

İş Sağlığı ve Güvenliği Sektöründe Bayes Ağları Uygulaması

Application of Bayesian Networks in Occupational Health and Safety Sector

Zinnet Duygu AKŞEHİR, Ebru PEKEL,
Sedat AKLEYLEK, Erdal KILIÇ Ondokuz
Mayıs Üniversitesi Mühendislik Fakültesi
Bilgisayar Mühendisliği Bölümü
{duygu.aksehir, ebru.pekel, erdal.kilic,
sedat.akleylek@bil.omu.edu.tr}

Yalçın ORUÇ
Rönesans Holding
{yalcin.oruc@ronesans.com}

Öz

Günümüzde inşaat sektöründeki gelişmeyle beraber iş kazalarının da sayısı artmıştır. Teknolojinin gelişimi, iş güvenliğindeki önlemlerde eksiklikler ve çalışanların eğitimsiz oluşu bu iş kazalarındaki ana nedenlerdir. Bu çalışmada iş kazası verileri önce veri ön işleme adımından geçirilip daha sonra elde edilen veriler üzerinde tek değişkenli sıklık ve çapraz tablolama analizi yapılmıştır. Yapılan analizler sonucunda iş kazalarının meydana gelmesinde yüksek risk taşıyan değişkenler tespit edilmiştir. Ardından bu değişkenlerin iş kazasına etkileri Bayes ağları (BNs) ile analiz edilmiştir. Bayes ağı, değişkenler arasındaki koşullu bağımlılık ilişkilerini yansıtan grafiksel bir modeldir. Bayes ağı, uluslararası bir inşaat firmasından bir veri kümesi üzerinde uygulanmıştır. Kurulan Bayes ağının doğruluk oranı ve diğer performans ölçütleri analiz edilmiş ve yapılan modelin etkinliği yorumlanmıştır. Deneysel sonuçlara göre, bazı iş kazası vakalarının makine öğrenme tekniklerini kullanarak yüksek doğruluk oranları ile önceden tahmin edilebileceği gösterilmiştir.

Anahtar Sözcükler: Bayes ağı, sınıflandırma, makine öğrenmesi, iş kazası, tek değişkenli frekans analizi, çapraz tablolama

Abstract

Nowadays with the development in construction sector, there has been an increase in production and the number of accidents has increased. Progressive technology, the development of the industry and the lack of protective precaution, and the responsibility for the uneducated employees are the main causes of occupational accidents. In this study, occupational accident data were passed through the data preprocessing step and then the univariate frequency and cross tabulation analysis were performed. As a result of the analysis, high risk variables were identified in the occurrence of occupational accidents. The effects of the se variables on the occupational accident were the analyzed by Bayesian Networks (BNs). The BNs are graphical model which reflects the conditional dependency relations between variables. BNs are constructed on a data set from an international construction company. The accuracy rate and other performance measures of the constructed Bayesian network are analyzed and the effectiveness of the constructed model is analyzed. According to the experimental results, it's explicit that some cases of occupational accidents can be predicted before hand with high accuracies by using machine learning techniques.

Gönderme ve kabul tarihi: 06.02.2019-13.03.2019

S. Akleylek: orcid.org/0000-0001-7005-6489

E. Kılıç: orcid.org/0000-0003-1585-0991 Y.

Oruç: orcid.org/0000-0002-6305-5653

Z. D. Akşehir: orcid.org/0000-0002-6834-6847

E. Pekel: orcid.org/0000-0001-7717-6790

Keywords: Bayesian Network, classification, machine learning, occupational accident, univariate analysis, cross tabulation

1. Giriş

Dünyada her yıl, başta inşaat sektörü olmak üzere çeşitli sektörlerde iş kazaları nedeniyle milyonlarca insan yaralanmakta ya da yaşamını yitirmektedir. Bu iş kazalarının tekrarlarının yaşanmaması için geçmişte meydana gelen iş kazalarının analizi yapılarak öncelikle kaza risk modelleri oluşturulmalı ve ardından kaza olasılıklarının tespiti sağlanmalıdır. Kaza analizi üzerinde var olan çalışmalar incelendiğinde, araştırmacılar genel olarak iki farklı analiz üzerinde yoğunlaşmaya çalışmışlardır. Bunlardan birincisi temel istatistiksel analiz olurken ikincisi veri madenciliği ve makine öğrenmesini içeren analiz olmuştur.

Baradan ve arkadaşları bir çalışmada Ege bölgesinde 2009-2010 yılları arasında inşaat sektöründe meydana gelen iş kazalarını ele alarak tek değişkenli sıklık analizi ve çapraz tablolama yöntemleri(SAÇT) ile kaza risk analizi yapmıştır[1]. Baradan ve Akboğa yapmış olduğu başka bir çalışmada ise yine tek değişkenli sıklık analizi ve çapraz tablolama yöntemini kullanarak farklı veri setleri üzerinde iş kazası analizleri yapmışlardır[2]. [1] ve [2]'de kullanılan SAÇT yöntemi bu çalışmadaki veri kümesinin yapısı ve çalışılan problemin benzerliğinden dolayı önışleme aşamasında bu yöntem kullanılmıştır.

Erginel ve Toptancı çalışmalarında 2005-2015 yılları arasında İç Anadolu Bölgesi inşaat iş kolunda meydana gelen iş kazası verilerini olasılık dağılımlarıyla modelleyerek kaza risk analizi yapmıştır [3].

Gümüş çalışmasında 2015 Sosyal Güvenlik Kurumu yıllık raporlarını incelemiştir. Ardından iş kazalarına bağlı verileri Excel tablolarından süzdürerek ölümlü iş kazalarının görülme sıklıklarının analizini yapmıştır [4].

Eratak ise yaptığı doktora tezi çalışmasında iki adet linyit kömür madeni iş kazası verilerini inceleyerek ilerde meydana gelebilecek kazaların önlenmesi veya en aza indirgenmesini amaçlamıştır. Bunu yaparken hem temel istatistiksel tekniklerden hem de makine öğrenimi tekniklerinden faydalanmıştır. Basit istatistik yöntemleriyle değişkenler analiz edilip en tehlikeli durumlar hakkında genel bilgi edinilmiştir. Makine öğrenimi teknikleriyle (regresyon, sinir ağları

ve bulanık mantık) de gelecekteki kaza tahmini yapılmaya çalışılmıştır [5].

Geçer yapmış olduğu yüksek lisans tez çalışmasında Sakarya İl Emniyet Müdürlüğü tarafından tutulan trafik kaza raporlarından yararlanarak bir ilişkisel veri tabanı yönetim sistemi oluşturmuştur. Bu sayede çeşitli raporlama ve sorgulama teknikleri ile kazanın hangi saatte yapıldığı, kaza oluşum şekli, kazayı yapan kişinin eğitim seviyesi bilgileri elde edilmiştir. İlerleyen aşamalarda bu sistem üzerine veri madenciliği yöntemlerinin eklenmesini hedeflemiştir [6].

Literatürde inşaat sektöründe yeterli sayıda iş kazası türleri ve etkileri üzerine Bayes ağları ile tahminleme çalışmaları mevcut değildir. Bu nedenle bu çalışmada iş kazası etkileri ve şiddetleri üzerine Bayes ağları kullanılarak bir tahminleme çalışılması yapılmak istenmiştir. Bu amaç doğrultusunda bir makine öğrenmesi algoritması olan Bayes Ağları(BA) uygulanarak veri setindeki bazı değişkenler tahmin edilmek istenmiştir.

Bayes ağları olasılıksal bir çizge yöntemi olmasından dolayı gerçek hayat problemlerine daha çok yakınsamaktadır. Bu nedenle literatürde Bayes ağları ile ilgili çalışmalar bulunmaktadır[7]. Çalışmanın bu bölümünde öncelikle iş kazalarının analizi üzerine yapılan çalışmalar incelenmiştir. Ardından iş sağlığı ve güvenliği alanında Bayes ağları kullanılarak yapılan tahminleme çalışmaları özetlenecektir.

Garcia-Herrera ve arkadaşları sosyal desteğin iş taleplerinin ve mesleki stresin neden olduğunu, iş kazalarını nasıl azalttığını analiz eden bir Bayes ağı modeli oluşturmuştur[8]. Martin ve arkadaşları işyeri kazalarını daha gerçekçi bir şekilde tanımlamak için destek ekipmanlarını kullanarak gerçekleştirilen işyeri görevlerini çevreleyen koşulları analiz etmiştir [9]. Leu ve arkadaşlarının yapmış olduğu bir çalışmada ise çelik bina yapımındaki olası iş kazaları Bayes ağları kullanarak analiz edilmek istenmiştir[10].

İş kazaları başlığı altında incelemiş olduğumuz çalışmalar, uygulanan alan ve kullanılan analiz yöntemine göre Çizelge-1'deki gibi özetlenmiştir. Çizelge-1'deki bilgiler doğrultusunda, bu çalışmada kullanılacak veri kümesi için en uygun yöntemin ne olacağı araştırılmıştır. Yapılan araştırmalar sonucunda en uygun yöntemin, temel istatistiksel analiz ve Bayes ağlarından oluşan melez bir yaklaşım olduğu kanaatine varılmıştır.

Bu çalışmada iş kazası verileri bir önışleme adımından geçirilerek iş kazası risk analizi için uygun hale getirilmiştir. Önışleme aşaması sonrasında elde edilen bu veriler üzerine temel istatistiksel analizlerden olan tek deęişkenli sıklık ve çapraz tablolama analizleri uygulanmıştır. Çapraz tablolama analizinin ardından bağımlı deęişken ve bağımsız deęişkenler arasındaki ilişkinin en güçlü olduęu deęişkenler belirlenerek bu deęişkenler üzerinde Bayes aęları uygulanmış ve çıktı olarak seçilen deęişkenler için sınıflandırma analizi yapılmıştır.

Çizelge-1. İş kazaları kapsamında yapılan çalışmaların özetlenmesi

<i>Uygulanan Alan</i>	<i>Kullanılan Analiz Yöntemi</i>
İnşaat sektöründeki iş kazaları [1,2]	Temel istatistiksel analiz
İnşaat sektöründeki iş kazaları [3]	Olasılık dağılım analizi
SGK raporlarında kayıtlı iş kazaları [4]	Excel tabloları ile analiz
Maden iş kazaları [5]	Temel istatistiksel analiz ve makine öğrenimi teknikleri
Trafik kazaları [6]	Sorgulama ve raporlama analizi
Mesleki stresin neden olduęu iş kazaları [8]	Bayes aęları
İşyeri kazaları [9]	Bayes aęları
İnşaat sektöründeki iş kazaları [10]	Bayes aęları

Bu çalışma öncesinde yürütölen başka bir çalışmada ise veriler ham hali ile sınıflandırma analizine tabi tutulmuştur [11]. Bu çalışmada amaç verilerdeki tekrarlı örüntülerden kurtularak sınıflandırma analizine daha makul seviyede bir veri sunmaktır.

Bu çalışma şu şekilde organize edilmiştir: Bölüm 2’de verinin temizlenmesi ve temizlenen verinin tablo uygulamasında nasıl tutulduęu hakkında bilgi verilecektir. Bölüm 3’de uygulamada kullanılan tek deęişkenli sıklık analizi, çapraz tablolama analizi ve Bayes aęları yöntemlerine değinilecektir. Son bölümde ise bu analizler sonucunda elde edilen bulgular tablo ve grafikler halinde verilerek yorumlanacaktır.

2. Veri Kümesinin Hazırlanması

Bu çalışmada gerçek iş kazası verileri kullanılmıştır. Sınıflandırma aşamasında kullanılacak olan veriler bir

önışleme sürecine tabi tutulmuş ve sonrasında bu veriler Bayes aęı ile analiz edilmiştir. Kullanılan veri kümesi toplam 71 tane öznitelik olmak üzere 424767 iş kazası raporlarından oluşmaktadır. Fakat veri kümesi incelendiğinde raporların tutulması sırasında gerekli bilgilerin eksik ya da hatalı girildięi gözlemlenmiştir. Bu yüzden ilk önce bu eksik ya da hatalı girişlerin bulunduęu veriler R programı ile veri kümesinden temizlenmiştir. Ardından veri kümesinde yer alan rapor numarası, sorumlu mühendis, maliyet, dięer faaliyetler, notlar gibi alanların yapacađımız analiz için bir önem teşkil etmediğinden bu alanlar veri kümesinden silinmiştir. Bu işlemlerin ardından toplam 17 tane öznitelik olmak üzere 148410 veriden oluşan veri kümesi elde edilmiştir. Verilerin temizlenmesinin ardından kaza raporları tekrar incelenerek kaza, kazazede ve proje olmak üzere üç kategorik gruba ayrılmıştır.

Bu bölümde, çalışmada kullanılan veri kümesinin tablo uygulamasında nasıl tutulduęu anlatılmıştır.

Kaza Bilgileri

Raporlar sonucunda elde edilen bilgilere dayanarak kaza bilgilerinde tutulan önemli deęişken potansiyel şiddet olmuştur. Potansiyel şiddet üç alt kategoride tutulmaktadır: yüksek, orta, düşük. Buradaki alt kategoriler, meydana gelen kazanın ölüm ile sonuçlanma olasılığına göre belirlenmiştir. Aynı zamanda bu deęişken çapraz tablolama analizinde bağımlı deęişken olarak kullanılmıştır.

Kaza bilgileri kategorisi altında, iş kazasının hangi proje ve saat diliminde meydana geldiđi, kazanın çeşidinin ve meydana gelmesindeki etken faktörlerin ne olduęu hakkında bilgiler yer almaktadır. Kaza bilgilerine ait deęişkenler ve bu deęişkenlerin alt kategorileri Çizelge-2’de verilmektedir.

Kazazede Bilgileri

Kazazede kategorisi altında iş kazası geçirenlerin öğrenim durumu, yaşı, mesleđi, çalışma süresi, yaralanma türü gibi bilgiler yer almaktadır. Bu bilgiler sayesinde hangi işlerin hangi kişilere uygun olduğunun veya hangi nitelikteki kişilerin kaza geçirme ihtimalinin daha yüksek olduğunun tespiti kolayca yapılabilmektedir. Kazazede bilgilerine ait deęişkenler ve bu deęişkenlerin alt kategorileri Çizelge-3’te verilmektedir.

Çizelge-2. Kazalara ilişkin bilgiler

Değişken	Alt Kategori	Değişken	Alt Kategori	
Potansiyel Şiddet	Yüksek	Kaza Faaliyeti	Transport / araç kullanımı	
	Orta		Demir bağlama işleri	
	Düşük		El aleti / ekipman kullanımı	
Kaza Saati	00:00-04:00		Elektriksel montaj / kontrol	
	04:01-08:00		İnce işler (genel)	
	08:01-12:00		Kaynak / Isıtma / Sıcak Çalışmalar	
	12:01-16:00		Mekanik kaldırma(vinç)	
	16:01-20:00		Sıva / Alçıpan işleri	
	20:01-23:59		Kalıp montajı / sökümü	
Kaza Çeşidi	Kayma, takılma & düşme (aynı seviyeden)		Mekanik montaj	
	Yüksekten düşme		Yüksekte çalışma	
	Hareketli bir cismin / ekipmanın çarpması		Test işleri	Bakım/Onarım
	Transport / Araç hareketi		Temizlik	
	Malzeme düşmesi		Yürütme	
	Kimyasal / Biyolojik maazuzyet		Dinlenme	
	Göze yabancı cisim kaçması		Diğer	
	Keskin, sivri nesnelere temas		Kaza İş Faktörü	Yetersiz Referans Materyalleri / El Kitapçıkları
	Kaynak ışığı / UV ışığı			Yetersiz Risk / İhtiyaç Değerlendirmesi
	El aleti kullanımı	Olası Sistem Hatalarının Değerlendirilmemesi		
	Çökme / Devrilme	Yetersiz İş Planlaması		
	Yangın / Patlama	Farklı Çalışma Grupları Arasındaki İletişim Yetersizliği		
	İki nesne arasında, altında kalma / sıkışma	Yetersiz Liderlik		
	Diğer	Yetersiz Performans Ölçme ve İzleme		
	Sabit bir cisme / ekipmana çarpma	Çalışanlar Arasında Yatay İletişim Eksikliği		
	Ark Parlama	Yetersiz İşe Giriş Oryantasyonu		
	Elektrik çarpması	Üretimin Devamına Yönelik İş Baskısı		
	Aşırı sıcaklıklara (sıcak / soğuk) maruz kalma	Ekipman Kontrol Sistem Eksikliği		
Basınç salınımı	Yetersiz Kontrol, Denetim ve İzleme			
İş Makinesi Çalışmaları	İşe Yerleştirmede Yetersizlik			
Kalp Olayları	Operasyonel Hazırlığın Değerlendirilmemesi			
Duman & Gaz	Yetersiz Eğitim İçeriği			
Ülke	Cezayir	Türkiye	Tehlikelerin Belirlenememesi	
	Kazakistan	Türkmenistan	Diğer	
	Rusya	Venezuela	Mevcut Yara / Hastalık	
Kazada Ortam Faktörü	Tehlikeli Kimyasallar	Kaza Kişisel Davranış	Riski / Tehlikeyi Değerlendirememesi	
	Radyasyon		Eğitimin Yenilenmemesi	
	Enerjilendirilmiş Elektrik Sistemi		Süpervizör / Yönetici Tarafından Acele Ettirilmesi	
	Uygun Olmayan Çalışma Ortamı		Diğer	
	Hasarlı sistem / alet / ekipman		Çalışanın Acele Etmesi Gerektiğini Düşünmesi	
	Doğal Tehlikeler		Eğitimin Pratikte Uygulanmaması	
	Mekanik Tehlikeler		Güvelli Davranışların Olumsuz Pekleştirilmesi	
	Kalabalık / Kısıtlı Hareket			
	İşe uygun KKD olmaması			
	Amaca Uygun Olmayan Araç			
	Gürültü			
	Yangın / Patlama			
Koruma Altına Alınmamış Yükseklik				
Diğer				

Çizelge-3. Kazazedeye ilişkin bilgiler

Değişken	Alt Kategori		Değişken	Alt Kategori		
Kazazede Mesleği	Düz işçi		Kazazede Uyruk	Türk	Beyaz Rus	
	Kalıpçı			Özbek	Ukraynalı	
	Bekçi / Güvenlik görevlisi			Tacik	Türkmen	
	İş güvenliği ekip çalışanı			Kazak	Hintli	
	Üstyapı çalışanı			Kırgız	Azeri	
	Altyapı çalışanı			Rus	Diğer	
	Mekanik ekip çalışanı		Kazazede Çalışma Süresi	< 1 Ay	6-12 Ay	
	Havalandırmacı			1-3 Ay	12-24 Ay	
	İzolasyoncu			3-6 Ay	>24 Ay	
	Montaj Çalışmanı		Kazazede Personel Tipi	Taşeron Personeli		
	Ofis personeli			Firma Personeli		
	Marangoz			Diğer		
	İskeleci		Kazazede Yaralanma Türü	Zedelenme / Aşınma		
	Topograf			Alerjik Reaksiyon		
	Elektrikçi			Uzuv kopması		
	Kaynakçı			Gözü kaynak / UV ışığı alması		
	Sapancı			Yanık		
	Tamirci			Ezilme / Morarma		
	Şoför			Kesik / Yarık		
Operatör		Çıkık				
Mühendis		Elektrik çarpması				
Tekniker		Yabancı cisim				
Kazazede Vücut Yaralanma Bölgesi	Üst Ekstremit	Kafa / Kafatası		Kazazede Yaralanma Türü	Kırık / Çatlak	
	Alt Ekstremit	Bel / Sırt			Tahriş	
	Bütün Vücut	Boyun / Boğaz			Delinme / Çivi batması	
	Göğüs / Kaburga	Omuz			Burkulma / İncinme	
	İç Organlar	Karın	Göz Yaralanması			
	Yüz	Omurga	İç Organ Yaralanması			
	Kalça	Cinsel Organ	Kalp Krizi / Kalp Durması			
Kazazede Olay Sınıfı	Çevre Kazası		Kazazede Yaşı	20 yaş ve altı	36-40	
	İlk Yardımlı Kaza			21-25	41-45	
	İşle İlişkisiz Ölümlü Olay			26-30	46-50	
	Kayıp İş Günlü Kaza			31-35	50 yaş üstü	
	Kısıtlı İş Görememezlik					
	Tıbbi Müdahaleli Kaza					
	Yangın					
Ölümlü Kaza						

Çizelge-4: Projelere ilişkin bilgiler

bağımsızlık testi yapılmaktadır. Ki-kare değeri hesaplanırken (1) numaralı denklem kullanılmaktadır.

Değişken	Alt Kategori
Proje Tipi	Tesis
	Şehir Hastaneleri
	Alışveriş Merkezi
	Konut / Ofis
	Entegre Sağlık Kampüsü
	Yüksek Katlı Bina
	Diğer

Bağımsızlık testi için χ^2 test edilmesi gereken hipotezler şu şekilde verilmektedir

$$\chi^2 = \sum \frac{(\text{Gözlenen Sıklık} - \text{Beklenen Sıklık})^2}{\text{Beklenen Sıklık}} \quad (1)$$

3. Kullanılan Yöntemler

İş kazası istatistiklerinde kullanılan verideki değişkenler ile bu değişkenlerin alt kategorilerinin dağılımı incelenerek geçmişte meydana gelen iş kazaları analiz edilmiştir. Bu analiz, iş kazalarından dersler çıkartılarak gelecekte olabilecek iş kazalarının azaltılması açısından önem teşkil etmektedir[12]. Bu çalışmada öncelikle iş kazası verileri üzerinde temel istatistiksel analizlerden tek değişkenli sıklık ve çapraz tablolama analizi yapılmıştır. Bayes ağlarının stokastik hesaplamaları sayesinde gerçek hayattaki belirsiz durumları daha iyi temsil ettiği düşünülmektedir. Bundan dolayı temel istatistiksel analizlerin ardından belirlenen değişkenler üzerinde Bayes ağları analizi yürütülmüştür.

Tek Değişkenli Sıklık Analizi

Veri kümesinde değişkenler kategorik olarak tutulmaktadır. Temel istatistiksel analizlerden tek değişkenli sıklık analizi ile değişkenlerin alt kategorileri arasındaki dağılımlar incelenmektedir [13]. Eldeki verilerin yorumlanması ve kolay anlaşılması açısından tek değişkenli sıklık analizi önemlidir [14]. Aynı zamanda bu analiz, bir sonraki aşamada yapılacak olan çok değişkenli analizin ilk basamağını oluşturmaktadır.

Yapmış olduğumuz çalışmada tek değişkenli sıklık analizi için elektronik tablolama programından yararlanılmıştır. Bu analiz ile her bir değişkenin kategorileri arasındaki sıklık dağılımları incelenip sonuçlar grafik ve tablolar ile aktarılmıştır.

Çapraz Tablolama Analizi

Veriler üzerinde tek değişkenli sıklık analizi uygulandıktan sonra çapraz tablolama ile ikili analizler yapılmaktadır. Bu ikili analizler, her bir bağımsız değişken ile bağımlı değişken arasında ilişki olup olmadığını araştırmak için yapılmaktadır [13]. Bu ilişkinin sorgulanıp yorumlanması kısmında başarılı testlerden biri Pearson ki-kare testidir [15]. Pearson ki-kare testi ile değişkenler arasında

H_0 = Bağımlı ve bağımsız değişken arasında ilişki yoktur.

H_1 = Bağımlı ve bağımsız değişken arasında ilişki vardır.

Pearson ki-kare değeri hesaplanırken serbestlik derecesi (2) numaralı denklem yardımıyla elde edilmektedir. Serbestlik derecesinin hesabında sütunlar öznitelikleri, satırlar ise örnekleri ifade etmektedir. Bu hesaplamaların ardından hipotez testinde kullanılacak olan p değeri (olasılık değeri) hesaplanmaktadır.

$$d_f = (\text{Sütun sayısı} - 1) * (\text{Satır sayısı} - 1) \quad (2)$$

Hesaplanan p değeri 0,05'in altında ise H_0 hipotezi reddedilir ve değişkenler arasında anlamlı bir ilişkinin olduğu kabul edilir. Değişkenler arasında ilişkinin olduğu tespit edildikten sonra yapılması gereken bu ilişkinin ne kadar güçlü olduğudur. Bunun tespit edilmesi için de "Phi&Cramer's v" değeri hesaplanmalıdır [16]. Hesaplanan bu değere göre değişkenler arasındaki ilişki Tablo 5'deki gibi yorumlanmaktadır.

Bayes Ağları

Literatürde Bayes ağlarının inşaat sektöründeki iş kazası etkileri ve şiddetleri üzerine yeterli sayıda çalışma olmadığından dolayı mevcut problem üzerinde Bayes ağları kullanılarak bir tahminleme çalışması yapılmak istenmiştir. Bayes ağları, değişkenlerin düğüm ve değişkenler arasındaki ilişkilerin ise yön okları olarak gösterildiği yönlü döngüsel olasılık ağlarıdır. Genel olarak bir Bayes ağı iki önemli bölümden oluşur. Bunlardan biri değişkenlerin izdüşümünün ve bu değişkenler

Çizelge-5: Phi&Cramer's değeri yorumlama

Değer	İlişki
0,00	İlişki yok
0,00 – 0,10	Çok zayıf
0,10 – 0,20	Zayıf
0,20 – 0,25	Makul
0,25 – 0,30	Orta
0,30 – 0,35	Güçlü
0,35 – 0,40	Orta güçlü
0,40 – 0,50	Oldukça güçlü
0,50 – 0,99	Mükemmele yakın
1,00	Mükemmel

Arasındaki olasılıksal ilişkilerinin düğüm ve oklar üzerinden gösterildiği grafiksel bölümdür ve diğer bölüm değişkenlerin koşullu olasılık tablolarıdır. Bayes ağlarında grafik bölüm ağ yapısını oluşturur. İki düğümün her biri bir okla birbirine bağlandığı zaman, okun başlangıcındaki düğüm ana düğümdür ve okun sonundaki düğüm alt düğümdür [17].

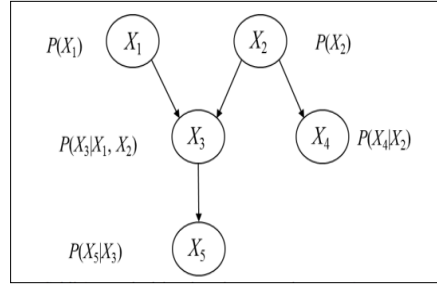
Şekil 1'de X_1 , X_2 , X_3 , X_4 ve X_5 gibi beş değişkene sahip olan bir Bayes ağı örneği gösterilmiştir. Bu ağda, X_1 ve X_2 değişkenleri, X_3 değişkeninin ebeveynidir ve X_3 değişkeni, X_5 değişkeninin ebeveynidir. Ek olarak, X_4 değişkeni, X_2 değişkeninin çocuğudur. Değişkenlerin koşullu olasılık dağılımları $P(X_1)$, $P(X_2)$, $P(X_3 | X_1, X_2)$, $P(X_4 | X_2)$ ve $P(X_5 | X_3)$.

Bir ağdaki bir değişken ile başka bir değişken arasındaki bir okun bulunmaması, bu değişkenler arasında olası bir bağlantı olmadığını ve dolayısıyla marjinal olasılık dağılımına sahip olduğunu gösterir.

Bayes ağlarında çocuk veya ebeveyn sayısı konusunda herhangi bir kısıtlama yoktur. Denklem 3'de gösterildiği gibi; N , ağdaki değişkenlerin sayısını belirtir. Bayes ağlarında her bir değişkenin koşullu olasılık dağılımlarının çarpımı ağın birleşik olasılık dağılımını vermektedir [18].

$$P(X_1, X_2, \dots, X_N) = \prod_{i=1}^N P(X_i | \text{Parent}(X_i)) \quad (3)$$

Bayes ağ yapısının oluşturulmasında iki ana yaklaşım bulunmaktadır. İlki ağ yapısının, yani ağdaki değişkenler arasındaki nedensel ilişkilerin uzmanlar tarafından belirlenmesidir.



Şekil-1: Bayes ağı

Bu yaklaşımda, değişkenler arasındaki bağımlılık ilişkilerini ve bu ilişkilerin yönlerini (ebeveyn-çocuk) belirledikten sonra, ağdaki değişkenlerin koşullu olasılıkları mevcut bir algoritma ile öğrenilir. İkinci yaklaşım ise ağ yapısını mevcut bir algoritma ile doğrudan veri setinden öğrenmektir.

Ağ yapısının öğrenilmesine ilişkin algoritmalar, kullandıkları metotlara göre kısıtlama esaslı yöntemler, arama ve puanlama esaslı yöntemler olarak ayrılmıştır. Kısıtlama esaslı yöntemlerde değişkenler arasındaki koşullu olasılık ilişkilerinin istatistiksel testler kullanılarak geçerli olup olmadığına karar verilebilir. Arama ve puanlamaya dayalı yöntemlerde, veri seti ile kurulacak olan ağ yapısı arasındaki benzerlik istatistiksel bir puanla test edilir ve veri setine en uygun ağ yapısı seçilir. Bu çalışmada ağ yapısı ise Weka programının yardımı ile belirlenen verilerden öğrenilecektir [19].

4. Bulgular

Tek Değişkenli Sıklık Analiz Bulguları

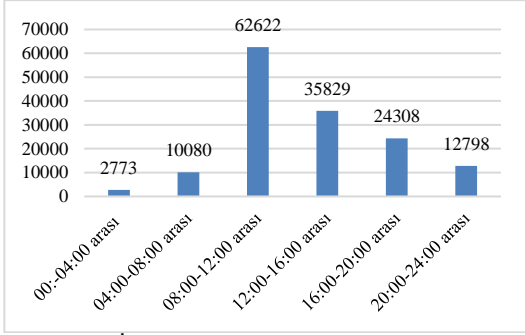
İncelenen 148410 iş kazasının %51'inin ölümle sonuçlanma ihtimali orta, %28'inin yüksek, %21'inin ise düşük olduğu görülmüştür.

Şekil-2' de gösterilen kaza saati analizine göre kazaların çoğunlukla 08:00-12:00 (%42) ve 12:00-16:00 (%24) saatleri arasında meydana geldiği görülmüştür.

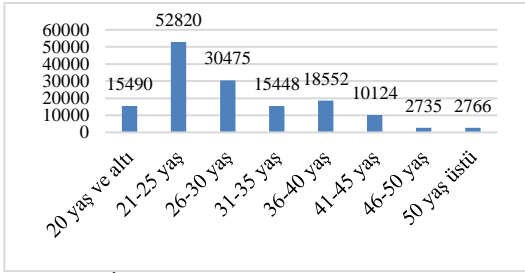
Kaza geçirenlerin %36'sını 21-25 yaş grubundaki kişiler olduğu görülmüştür. Şekil 3' deki dağılıma bakıldığında 50 yaş üstü kişilerin de bu projelerde yer alıp %2'lik kısmının iş kazası geçirdiği göze çarpmaktadır.

Kazazedenin yarasının vücudunun hangi bölgesinde meydana geldiği incelenmiş ve Çizelge-6'dan da

görülebileceği üzere yaklaşık %48’lik oranla üst ekstremité bölgesinin olduđu gözlemlenmiştir. İkinci bir analiz yapıldığında ise üst ekstremité bölgesinden en çok kol kısmından yaralanıldığı, yaralanma nedeninin ise kırık/çıkık ve kesik/yarık olduđu görülmüştür.



Şekil-2: İş kazalarının kaza saatlerine göre sıklık dağılımı



Şekil-3: İş kazalarının kazazede yaşına göre sıklık dağılımı

Çizelge-1: Kazazedenin yaralanma vücut bölgesine göre sıklık dağılımı ve yüzde oranları

Yaralanma Vücut Bölgesi	Sıklık Değeri	Yüzde Oranları
Üst Ekstremité	70894	47,769
Alt Ekstremité	20058	13,515
Karın	330	0,222
Göğüs/Kaburga	3250	2,190
İç Organlar	5918	3,988
Bütün Vücut	3	0,002
Yüz	14425	9,720
Omuz	1081	0,728
Boyun/Boğaz	1436	0,968
Omurga	3378	2,276
Cinsel Organ	186	0,125
Kalça	2228	1,501
Kafa/Kafatası	6068	4,089
Bel/Sırt	15538	10,470
Diğer	3617	2,437

Meydana gelen kazaların türleri incelendiğinde kayıp iş günlü ve tıbbi müdahaleli kazaların daha sık yaşandığı görülmüştür. Diğer alt kategoriler incelendiğinde ise %1’lik kısmının ölümlü, %0,007’lik kısmının ise işle ilişkisiz bir ölümlü sonuçlandığı Çizelge-7’de görülmektedir.

Çizelge-2: Kazazede olay sınıfına göre sıklık dağılımı ve yüzde oranları

Kazazede Olay Sınıfı	Sıklık Değeri	Yüzde Oranları
Kayıp İş Günlü Kaza	51541	34,729
Ölümlü Kaza	1620	1,092
Kısıtlı İş Görememezlik	18028	12,147
İşle İlişkisiz Ölümlü Olay	11	0,007
İlk Yardımlı Kaza	18291	12,325
Yangın	5832	3,930
Çevre Kazası	4032	2,717
Tıbbi Müdahaleli Kaza	49055	33,054

Kazaların büyük çoğunluğu (%31’i) tesis projelerinin yürütülmesi sırasında meydana gelmiştir. Bu projeler yürütülürken en çok Türk çalışanlar/işçilerin (%44) iş kazası geçirdiği ve onları da Özbek çalışanlar/işçilerin (%20,7) takip ettiği görülmüştür.

Yürütülen projelerde kaza geçirenlerin meslekleri incelendiğinde elektrikçi, kalıpcı, izolasyoncu, demirci ve iskeleci meslek gruplarının sıklıkla iş kazası geçirdiği gözlemlenmiştir. İkinci bir analizle bu meslek gruplarındaki kişilerin çalışma süreleri incelenmiştir. Analiz sonuçlarına göre kişiler işe başladığı ilk üç ay içerisinde sıklıkla iş kazası geçirmiştir. Bu, büyük olasılıkla yeterli eğitimin verilmemesi ve işe adaptasyon sürecinin zorluğundan kaynaklanmaktadır.

Çapraz Tablolama Analiz Bulguları

Bu bölümde çapraz tablolama analizi yapılarak elde edilen sonuçlar yorumlanmıştır.

Çapraz tablolama analizinde potansiyel şiddet olarak belirlenen bağımlı değişken ve 16 bağımsız değişken arasındaki ilişki sorgulanmıştır. Bu ilişki sorgulanırken bağımsız değişkenlerin “p” ve “Phi&Cramer’s V” değerleri dikkate alınmıştır. Her bir bağımsız değişken ve bağımlı değişkenin ki-kare testi sonucundaki “p” değerine bakılmıştır. İncelenen “p” değerlerinin hepsi 0,05’den küçük olduğu için H_0 hipotezi reddedilerek değişkenler arasında bir ilişkinin olduğu kabul edilmiştir. Bu 16 bağımsız değişkende bir indirgeme işlemi yapmak için sıradaki adım bağımlı ve bağımsız değişkenler arasındaki ilişkinin ne kadar güçlü olduğunu “Cramer’sV” değerlerine bakarak tespit etmektir. Bu Çizelge-8’deki gibi seçilmiştir.

Çizelge-3: Potansiyel şiddet ve anlamlı değişkenler arasındaki çapraz tablolama sonuçları

Bağımsız Değişkenler	Pearson's X^2 (df), p	Phi & Cramer's V
Kazazede Mesleği	$X^2(62)=90657,126$ p=0.000	Crv(62)=0.553
Kazazede Yaralanma	$X^2(38)=74142,132$ p=0.000	Crv(38)=0.500
Kazazede Olay Sınıfı	$X^2(14)=54403,433$ p=0.000	Crv(14)=0.428

Kazazede mesleği ve potansiyel şiddet arasındaki ilişki incelendiğinde ölümle sonuçlanma ihtimalinin en yüksek olduğu meslek grubunun %8,18'lik oranla elektrikçi olduğu ve bunu %5,15'lik oranla iskeleci meslek grubunun takip ettiği Tablo 9'da görülmektedir. Ayrıca yaşanan 148410 iş kazasının %14,32'sinin elektrikçi, %11,07'sinin kalıpcı ve %11,53'ünün diğer meslek grupları tarafından meydana geldiği tespit edilmiştir.

Kazazede yaralanma türü ve potansiyel şiddet arasındaki ilişki incelendiğinde en çok yaşanan yaralanma türünün kesik/yarık olduğu Tablo 10'da görülmektedir. Bu yaralanma türünün %13,88'inin ölümle yaralanma ihtimali orta iken %8,21'inin düşük olduğu görülmektedir. Ayrıca ölümle sonuçlanma ihtimalinin en yüksek olduğu yaralanma türünün %12,8'lik oranla kırık/çatlak yaralanmalarının olduğu analiz edilmiştir.

Kazazede olay sınıfı ve potansiyel şiddet arasındaki ilişki incelendiğinde ise ölümle sonuçlanma ihtimalinin yüksek olduğu iş kazalarının %17,93'ünün kayıp iş günlü kaza olarak nitelendirildiği Tablo 11'de görülmektedir. Meydana gelen 1620 ölümlü iş kazasının potansiyel şiddetin yüksek olarak nitelendirilmesi ile beraber tüm iş kazalarının %1,1'ini oluşturduğu görülmektedir. Ayrıca çevre kazalarının ve yangın olaylarının hepsinin ölümle sonuçlanma ihtimalinin orta olduğu dikkat çekmiştir.

Bayes Ağları Bulguları

İş kazalarını Bayes ağları ile analiz edebilmek için veri setindeki kazazede olay sınıfı, kazazede mesleği ve kazazede yaralanma türü değişkenleri çıktı değişkenleri olarak belirlenmiştir. Analizde kullanılan kazazede olay sınıfı değişkeni gerçekleşen kazaların niteliğine ilişkin bilgiler verirken, kazazede mesleği hangi çalışanlar arasında kazaların yaygın olduğu ve yaralanma türü ise hangi yaralanma türlerinin meydana geldiği hakkında yorumlama imkanı sağlayacaktır. Bu bölümdeki analizler için Weka 3.6 programından faydalanılmıştır. Çalışmada performans ölçütlerinin hesaplanmasında 10-katmanlı çapraz doğrulama

yönteminden faydalanılmıştır. Tablo 12'de kazazede olay sınıfına ait karmaşıklık matrisi verilmiş olup doğru tahmin edilen veri sayısı 119079'dur.

Kazazede olay sınıfı değişkeni altında 51541 tane kayıp iş günlü kaza verisi bulunmakta olup bu verilerin yaklaşık %83'ü doğru tahmin edilmiştir. Kalan %17'si ise kısıtlı iş görememezlik, ilk yardımcı kaza ve tıbbi müdahaleli kaza olarak yanlış tahmin edilmiştir. Ölümlü kaza, yangın, çevre kazası ve işle ilişkisiz ölümlü olay alt değişkenlerine bakıldığında bu değişkenlere ait verilerin tamamının doğru tahmin edildiği görülmektedir. En çok yanlış tahmin tıbbi müdahaleli kaza değişkeninde olmakla beraber verilerin %70'i doğru tahmin edilebilmiştir. Değişkene ait tahminler genelinde %80,23 oranında doğruluk değeri elde edilmiştir.

Yaralanma türü değişkeninin 20 alt kategorisi olup karmaşıklık matrisi 20x20 boyutundadır. Kazazede yaralanma türü karmaşıklık matrisinde doğru tahmin edilen veri sayısı diyagonal toplandan dolayı 111636'dır. Bunların en büyük kısmını 37536 veri arasından 32957 verinin doğru tahmin edilmesiyle kırık/çatlak değişkenindeki doğru tahminler oluşturmaktadır.

Doğruluk yüzdesi olarak en iyi performans elektrik çarpması ve kalp krizi yaralanma türü olmuştur ve tüm değerler doğru tahmin edilmiştir. En kötü tahmin performansı değişkene ait verilerin sadece %60'lık kısmının doğru tahmin edilmesiyle ezilme/morarma değişkeni oluşmuştur. Değişkene ait tahminler genelinde %81,96 oranında doğruluk değeri elde edilmiştir.

Kazazede mesleği değişkeninin 32 alt kategorisi olup karmaşıklık matrisi 32x32 boyutundadır. 21252 tane veri sayısı ile elektrikçi mesleğinin verisi bulunmaktadır. Ve bunların %75'lik kısmı doğru tahmin edilmiştir. 102963 doğru tahmin arasından 15977 tahmin sayısı ile elektrikçi değişkenindeki doğru tahminler sayısal olarak en büyük kısmı oluşturmaktadır. Ancak doğruluk yüzdesi olarak en iyi performans üst yapı çalışanı ve ofis personeli değişkenlerinde bulunmaktadır. Bu değişkenlerdeki tüm değerler doğru tahmin edilmiştir. En çok yanlış tahmin havalandırmacı değişkeninde oluşmuştur, verilerin %45'i doğru tahmin edilebilmiştir. Değişkene ait tahminler genelinde %76,11 oranında doğruluk değeri elde edilmiştir.

5. Sonuç ve Tartışma

Bölüm 3'te yapılan temel istatistiksel analizlerin sonucunda bu çalışmada çıktı değişkeni olarak kazazede olay sınıfı, kazazede mesleği ve yaralanma türü değişkenleri seçilmiştir. Bu değişkenler üzerinde yapılan analizler sonucunda en iyi doğruluk performansı%81,96 ile Yaralanma Türü değişkeninde deneyimlenmiştir. Tablo 13'deki Kappa istatistiği değeri analizin verdiği sonuçların tutarlı olup olmadığını söylemektedir, sonuçlara bakıldığında 0,74-0,78 oranlarıyla analizleri nütrettiği sonuçların tutarlı olduğu görülmektedir. En yüksek hata performansı olay sınıfı değişkeninde olduğu görülmektedir.

Olay sınıfı ve yaralanma türü değişkenlerinde Pekel ve arkadaşlarının yapmış oldukları çalışmaya göre daha düşük performanslar elde edilmesinin sebebi bu çalışmada bölüm 3'teki veri ön işleme sürecinin gerçekleşmiş olmasıdır [11]. Ön işleme süreci ile birlikte veri setindeki tekrar eden örüntüler çıkarılmıştır. Bu da aşırı öğrenme handikabının önüne geçmiştir.

Nihai olarak bu çalışma iş sağlığı ve güvenliği sektöründe makine öğrenmesi tekniklerinin uygulanabilirliğini göstermektedir. Gelecek çalışmalarda sadece geleneksel makine öğrenmesi algoritmaları ile çalışmak değil farklı optimizasyon teknikleri ile makine öğrenmesi algoritmalarını melezleyerek daha iyi sonuçlar üretebileceği düşünülmektedir.

Tablo 13. Performans değerleri

	Yaralanma Türü	Kazazede Mesleği	Kazazede Olay Sınıfı
Doğruluk (%)	81,96	76,12	80,24
Kappa İstatistiği	0,78	0,74	0,74
Mutlak Ortalama Hata	0,02	0,02	0,05

Bilgilendirme

Rönesans Holding tarafından desteklenen bu çalışma [12] numaralı çalışmanın genel bir yapıya uyarlanmış halidir.

Çalışmanın daha iyi hale gelmesi için görüşlerini ileten hakemlere teşekkürü borç biliriz.

Çizelge-40: Kazazede yaralanma türü ve potansiyel şiddet arasında yapılan çapraz tablolama analiz sonuçları

		Potansiyel Şiddet			Toplam
		Düşük	Orta	Yüksek	
Kazazede Yaralanma Türü	Zedelenme / Aşınma	673 (%0,45)	1578 (%1,06)	2219 (%1,5)	4470 (%3,01)
	Alerjik Reaksiyon	7 (%0,0047)	0 (%0)	1 (%0,0006)	8 (%0,0053)
	Uzuv kopması	40 (%0,03)	272 (%0,18)	503 (%0,34)	815 (%0,55)
	Gözü kaynak / UV ışığı alması	989 (%0,67)	30 (%0,02)	0 (%0)	1019 (%0,69)
	Yanık	368 (%0,25)	6090 (%4,1)	13128 (%8,85)	19586 (%13,2)
	Ezilme / Morarma	4913 (%3,31)	16961 (%11,43)	868 (%0,58)	22742 (%15,32)
	Kesik / Yarık	12190 (%8,21)	20597 (%13,88)	3599 (%2,43)	36386 (%24,52)
	Çıkkık	386 (%0,26)	806 (%0,54)	0 (%0)	1192 (%0,8)
	Elektrik çarpması	180 (%0,57)	510 (%0,68)	134 (%0,31)	824 (%0,55)
	Yabancı cisim	573 (%0,39)	1056 (%0,71)	0 (%0)	1629 (%1,1)
	Kırık / Çatlak	1987 (%1,34)	16556 (%11,16)	18993 (%12,8)	37536 (%25,3)
	Tahriş	48 (%0,03)	48 (%0,03)	0 (%0)	96 (%0,06)
	Delinme / Çivi batması	1146 (%0,77)	200 (%0,13)	0 (%0)	1346 (%0,9)
	Burkulma / İcinme	1123 (%0,76)	2380 (%1,6)	179 (%0,12)	3682 (%2,48)
	Göz Yaralanması	2150 (%1,45)	1311 (%0,88)	1020 (%0,69)	4481 (%3,02)
	İç Organ Yaralanması	0 (%0)	0 (%0)	858 (%0,58)	858 (%0,58)
	Kalp Krizi / Kalp Durması	0 (%0)	1 (%0,0006)	7 (%0,0047)	8 (%0,0053)
	Şok / Travma	1 (%0,0006)	13 (%0,008)	528 (%0,36)	542 (%0,36)
	Zehirlenme	130 (%0,09)	5420 (%3,65)	0 (%0)	5550 (%3,74)
	Diğer	4510 (%3,04)	1130 (%0,76)	0 (%0)	5640 (%3,8)
Toplam		31414 (%21,18)	74959 (%50,51)	42037 (%28,34)	148410 (%100)

Çizelge-51: Kazazede olay sınıfı ve potansiyel şiddet arasında yapılan çapraz tablolama analiz sonuçları

		Potansiyel Şiddet			Toplam
		Düşük	Orta	Yüksek	
Kazazede Olay Sınıfı	Çevre Kazası	0 (%0)	4032 (%2,72)	0 (%0)	4032 (%2,72)
	İlk Yardımlı Kaza	10530 (%7,1)	7662 (%5,16)	99 (%0,07)	18291 (%12,33)
	İşle İlişkisiz Ölümlü Olay	3 (%0,002)	1 (%0,0006)	7 (%0,004)	11 (%0,007)
	Kayıp İş Günlü Kaza	2613 (%1,77)	22315 (%15,04)	26613 (%17,93)	51541 (%34,74)
	Kısıtlı İş Görememezlik	4360 (%2,94)	6556 (%4,42)	7112 (%4,8)	18028 (%12,16)
	Tıbbi Müdahaleli Kaza	13908 (%9,37)	28561 (%19,24)	6586 (%4,44)	49055 (%33,05)
	Yangın	0 (%0)	5832 (%3,93)	0 (%0)	5832 (%3,93)
	Ölümlü Kaza	0 (%0)	0 (%0)	1620 (%1,1)	1620 (%1,1)
Toplam		31414 (%21,18)	74959 (%50,51)	42037 (%28,34)	148410 (%100)

Çizelge-12: Kazazede olay sınıfı karmaşıklık matrisi

	Ölümlü kaza	Kayıp iş günlü kaza	Kısıtlı iş görememezlik	Tıbbi müdahaleli kaza	İlk yardımcı kaza	Yangın	Çevre kazası	İşle ilişkisiz olay
Ölümlü kaza	1620	0	0	0	0	0	0	0
Kayıp iş günlü kaza	0	42806	1250	3556	3929	0	0	0
Kısıtlı iş görememezlik	0	1556	14866	1196	409	0	0	1
Tıbbi müdahaleli kaza	0	4964	480	34352	9256	0	0	3
İlk yardımcı kaza	0	367	283	2078	15561	0	0	2
Yangın	0	0	0	0	0	5832	0	0
Çevre kazası	0	0	0	0	0	0	4032	0
İşle ilişkisiz olay	0	0	1	0	0	0	0	10

Kaynakça

- [1] Baradan, S., et al., *Ege Bölgesindeki İnşaat İş Kazalarının Sıklık ve Çapraz Tablolama Analizleri*. İMO Teknik Dergi, 2016. **7345**(7370): p. 448.
- [2] Akboğa, Ö. and S. Baradan, *İnşaat Sektöründeki Ölümlü İş Kazalarının Karakteristiklerinin İncelenmesi: İzmir Alan Çalışması*, 5. İşçi Sağlığı ve İş Güvenliği Sempozyumu, İzmir, 2015: p. 215-224.
- [3] Erginel, N. and Ş. Toptancı, *İş Kazası Verilerinin Olasılık Dağılımları ile Modellenmesi*. Mühendislik Bilimleri ve Tasarım Dergisi, 2017. **5**(SI): p. 201-212.
- [4] Gümüş, R., *Türkiye’de 2015 Yılında Meydana Gelen İş Kazalarının Analizi Ve 2014 Yılı Verileri İle Karşılaştırılması*. The Journal of Academic Social Science Studies International Journal of Social Science, 2017. **55**: p. 277-287.
- [5] Eratak, Ö.D., *Yeraltı Kömür Madenciliğinde Güvenlik İçin Risk Yönetimde Analiz Ve Modelleme*. Doktora Tezi, ODTÜ.
- [6] Geçer, H.S., *Trafik Kaza Analizleri İçin Web Tabanlı Bir Karar Destek Sistemi Geliştirilmesi*. Yüksek Lisans Tezi, Sakarya Üniversitesi, 2013.
- [7] García-Herrero, S., et al., *Working Conditions, Psychological/Physical Symptoms and Occupational Accidents. Bayesian Network Models*. Safety Science, 2012. **50**(9): p. 1760-1774.
- [8] García-Herrero, S., et al., *Using Bayesian Networks to Analyze Occupational Stress Caused by Work Demands: Preventing Stress Through Social Support*. Accident Analysis & Prevention, 2013. **57**: p. 114-123.
- [9] Martín, J.E., et al., *A Bayesian Network Analysis of Workplace Accidents Caused by Falls From AHeight*. Safety Science, 2009. **47**(2): p. 206-214.
- [10] Leu, S.-S. and C.-M. Chang, *Bayesian- Network-Based Safety Risk Assessment For Steel Construction Projects*. Accident Analysis & Prevention, 2013. **54**: p. 122-133.
- [11] Pekel, E., et al. *A Bayesian Network Application in Occupational Health and Safety*. in *2018 3rd International Conference on Computer Science and Engineering (UBMK)*. 2018: pp. 239-243. IEEE.
- [12] Akşehir, Z.D., et al. *On the Analysis of Work Accidents Data by Using Data Preprocessing and Statistical Techniques*. in *2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. 2018: pp. 1-6. IEEE.
- [13] Babbie, E.R., *The basics of social research*. 2013: Cengage Learning.
- [14] Akboğa, Ö., *İnşaat İş Kazalarında Lojistik Regresyon İle Kaza Şiddetinin Modellenmesi*. Ege Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, 2014.
- [15] Sims, R.L., *Bivariate data analysis: A practical guide*. 2000: Nova Publishers.
- [16] Healey, J.F., *Statistics: A tool for social research*. 2014: Cengage Learning.
- [17] Friedman, N., I. Nachman, and D. Peér. *Learning Bayesian Network Structure From Massive Datasets: the «Sparse Candidate Algorithm*. in *Proceedings of The Fifteenth Conference on Uncertainty in Artificial Intelligence*. 1999: pp. 206-215.
- [18] Jensen, F.V., *An Introduction to Bayesian Networks*. Vol. 210. 1996: UCL Press London.
- [19] Hall, M., et al., *The WEKA Data Mining Software: an Update*. ACM SIGKDD Explorations Newsletter, 2009. **11**(1): p. 10-18