıtma radyatörleri (elektrikli radyatörler hariç) ve sıcak su kazanları (boylerleri) imalatı
44	C2562	5	Metallerin makinede işlenmesi ve şekil verilmesi
45	C2571	2	Çatal-bıçak takımları ve diğer kesici aletlerin imalatı
46	C2592	3	Metalden hafif paketleme malzemeleri imalatı
47	C2599	6	Başka yerde sınıflandırılmamış diğer fabrikasyon metal ürünlerin imalatı
48	C2826(C2823)	5	Büro makineleri ve ekipmanları imalatı (bilgisayarlar ve çevre birimleri hariç)
49	C2829	2	Başka yerde sınıflandırılmamış diğer genel amaçlı makinelerin imalatı
50	C2830	2	Tarım ve ormancılık makinelerinin imalatı
51	C2932	2	Motorlu kara taşıtları için diğer parça ve aksesuarların imalatı
52	C3012	3	Eğlence ve spor amaçlı teknelerin yapımı
53	C3020	2	Demir yolu lokomotifleri ve vagonlarının imalatı
54	C3101	3	Büro ve mağaza mobilyaları imalatı
55	C3109	3	Diğer mobilyaların imalatı
56	C3312	3	Makinelerin onarımı
57	C3315	2	Gemilerin ve teknelerin bakım ve onarımı
58	C3317	3	Diğer ulaşım ekipmanlarının bakım ve onarımı

59	D3511	3	Elektrik enerjisi üretimi
60	D3530	3	Buhar ve iklimlendirme temini
61	E3821	4	Tehlikesiz atıkların ıslahı ve bertaraf edilmesi
62	F4120	24	İkamet amaçlı olan veya ikamet amaçlı olmayan binaların inşaatı
63	F4211	9	Kara yolları ve otoyolların inşaatı
64	F4212	3	Demir yolları ve metroların inşaatı
65	F4221	12	Akışkanlar için hizmet projelerinin inşaatı
66	F4222	2	Elektrik ve telekomünikasyon için hizmet projelerinin inşaatı
67	F4291	9	Su projeleri inşaatı
68	F4299	4	Başka yerde sınıflandırılmamış bina dışı diğer yapılara ait projelerin inşaatı
69	F4312	3	Şantiyenin hazırlanması
70	F4321	4	Elektrik tesisatı
71	F4322	4	Sıhhi tesisat, ısıtma ve iklimlendirme tesisatı
72	F4323	12	
73	F4329	2	Diğer inşaat tesisatı
74	F4331	4	Sıva işleri
75	F4334	2	Boya ve cam işleri
76	F4339	4	İnşaatlardaki diğer bütünleyici ve tamamlayıcı işler
77	F4399	9	Başka yerde sınıflandırılmamış diğer özel inşaat faaliyetleri
78	G4520	5	Motorlu kara taşıtlarının bakım ve onarımı
79	G4532	2	Motorlu kara taşıtlarının parça ve aksesuarlarının perakende ticareti
80	G4612	3	Yakıtların, maden cevherlerinin, metallerin ve endüstriyel kimyasalların satışı ile ilgili araçlar
81	G4613	2	Kereste ve inşaat malzemelerinin satışı ile ilgili araçlar
82	G4618	3	Belirli diğer ürünlerin satışı ile ilgili uzmanlaşmış araçlar
83	G4673	2	Ağaç, inşaat malzemesi ve sıhhi teçhizat toptan ticareti
84	G4674	2	Hırdavat, sıhhi tesisat ve ısıtma tesisatı malzemelerinin toptan ticareti
85	G4711	4	Belirli bir mala tahsis edilmemiş mağazalarda gıda, içecek veya tütün ağırlıklı perakende ticaret
86	G4730	5	Belirli bir mala tahsis edilmiş mağazalarda otomotiv yakıtının perakende ticareti
87	G4751	2	Belirli bir mala tahsis edilmiş mağazalarda tekstil ürünleri perakende ticareti
88	G4752	4	Belirli bir mala tahsis edilmiş mağazalarda hırdavat, boya ve cam perakende ticareti
89	G4759	3	Belirli bir mala tahsis edilmiş mağazalarda mobilya, aydınlatma teçhizatı ve diğer ev eşyalarının perakende ticareti
90	H4931	2	Kara taşımacılığı ile yapılan şehir içi ve banliyö yolcu taşımacılığı
91	H4932	2	Taksi taşımacılığı
92	H4936	4	
93	H4941	13	Kara yolu ile yük taşımacılığı
94	H4942	10	Ev ve iş yerlerine verilen taşımacılık hizmetleri
95	H5019	2	
96	H5210	3	Depolama ve ambarlama
97	H5221	4	Kara taşımacılığını destekleyici hizmet faaliyetleri
98	H5222	4	Su yolu taşımacılığını destekleyici hizmet faaliyetleri
99	H5224	4	Kargo yükleme boşaltma hizmetleri
100	H5229	2	Taşımacılığı destekleyici diğer faaliyetler
101	I5510	2	Oteller ve benzeri konaklama yerleri
102	I5610	6	Lokantalar ve seyyar yemek hizmeti faaliyetleri
103	I5629	2	Diğer yiyecek hizmeti faaliyetleri
104	M7022	8	İşletme ve diğer idari danışmanlık faaliyetleri
105	M7111	2	Mimarlık faaliyetleri
106	M7311	2	Reklam ajanslarının faaliyetleri
107	N7711	2	Motorlu hafif kara taşıtlarının ve arabaların kiralanması ve leasingi
108	N8010	4	Özel güvenlik faaliyetleri
109	N8110	2	Tesis bünyesindeki kombine destek hizmetleri
110	N8121	5	Binaların genel temizliği

111	N8130	3	Çevre düzenlemesi ve bakımı faaliyetleri
112	N8211	7	Kombine büro yönetim hizmeti faaliyetleri
113	O8510	2	Okul öncesi eğitim
114	Q9200	6	Kumar ve müşterek bahis faaliyetleri
115	R9512	2	İletişim araç ve gereçlerinin onarımı
116	R9521	3	Tüketici elektroniği ürünlerinin onarımı
117	R9605	7	

CLUSTERING BY SECTORS ACCORDING TO THE TYPE OF FATAL ACCIDENTS IN TURKEY: THE BICLUSTERING METHOD

Extended Abstract

Aim: The aim of the study is to cluster the sectors in Turkey and the types of fatal accidents by biclustering method. Conventional clustering methods can not simultaneously cluster sectors and types of accidents. Therefore biclustering method developed from clustering method is discussed. In this way, it is aimed to cluster the sectors taking into consider the types of accidents related to each other. At the same time, it can be determined which sector clusters are associated with accident clusters. And it can also be determined that the types of accidents that are related to each other have a common impact on the types of accidents.

Method(s): Bimax bicluster method, which is one of the most frequently used methods of biclustering methods, is used in the study. And it is known as the reference model in biclustering algorithms. Bimax allows binary data matrix to be clustered in terms of rows and columns at the same time. Bimax analyzes the sub-matrices of the data matrix and is shown as the most efficient algorithm in the existing binary clustering algorithms. The data set covered in the study is a binary matrix consisting of 1 for mortality accident and 0 for other cases. Matrix lines indicate sectors, while columns show mortal accident types. 117x32 matrix of data was used in the study. A total of 3744 elements in the data matrix had a value of "1" representing 1100 of them were fatal accidents. In order to find biclusters in practice, R package program was used. Also the "biclust" package has been added in the R program. The results are given in biclustering graphs.

Findings: Among the fatal accidents, the first three respectively "a deep fall from a high place", "motor vehicle crashing into another vehicle or a human being" and "electric current accidents". The first three sectors where mortality accidents are most frequently seen are "construction of residential or non-residential buildings", "ornaments and building stones, limestone, gypsum, chalk and rock stone quarrying" and "road and freight transport". As a result of the Bimax algorithm, sectors are divided into 12 clusters. The sizes of the obtained biclusters are between 6 and 10 members for the sectors. For accident types, it has 4 and 5 members. The largest bicluster is a binary cluster of 10 sector types and 4 accident types. While the smallest cluster is composed of 6 sectors, the largest cluster is composed of 10 sectors. Some of the findings of these clusters are as follows: in the clusters including the construction sectors, accidents in the form of falling from high ground are seen intensely. Accidents arising from motor vehicles are concentrated in the heavily of the transportation sectors (this variable is found in all biclusters). The sectors that include food services are attracted to the accidents that cause the flammable materials to shoot and the explosion to take place.

Conclusion: In this study, sectors and types of accidents were clustered together to form risky sector groups. Sectors and types of accidents are used simultaneously from biclustering algorithms for clustering. The implemented Bimax algorithm provides more detailed and realistic results than the conventional clustering methods. With this practical implementation, measures can be taken to take precautions against mortal accidents by making risk classifications for the sectors. In addition, different premium applications for sectors may be developed against these measures and sanctions may be more consistent and correct against the clustered sectors. It can also be seen in which sector groups the types of accidents within a certain group are effective.