

**T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**SÜRÜCÜ KAYNAKLI TRAFİK KAZALARININ VERİ
MADENCİLİĞİ YAKLAŞIMI İLE MODELLENMESİ: SAKARYA
ÖRNEĞİ**

DOKTORA TEZİ

Zeliha Çağla KUYUMCU

İnşaat Mühendisliği Anabilim Dalı

Ulaştırma Bilim Dalı

MART 2024

**T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**SÜRÜCÜ KAYNAKLI TRAFİK KAZALARININ VERİ
MADENCİLİĞİ YAKLAŞIMI İLE MODELLENMESİ: SAKARYA
ÖRNEĞİ**

DOKTORA TEZİ

Zeliha Çağla KUYUMCU

İnşaat Mühendisliği Anabilim Dalı

Ulaştırma Bilim Dalı

Tez Danışmanı: Dr. Öğr. Üyesi Hakan ASLAN

Ortak Danışman: Prof. Dr. Nilüfer YURTAY

MART 2024

Zeliha Çağla KUYUMCU tarafından hazırlanan “Sürücü Kaynaklı Trafik Kazalarının Veri Madenciliği Yaklaşımı ile Modellenmesi: Sakarya Örneği” adlı tez çalışması 15.03.2024 tarihinde aşağıdaki jüri tarafından oy birliği ile Sakarya Üniversitesi Fen Bilimleri Enstitüsü İnşaat Mühendisliği Anabilim Dalı **Ulaştırma** Bilim Dalı’nda Doktora tezi olarak kabul edilmiştir.

Tez Jürisi

Jüri Başkanı: **Dr. Öğr. Üyesi Hakan ASLAN (Danışman)**
Sakarya Üniversitesi

Jüri Üyesi: **Prof. Dr. Ertan BOL**
Sakarya Üniversitesi

Jüri Üyesi: **Dr. Öğr. Üyesi Abdülkadir ÖZDEN**
Sakarya Uygulamalı Bilimler Üniversitesi

Jüri Üyesi: **Doç. Dr. Muhammed Yasin ÇODUR**
Erzurum Teknik Üniversitesi

Jüri Üyesi: **Doç. Dr. Ufuk ÇELİK**
Bandırma Onyedli Eylül Üniversitesi

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Sakarya Üniversitesi Fen Bilimleri Enstitüsü Lisansüstü Eğitim-Öğretim Yönetmeliğine ve Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesine uygun olarak hazırlamış olduğum “Sürücü Kaynaklı Trafik Kazalarının Veri Madenciliği Yaklaşımı ile Modellenmesi: Sakarya Örneği” başlıklı tezin bana ait, özgün bir çalışma olduğunu; çalışmamın tüm aşamalarında yukarıda belirtilen yönetmelik ve yönergeye uygun davrandığımı, tezin içerdiği yenilik ve sonuçları başka bir yerden almadığımı, tezde kullandığım eserleri usulüne göre kaynak olarak gösterdiğimi, bu tezi başka bir bilim kuruluna akademik amaç ve unvan almak amacıyla vermediğimi ve 20.04.2016 tarihli Resmi Gazete’de yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince Sakarya Üniversitesi’nin abonesi olduğu intihal yazılım programı kullanılarak Enstitü tarafından belirlenmiş ölçütlere uygun rapor alındığını, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun ortaya çıkması halinde doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi beyan ederim.

(15/03/2024)

Zeliha Çağla Kuyumcu

Destekleyen herkese...

TEŞEKKÜR

Farklı açılardan ve eleştirel bakış açısıyla ufukumun açılmasına katkı sağlayan, doktora sürecim boyunca kişisel ve akademik gelişimime yardımcı olan, iyi ki tanışmışım dediğim danışman hocam Sayın Dr. Hakan Aslan'a çok teşekkür ederim.

Tezin yöntemi, doktora öncesi çalışmadığım bir alan olduğundan, zorlandığım aşamalarda rahatlıkla danışabildiğim, bu zorlu süreçte bana her konuda destek olan sevgili ortak danışman hocam Prof. Dr. Nilüfer Yurtay'a çok teşekkür ederim.

İçinden çıkamadığım durumlarda kapısını çalabildiğim, ders notlarını karıştırarak yol bulabildiğim, iyiliğini ve olumlu düşünce biçimini örnek almaya çalıştığım hocam Prof. Dr. Kemal Selçuk Ögüt'e çok teşekkür ederim.

Her zaman yanımda olan babam Nedim Çağlar'a, annem Şengül Çağlar'a, kardeşlerim Umut Çağlar ve Eymen Hamza Çağlar'a çok teşekkür ederim.

Doktora serüvenim boyunca tüm ruhsal değişimlerimi sevgi ile karşılamaya çalışan, sabreden, yanımda olduğunu her zaman hissettiren, bu yolculukta birlikte büyüdüğüm sevgili eşim Ali Kuyumcu'ya çok teşekkür ederim.

Dedem, anneannem ve teyzem başta olmak üzere dua ve desteklerini eksik etmeyen geniş aileme, arkadaşlarıma, bu süreçte kaybettiğim ancak hep yanımda hissettiğim babaannem ve halama çok teşekkür ederim.

Doktora tez önerimden sonra yaptığımız tüm resmî ve resmî olmayan tez izleme toplantılarında kendi tecrübe ve bilgi birikimleriyle katkı sunarak çalışmama farklı bakış açıları kazandıran hocalarım Prof. Dr. Ertan Bol ve Dr. Abdülkadir Özden'e çok teşekkür ederim.

Tez savunma sınav jürimde yer alarak sundukları tüm katkılar için hocalarım Doç. Dr. Muhammed Yasin Çodur ve Doç. Dr. Ufuk Çelik'e çok teşekkür ederim.

Tezimde ve diğer akademik çalışmalarımda kullanmak üzere istediğim trafik kaza verilerini sağlayan T.C. Emniyet Genel Müdürlüğü Trafik Başkanlığı'na teşekkür ederim.

Bu çalışma, Sakarya Üniversitesi Bilimsel Araştırma Projeleri-Lisansüstü Tez Projesi kapsamında (Proje No: 2023-19-44-26) desteklenmiştir. Sakarya Üniversitesi Araştırma Dekanlığı'na, bu vesileyle katkıları için hocam Prof. Dr. Mehmet Emiroğlu'na teşekkür ederim.

Zeliha Çağla Kuyumcu

İÇİNDEKİLER

Sayfa

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ	v
TEŞEKKÜR	ix
İÇİNDEKİLER	xi
KISALTMALAR	xiii
SİMGELER	xv
TABLO LİSTESİ	xvii
ŞEKİL LİSTESİ	xix
ÖZET	xxiii
SUMMARY	xxvii
1. GİRİŞ	1
1.1. Tezin Kapsamı	1
1.2. Tezin Amaç ve Hedefleri	2
1.3. Dünya’da Trafik Kazaları	3
1.3.1. Trafik güvenliği ve trafik kazaları ile ilgili çalışmalar	4
1.3.1.1. Birliktelik Kuralları kullanılarak yapılan çalışmalar	10
1.3.1.2. Karar Ağaçları kullanılarak yapılan sınıflandırma çalışmaları	12
1.4. Türkiye’de Trafik Kazaları	15
1.4.1. Türkiye’de trafik güvenliği alanında yapılan çalışmalar	17
1.5. Sakarya’da Trafik Kazaları	20
1.6. Çalışmanın Crisp-Dm Süreci ile Yönetimi	21
1.6.1. Amacın anlaşılması	22
1.6.2. Verinin anlaşılması	22
1.6.3. Verinin hazırlanması	24
1.6.4. Modelleme	24
1.6.5. Değerlendirme	25
1.6.6. Uygulama	25
2. YÖNTEM VE VERİ	27
2.1. Veri Madenciliği	27
2.1.1. Birliktelik Kuralları	29
2.1.2. Karar Ağaçları	31
2.2. Çalışma Alanı	32
2.3. Veri Seti	35
2.3.1. Veriye ait tanımlayıcı istatistikler	35
2.4. Veri Ön işleme Süreci	54
2.4.1. Veri temizleme	56
2.4.2. Veri birleştirme	57
2.4.3. Veri dönüştürme	58
2.5. Birliktelik Kuralları İçin Veri Ön İşleme Süreci	59
2.5.1. Birliktelik Kuralları elde edilirken kullanılan veriye ait tanımlayıcı istatistikler	59
2.6. Karar Ağaçları İçin Veri Ön İşleme Süreci	63

2.6.1. Karar Ağacı elde edilirken kullanılan veriye ait tanımlayıcı istatistikler.	63
3. VERİ ANALİZLERİ.....	67
3.1. Birliktelik Kuralları ile Veri Analizi	67
3.2. Karar Ağaçları ile Veri Analizi	74
3.2.1. Elde edilen Karar Ağacı modeli başarımlı düzeyi.....	86
4. BULGULAR VE TARTIŞMA	93
4.1. Birliktelik Kuralları Sonuçlarının Literatürle Değerlendirilmesi	93
4.2. Karar Ağaçları Sonuçlarının Literatürle Değerlendirilmesi	95
5. SONUÇ VE ÖNERİLER.....	97
KAYNAKLAR.....	103
EKLER.....	111
ÖZGEÇMİŞ.....	115

KISALTMALAR

ABD	: Amerika Birleşik Devletleri
CHAID	: Chi-square Automatic Interaction Detection
CRISP-DM	: Cross industry standard process for data mining
DSÖ	: Dünya Sağlık Örgütü
EGM	: Emniyet Genel Müdürlüğü
KGM	: Karayolları Genel Müdürlüğü
MNL	: Multinomial Lojistik Regresyon Modeli
TÜİK	: Türkiye İstatistik Kurumu
DP	: Doğru pozitif, doğru tahmin edilen pozitif veri sayısı
DN	: Doğru negatif, doğru tahmin edilen negatif veri sayısı
YN	: Yanlış negative, yanlış tahmin edilen negatif veri sayısı
YOGT	: Yıllık ortalama günlük trafik
YP	: Yanlış pozitif, yanlış tahmin edilen pozitif veri sayısı

SİMGELER

G	: Gerçekleşen gözlem sayısı
B	: Beklenen gözlem sayısı
χ^2	: Ki-kare istatistiği
χ_0^2	: Ki-kare kritik
α	: Anlamlılık düzeyi

TABLO LİSTESİ

Sayfa

Tablo 1.1. Trafiğe kayıtlı taşıt, kaza sayısı, 2015-2022 (Türkiye İstatistik Kurumu, 2023).	15
Tablo 1.2. Trafik kazalarına bağlı olarak ölen ve yaralananların sayısı, 2015-2022 (Türkiye İstatistik Kurumu, 2023).	16
Tablo 1.3. Ölümlü yaralanmalı trafik kazasına neden olan kusur sayısı, 2015-2022 (Türkiye İstatistik Kurumu, 2023).	16
Tablo 1.4. Ölümlü yaralanmalı trafik kazasından etkilenen kazazedeler, 2022 (Türkiye İstatistik Kurumu, 2023).	17
Tablo 1.5. Sakarya’da 2015-2022 yıllarındaki taşıt ve kaza sayıları (Türkiye İstatistik Kurumu, 2023).	20
Tablo 1.6. Sakarya’da 2015-2022 yıllarında kazalarda ölen ve yaralanan sayıları (Türkiye İstatistik Kurumu, 2023).	21
Tablo 2.1. Trafik kaza veri setinin bir bölümü.	31
Tablo 2.2. Emniyetten alınan veri dosyalarının ilkinde yer alan öznitelikler.	55
Tablo 2.3. Emniyetten alınan veri dosyalarının ikincisinde yer alan öznitelikler.	56
Tablo 2.4. Birliktelik Kurallarının elde edildiği veri setinin tanımlayıcı istatistikleri.	60
Tablo 2.5. Karar Ağacı için hazırlanan veri seti.	65
Tablo 3.1. Kaza ile ilgili etkenler arasında lift değerlerine göre sıralanan 9 önemli kural.	69
Tablo 3.2. Kaza ve çevresel etkenler ile ilgili etkenler arasında lift değerlerine göre sıralanan 10 önemli kural.	70
Tablo 3.3. Kaza ve sürücü ile ilgili etkenler arasında lift değerlerine göre sıralanan 4 önemli kural.	72
Tablo 3.4. Kaza ve yol ile ilgili etkenler arasında lift değerlerine göre sıralanan 11 önemli kural.	72
Tablo 3.5. Kaza ve taşıt ile ilgili etkenler arasında lift değerlerine göre sıralanan 4 önemli kural.	74
Tablo 3.6. Sürücü kusuru ve sürücü kaza sonucu durumu özniteliğindeki gerçekleşen değerler.	80
Tablo 3.7. Sürücü kusuru ve sürücü kaza sonucu durumu özniteliğindeki gerçekleşen ve beklenen değerler.	81
Tablo 3.8. Tüm özniteliklerin Ki-kare istatistik, kritik ve serbestlik dereceleri.	82
Tablo 3.9. Tüm değişkenler için kurulan Karar Ağacı modelinin başarımı.	87
Tablo 3.10. Her sınıfa ait gerçekleşen ve tahmin edilen verilere göre doğru ve yanlış sınıflandırma değerleri.	87
Tablo 3.11. Seçilen değişkenler için kurulan Karar Ağacı modelinin başarımı.	91

ŞEKİL LİSTESİ

Sayfa

Şekil 1.1. 2009-2018 yılları arasında trafik kazalarında, 1000 kişide ortalama yaralanmalar (Injuries in the European Union 2009-2018, 2021).	3
Şekil 1.2. CRISP-DM sürecinin yaşam döngüsü, Eckerson ve ark. (2000)'dan uyarlanmıştır.	22
Şekil 2.1. Veri madenciliğini besleyen alanlar, Han ve ark. (2012)'den uyarlanmıştır.	27
Şekil 2.2. Veri madenciliği aşamaları, Han ve ark. (2012) ve Yu ve ark. (2019)'dan uyarlanmıştır.	28
Şekil 2.3. Veri madenciliği grupları, Fian Tabea (2021) ve Maimon ve Rokach (2005)'ten uyarlanmıştır (Fian Tabea, 2021; Maimon & Rokach, 2005).	29
Şekil 2.4. Karar Ağacı gösterimi. Anvari ve ark. (2017)'den uyarlanmıştır (Anvari vd., 2017).	32
Şekil 2.5. Sakarya şehrinin Dünya'daki konumu (Google Earth, 2023).	33
Şekil 2.6. Sakarya'nın Türkiye'deki konumu (Google Earth, 2023).	33
Şekil 2.7. Sakarya'nın ilçeleri (T.C. Milli Savunma Bakanlığı Harita Genel Müdürlüğü, 2023).	34
Şekil 2.8. Trafik kazalarının sonuçlarına göre histogramı.	36
Şekil 2.9. 2015-2020 yıllarında Sakarya'da gerçekleşen trafik kazalarının yıllara göre histogramı.	36
Şekil 2.10. Trafik kazalarının haftanın günlerine göre histogramı.	37
Şekil 2.11. Trafik kazalarının ilçelere göre histogramı.	37
Şekil 2.12. Trafik kazalarının yerleşim yerlerine göre histogramı.	38
Şekil 2.13. Trafik kazalarının yol tipine göre histogramı.	38
Şekil 2.14. Trafik kazalarının yol kaplamasına göre histogramı.	39
Şekil 2.15. Trafik kazalarının yol sınıfına göre histogramı.	39
Şekil 2.16. Trafik kazalarının yatay yol geometrisine göre histogramı.	40
Şekil 2.17. Trafik kazalarının düşey yol geometrisine göre histogramı.	40
Şekil 2.18. Trafik kazalarının kavşak türüne göre histogramı.	41
Şekil 2.19. Trafik kazalarının geçit türüne göre histogramı.	41
Şekil 2.20. Trafik kazalarının diğer yol unsurlarına göre histogramı.	42
Şekil 2.21. Trafik kazalarının gün ışığına göre histogramı.	42
Şekil 2.22. Trafik kazalarının hava durumuna göre histogramı.	43
Şekil 2.23. Trafik kazalarının yol yüzeyine göre histogramı.	43
Şekil 2.24. Trafik kazalarında ilk yardım histogramı.	44
Şekil 2.25. Trafik kazalarının oluş türüne göre histogramı.	44
Şekil 2.26. Trafik kazalarına karışan taşıt sayısına göre histogramı.	45
Şekil 2.27. Trafik kazalarının gerçekleştiği çarpışma yerine göre histogramı.	45
Şekil 2.28. Trafik kazalarının gerçekleştiği yerlerde yol kusurları histogramı.	46
Şekil 2.29. Trafik kazalarının gerçekleştiği yerlerde otokorkuluk varlığı histogramı.	46

Şekil 2.30. Trafik kazalarının gerçekleştiği yerlerde yaya kaldırımı varlığı histogramı.	47
Şekil 2.31. Trafik kazalarının gerçekleştiği yerlerde trafik ışığı varlığı histogramı.	47
Şekil 2.32. Trafik kazalarının gerçekleştiği yerlerde aydınlatma varlığı histogramı.	48
Şekil 2.33. Trafik kazası geçiren sürücülerin cinsiyet histogramı.	48
Şekil 2.34. Trafik kazası geçiren sürücülerin eğitim durumuna göre histogramı.	49
Şekil 2.35. Trafik kazası geçiren sürücülerin alkol kullanımına göre histogramı.	49
Şekil 2.36. Trafik kazası geçiren sürücülerin psikolojik/fiziksel durumuna göre histogramı.	50
Şekil 2.37. Trafik kazası geçiren sürücülerin emniyet kemeri/kask kullanımına göre histogramı.	50
Şekil 2.38. Trafik kazası geçiren sürücülerin kaza sonucuna göre histogramı.	51
Şekil 2.39. Trafik kazasına karışan taşıtların türel histogramı.	51
Şekil 2.40. Trafik kazasına karışan taşıtların kullanım tipine göre histogramı.	52
Şekil 2.41. Trafik kazası geçiren sürücülerin yaş histogramı.	52
Şekil 2.42. Sakarya’da ehliyet sahibi olanların yaş dağılımı – 2020 yılı.	53
Şekil 2.43. Trafik kazası geçiren sürücülerin yaş grubuna göre histogramı.	53
Şekil 2.44. Trafik kazası geçiren sürücülerin kusurlarına göre histogramı.	54
Şekil 2.45. Veri ön işleme süreci	56
Şekil 2.46. Trafik kazalarında hayatını kaybeden kazazede sayıları (2015-2020).	60
Şekil 2.47. Trafik kazalarında yaralanan kazazede sayıları (2015-2020).	61
Şekil 2.48. Trafik kazalarına karışan sürücülerin kaza sonucu durumları (2015-2020).	61
Şekil 2.49. Kaza oluş türlerinin yıllara göre dağılımı.	62
Şekil 2.50. Önemli Birliktelik Kuralları elde edebilmek için oluşturulan beş ana grup.	62
Şekil 2.51. Veri sınıflandırma süreci: (a) Öğrenme: Eğitim verileri bir sınıflandırma algoritması tarafından analiz edilmektedir. (b) Sınıflandırma: Test verileri, sınıflandırma kurallarının doğruluğunu tahmin etmek için kullanılmaktadır, Han ve ark. (2012)’den uyarlanmıştır.	64
Şekil 3.1. Kaza ile ilgili etkenler arasında bulunan 13 Birliktelik Kuralına ait saçılım grafiği.	69
Şekil 3.2. Kaza ve çevresel etkenler ile bulunan 72 Birliktelik Kuralına ait saçılım grafiği.	70
Şekil 3.3. Kaza ve sürücü ile ilgili etkenler ile bulunan 18 Birliktelik Kuralına ait saçılım grafiği.	71
Şekil 3.4. Kaza ve yol ile ilgili etkenler ile bulunan 128 Birliktelik Kuralına ait saçılım grafiği.	72
Şekil 3.5. Kaza ve taşıt ile ilgili etkenler ile bulunan 17 Birliktelik Kuralına ait saçılım grafiği.	73
Şekil 3.6. Karar Ağacında kullanılan veri setindeki sürücülerin kaza sonucu durumlarının yıllara göre oransal olarak dağılımı.	76
Şekil 3.7. Karar Ağacında kullanılan veri setindeki sürücü kusurlarının yıllara göre oransal olarak dağılımı.	76
Şekil 3.8. Entropi sayısı formülü kullanılarak elde edilen öznitelik ağırlıkları sıralaması.	77
Şekil 3.9. Gini indeksi formülleri kullanılarak elde edilen öznitelik ağırlıkları sıralaması.	78
Şekil 3.10. Ki-kare formülleri kullanılarak elde edilen öznitelik ağırlıkları sıralaması.	78

Şekil 3.11. Ki-kare kritik ve Ki-kare istatistik değerlerinin görüldüğü Ki-kare dağılım grafiği, Neil A. Weiss (2011)'den uyarlanmıştır.	83
Şekil 3.12. Tüm özniteliklerin yer aldığı veri setiyle oluşturulan Karar Ağacı.....	84
Şekil 3.13. Seçilen özniteliklerin yer aldığı veri setiyle oluşturulan Karar Ağacı.....	89

SÜRÜCÜ KAYNAKLI TRAFİK KAZALARININ VERİ MADENCİLİĞİ YAKLAŞIMI İLE MODELLENMESİ: SAKARYA ÖRNEĞİ

ÖZET

Ulaştırma; insanların, eşyaların ve hayvanların bir yerden başka bir yere güvenli taşınmasıdır. Nüfus artışına ve ekonomik gelişmeye bağlı olarak artan hareketlilik talepleri, karayollarındaki özel taşıt kullanımı ve yük taşıma artışını beraberinde getirmektedir. Buna bağlı olarak trafik güvenliğinin sağlanması daha karmaşık yaklaşımların, uygulamaların, analiz süreçlerinin geliştirilmesini zorunlu kılmaktadır. Dünyada ve Türkiye’de trafik kazaları insan hayatına doğrudan ve dolaylı olarak etki eden en büyük sorunlardan biridir. Bileşmiş Milletler’in hedefi, 2030 yılına kadar trafik kazası kaynaklı yaralanma ve ölümleri %50 oranında azaltmak, 2050 yılına kadar da mümkün ise sıfırlamaktır. Türkiye’de ulaştırmadan sorumlu kurumların (Üniversiteler, Karayolları Genel Müdürlüğü (KGM), Emniyet Genel Müdürlüğü (EGM), Jandarma Genel Komutanlığı, Yerel Belediyeler, Millî Eğitim Bakanlığı (MEB)) ulaşım planlaması, trafik yönetimi ve güvenliği ile ilgili çeşitli uygulamaları bulunmaktadır. Bununla beraber, “sıfır can kaybı vizyonu”na ulaşmak için yönetilmesi gereken sürecin, alınması gereken önlemlerin kapsamlı ve derinlikli bir yapıda olduğu, son yıllardaki kaza istatistiklerinden görülebilmektedir.

Bir veya birden çok etkenin bir araya gelmesiyle oluşan trafik kazalarının meydana getirdiği olumsuzlukların etkilediği insan sayısı oldukça fazladır. 2022 verilerine göre trafik kazaları; ülkemizde her gün 14 kişinin hayatını kaybetmesi, 790 kişinin yaralanması, 2837 maddi hasarlı kaza yaşanması ile sonuçlanan, halk sağlığı sorunu olarak adlandırılan çok önemli bir konudur. Can kayıpları, kaza sonrası bedensel engelli hale gelen insanlar ve maddi kayıplar ülke ekonomisini önemli ölçüde etkilemektedir. Bunun yanında sosyal yaşantının da etkilenmesi yollarda güvenlik kaygılarının oluşumuna, kazayı yaşayan kişi ve yakınlarında psikolojik rahatsızlıklar görülmesine de neden olabilmektedir. Günümüzde “affeden yollar” terimi ile yol kullanıcı hatalarına hoşgörü gösterebilecek yollar tasarlama ve inşa etme konusunda adımlar atılmaya başlansa da hala yol güvenliği için iyileştirilmesi gereken birçok unsur olduğu yeterince düşüş olmayan kaza sayılarından anlaşılmaktadır. Trafik kazalarının meydana gelmesine neden olan hatalar; insan (sürücü, yaya, yolcu), taşıt ve yol olarak 3 gruba ayrılmaktadır. Bu hatalardan en önemlisinin, 2011-2022 yılları arasında %97-99 arasında değişen oranlarda insan etkeni, bu etkenlerin içinde ise %87-90 arasında değişen oranlarda sürücü etkeni olduğu, KGM’nin 2022 yılında hazırladığı raporda belirtilmiştir.

Bu çalışmada, Sakarya ilinde ölümlü ve yaralanmalı kazaların veri madenciliği yöntemleriyle incelenmesi amaçlanmıştır. Çalışmada iki farklı analiz yapılmış; birinci aşamada Birliktelik Kuralları yöntemi kullanılarak, trafikle ilgili risk etkenlerinin olası ilişkileri araştırılırken, ikinci aşamada Karar Ağaçları kullanılarak geliştirilen model ile yalnızca otomobillerin karıştığı kazalarda, kazaya karışan ve hatalı olan sürücünün kaza sonucundaki durumu (yaralı, ölü, sağlam) tahmin edilmeye çalışılmıştır.

Bu çalışmada gerçekleştirilen Birliktelik Kuralları analizleri ile trafik kaza veri setinde yer alan hangi durumların birlikte en fazla olduğu, hangi ögelerin (öznitelikler içerisindeki sınıfların), bütün içerisinde daha yaygın olduğu belirlenmiştir.

Karar Ağaçları ise, veri analizi süreçlerinde sınıflandırma yapılacak durumlarda en çok kullanılan yöntemlerden biridir. Karar Ağaçları, her düğümden ayrılan kolların o alandaki soruya yanıt verdikleri, her yaprağın da bir karar ve sonucu ifade ettiği algoritmalarıdır. Kök düğümün seçilmesi, ağacın dallanması ve budanması kararları ve sonucunda sınıflandırmanın yapılmasına bağlı olarak Karar Ağacı kurulurken kullanılan algoritmaya bağlıdır. Bu çalışmada, kullanılan trafik kaza veri seti incelendiğinde, kaza şiddetinin sınıflandırılmasına karar verilmiştir. Veri setinde, kaza sonucu ile ilgili 4 farklı öznitelik bulunmaktadır. Bunlardan biri kazanın sonucu öz niteliği olup içinde ölümlü ve yaralanmalı olmak üzere 2 sınıf yer almaktadır. Kazalardan %0,9'u ölümlü geri kalanı yaralanmalı olduğu için sınıflandırma tahmini yapmak çok zordur. Kaza sonucu ile ilgili diğer öznitelikler (ayrı birer öznitelik olmak üzere) sürücü, yolcu ve yayanın kazaya bağlı olarak karşılaştıkları sonuçlardır. Her kazada yolcu ve yaya olmadığından, trafik kazalarında en çok etkilenen ve istatistiklere göre kazanın olmasına en fazla neden olan kişiler sürücüler olduğundan bu çalışmada sürücü kaza sonucu, sınıflandırma yöntemlerinden biri olan Karar Ağaçları ile tahmin edilmiştir. Veri kümesinde bulunan bağımlı ve bağımsız öznitelikler kategorik (nominal) içerikte olduklarından ve tezde anlatılan diğer nedenlerden dolayı Ki-kare otomatik etkileşim algılama (CHAID) algoritması kullanılmıştır. Bu tez çalışmasındaki ikinci analizin amacı kazaya karışan, kazada kusuru olan sürücülere ait özellikleri ve kaza ile ilgili diğer durumlara bağlı olarak sürücünün kaza sonucu durumunu sınıflandırmak olduğundan, literatürdeki çalışmaların da bulgusu olarak bilinen, sürücünün kaza sonucuna (yaralanma şiddetine) etkisi en büyük öz niteliğin taşıt türü olması nedeniyle bu öz niteliğin etkisi çalışma dışında bırakılmıştır. Bu yaklaşım doğrultusunda, taşıt türleri içerisinde veri setinde sahip olduğu %55 oran ile en yüksek değere sahip olan otomobil kazalarındaki sürücü kaza sonuçları, sürücü öz niteliklerine bağlı olarak analiz edilmiştir. Bu anlamda literatürde bulunan çalışmalar, tez içeriğinin ilgili bölümlerinde detaylı bir şekilde tartışılmıştır.

Sürücü kaza şiddeti sınıflandırması ve tahmin modeli geliştirmek için kullanılan CHAID Karar Ağacı oluşturma sürecinde, ilk önce tüm özniteliklerin olduğu veri seti ile Ki-kare değerleri hesaplanmış, sonrasında bazı özniteliklerin baskın etkilerinin sürücü ile ilgili durumları gizlemesine bağlı olarak bu baskın olan öznitelikler veri setinden çıkarılmıştır. Kalan öznitelikler ile çıktı öz niteliği (bağımlı değişken) olan sürücü kaza sonucu durumu (sağlam-yaralı-ölü) arasındaki Ki-kare değerleri hesaplanarak ağacın kök düğümünde yer alan yani bağımlı öz nitelikle arasındaki ilişkisi en büyük olan öznitelik sürücü kusuru olarak belirlenmiştir. Sürücü kusuru öz niteliğini, bağımlı öz nitelikle aralarındaki ilişki büyüklüğüne göre; sürücü alkol sonucu, yol bölümü, saat dilimi, sürücü yaşı, sürücü belgesi varlığı, otomobil yaşı, sürücü eğitim durumu ve gün ışığı öznitelikleri takip etmiştir. Oluşturulan Karar Ağacına göre; alkollü taşıt kullanan sürücülerin tamamının yaralandığı, öndeki taşıtı yakın takip eden sürücülerin kaza sonucunun sürücü yaş grubuna bağlı olarak sınıflara ayrıldığı, dönüş kurallarına uymayan sürücülerin kaza sonucunun cinsiyetlerine göre değişim gösterdiği, taşıtın hızını yol ve hava koşullarına uydurmayan sürücülerin kaza sonucunun ilk olarak yol bölümüne göre değiştiği, sonraki aşamada 3 ve 4 kollu kavşaklarda cinsiyete, erkek olanların içinde de yaş grubuna göre değiştiği, dönel kavşaklarda saat dilimine göre sürücü kaza sonucunun değiştiği, kavşak ve geçitlerde

yol verme kuralına uymayan sürücülerin kaza sonucunun ilk olarak sürücü eğitim durumuna, lise mezunu olan sürücüler içerisinde ise kullandıkları taşıtın yaşına göre değiştiği, sürücü kusurunun hatalı şerit değiştirme olduğu kazalarda sürücü kaza sonucunun öğrenim durumuna, ilkokul mezunu olan sürücülerde gün ışığı durumuna ve gündüz ise kullandıkları taşıtın yaşına göre değiştiği görülmektedir.

MODELLING DRIVER-RELATED TRAFFIC ACCIDENTS THROUGH DATA MINING APPROACH: THE CASE OF CITY OF SAKARYA

SUMMARY

Transportation is the purposed movement of people, goods, and animals from one place to another. Transportation types are classified as road, rail, sea, and air. In our country, the most widely used type of them has been road transportation for years. The heavy traffic on the roads, coupled with the increase in private vehicle usage, makes traffic safety as a challenging issue due to the increases in the risk of collisions.

Traffic accidents are a major issue affecting human life directly and indirectly, both globally and in Türkiye. The United Nations aims to reduce traffic accident-related injuries and deaths by 50% by 2030 and to eliminate them entirely by 2050. Responsible institutions for transportation in Türkiye (General Directorate of Highways, General Directorate of Security, The Gendarmerie General Command, Municipalities, Ministry of National Education) have various practices related to transportation planning, traffic management, and traffic safety. However, achieving the "vision of zero fatalities" still requires numerous measures as evident from recent accident statistics.

Traffic accidents, generally resulting from the combination of two or more factors, cause multidimensional damages. According to 2022 data in Türkiye, traffic accidents lead to a daily average of 14 fatalities, 790 injuries, and 2837 property-damage-only incidents, posing a significant public health problem. Loss of life, post-accident physical disabilities, and property losses significantly impact both the country's economy and the social life of the people leading to safety concerns on the roads and psychological distress for those involved in accidents and their families.

Despite initiatives in recent years such as the design of "forgiving roads" to tolerate user errors, there are still many aspects that need improvement as far as road safety is concerned, as indicated by recent accident statistics. Errors leading to traffic accidents are categorized into three main related groups: human (driver, pedestrian, passenger), vehicle, and road. The most crucial of these errors is the human related factors, ranging between 97-99% from 2011 to 2022, with the driver factor accounting for 87-90%, as stated in a report prepared by the KGM in 2022.

Türkiye with the population of 86,907,000.00 has 259,072 km-long highways. The number of vehicles registered in traffic is 28,183,745.00 among them 14,967,044.00 being automobiles according to the 2023 records.

The average numbers of cars per 1000 people for Türkiye and EU countries are 167 and 560, respectively. While the number of deaths per 1 million cars in Türkiye is 366 people, the related figure for EU countries is 76 reflecting the need to make significant progress in road safety engineering and applications.

This study aims to examine fatal and injury accidents in the city of Sakarya using data mining analysis methods. Two different analyses were conducted in the study. The

first involved exploring possible relationships among traffic-related risk factors using association rules. The second used decision trees to predict the severity of accidents involving only cars.

The study identified which conditions co-occur most frequently in the traffic accident dataset and which elements (classes within attributes) were more prevalent overall. Decision trees are widely used for classification in situations like traffic accidents. In decision trees, while each branch departing from a node answers a question in that area, each leaf represents a decision outcome. The selection of the root node, branching and pruning decisions, and the subsequent classification depend on the algorithm used when building the decision tree. In this study, the traffic accident dataset was examined, and it was concluded that accident severity could be classified. Among the four attributes related to the accident outcome in the dataset, one is the outcome attribute containing two classes: fatal and non-fatal injuries. Being only 0.9% of the accidents fatal makes predictions for the remaining injuries is challenging. The other attributes related to the accident outcome (treated as separate attributes) are associated with drivers, passengers, and pedestrians. Since numerous accident data sets do not involve information regarding passengers and pedestrians, drivers are regarded as the most affected and statistically contributing the most to the accidents. Therefore, in this study, the driver's accident outcomes were predicted using decision trees, a classification method. Some of the decision tree algorithms are CHAID (Chi Square Automatic Interaction Detection), CART (Classification and Regression Trees), C 4.5, ID3 and Random Tree. Since the dependent and independent variables are categorical (nominal), the CHAID algorithm was deemed suitable within the scope of this thesis.

As the purpose of the second analysis in this thesis is to classify the characteristics of the faulty drivers involved in the accidents, and since the vehicle-size attribute makes it difficult to fulfill this purpose, the vehicle with the largest weight in the dataset (55%) was selected as automobiles. The reason why this objective is challenging is that there are studies suggesting that vehicle size affects injury severity. It has been reported in the literature that the most closely related vehicle characteristic to injury severity is the size of the vehicle. Although there is a high correlation between weight and length, weight is reported to be more common in studies.

As far as the analysis is concerned, chi-square values were, first, calculated for all attributes, and then, due to the dominant effect of some attributes obscuring driver-related parameters, it was decided to exclude these attributes from the dataset. With this regard, chi-square values between the remaining and the output attributes (driver's accident outcome) were calculated and determined. The attribute with the greatest relationship magnitude at the root node of the tree, indicating the largest relationship with the dependent attribute, was obtained as driver error.

Data set is divided into two parts to create the decision trees of the model; 70% of which as training and 30% as test data. This distinction was made linearly and as the accidents in the data set were recorded and presented according to the date/year on/in which they occurred, the first 5 years of the available data have been taken as education, and the last 2 years as testing data.

To create a CHAID decision tree and classify the driver related severity of the accidents, chi-square values were first calculated with the dataset containing all attributes. Subsequently, due to the dominant influence of certain attributes leading to the concealment of driver-related situations, it was decided to remove these dominant attributes from the dataset. Remaining attributes were then used to calculate chi-square

values between the output attribute (dependent variable), which is the driver's posed severity level of the accident, and the remaining attributes. According to the magnitude of the relationship between the driver's fault attribute and the dependent attribute, the following attributes were ranked: driver under the influence of alcohol, road section, time zone, driver's age, presence of driver's license, age of the automobile, driver's education level, and daylight. According to the decision tree created, drivers using alcohol were all injured, and drivers closely following the preceding vehicle had their accident outcomes classified according to the driver's age group. The outcomes of drivers who did not adhere to turning rules varied according to their genders, and the accident outcomes of drivers who did not adjust their speed to the road and weather conditions primarily depended on the road section. In the next stage, the casualties were observed to change according to the gender for 3 and 4-way intersections. Among male drivers, they varied according to age group. In roundabouts, on the other hand, while the driver's accident outcome changed according to the time zone, at intersections and crossings, the outcomes of drivers not yielding through the right-of-way rule changed according to the driver's education level. Among high school graduates, the outcome also varied according to the age of the vehicle they used. In accidents where the driver's fault was improper lane changing, the educational level for accidents involving elementary school graduates played an important role, and within this group, the results varied according to the daylight conditions and the age of the vehicle used.

1. GİRİŞ

Ulaşım talebi, insanların temel gereksinimlerini karşılamak istemeleri sonucunda ortaya çıkmakla birlikte, özel gereksinimleri ile de bu talep artabilir. Bu talep, zorunlu yolculuklardan olan, ev-iş, ev-okul, diğer-iş veya isteğe bağlı (özel gereksinim) yolculuklar olan ev-hastane, ev-diğer, diğer-diğer vb. amaçlı hareketliliklerden oluşmaktadır. Ulaştırma türleri karayolu, demiryolu, denizyolu ve havayolu olarak sınıflandırılmaktadır. Ülkemizde en çok kullanılan ulaşım türü karayoludur. 2021 yılı verilerine göre ülkemizde yurt içi yolcu taşımacılığının %92,7'si, yük taşımacılığının %90'ı karayolu ile yapılmıştır (Ulaşan ve Erişen Türkiye, 2022).

Nüfus ve motorlu taşıt sayısının artışı, doğrudan trafikteki taşıt ve insan hareketliliğinin artmasına, özellikle kent içinde trafik sıkışıklığı, gürültü, çevre kirliliği gibi dolaylı etkilerin yanı sıra insan hayatını doğrudan etkileyen trafik kazalarının da artmasına neden olmaktadır.

Trafik kazası, karayolunda bir taşıtın başka bir taşıt ya da yayaya, hayvana veya bir engele çarpması olayıdır. Trafik kazaları, maddi hasar, yaralanma ya da ölümle sonuçlanabilir ('Trafik kazası', 2022). Bu somut fiziksel neticeler dışında, kazalar sonucunda kazazedelerde psikolojik sarsıntılar da (travma) meydana gelmektedir (The Organisation and Content of Trauma Memories in Survivors of Road Traffic Accidents | Elsevier Enhanced Reader, 2022).

Tarihte ilk ölümlü trafik kazası 1869 yılında İrlandalı gökbilimci Mary Ward'ın buharlı arabadan düşmesi ve taşıt tarafından ezilmesi sonucunda gerçekleşmiştir (Fallon & O'Neill, 2005).

Bu tezde, kaza kara noktaları yerine kazanın konumundan bağımsız olarak kaza kara örüntülerinin (birlikte en sık meydana gelen durumların) bulunması amaçlanmıştır.

1.1. Tezin Kapsamı

Tez çalışması kapsamında, Sakarya'da 2015-2020 yılları arasında, Emniyet Genel Müdürlüğü'nden (EGM) alınan ölümlü-yaralanmalı trafik kazalarına ait tutanakların işlendiği excel dosyasındaki veriler analiz edilmiştir. Jandarma Genel Komutanlığı'nın

veri arşivinde olan ölümlü-yaralanmalı trafik kazaları bu çalışma kapsamında yer almamaktadır. Çalışmada, trafik kazalarına ait genel bilgiler, kazaların oluşmasına etki eden özniteliklerin tipleri ve özellikleri, kullanılan trafik kaza verilerine ait tanımlayıcı istatistikler, ilgili veri setinin analizler için hazırlığı (ön işleme süreci), kaza örüntülerinin Birliktelik Kuralları ile elde edilmesi ve otomobillerin karıştığı kazalarda sürücü kaza şiddetinin tahmin edilmesi için kullanılan yöntem olan Karar Ağaçları analizi yer almaktadır.

1.2. Tezin Amaç ve Hedefleri

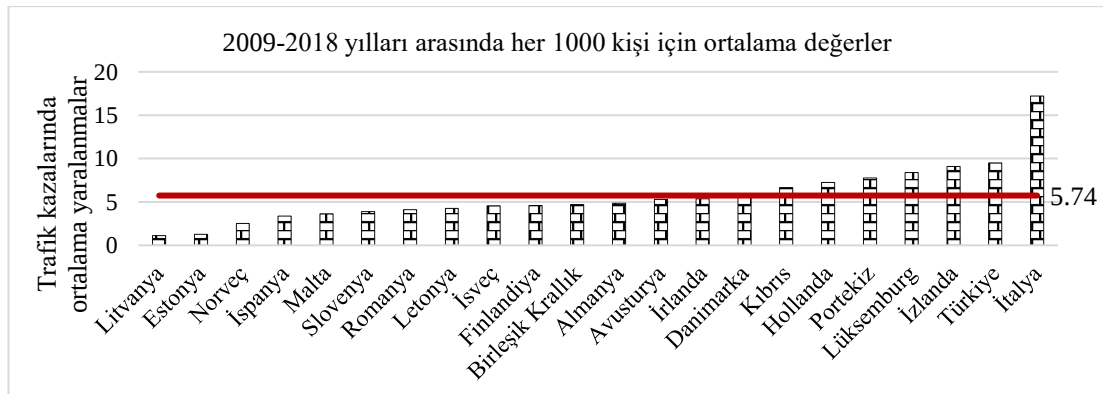
Bu tezin amacı, trafik kazalarına neden olan etkenlerin birlikteliklerini incelemek, sürücü kaza şiddetini belirlemek için sınıflandırma tahmin modeli kurmaktır. Böylelikle, meydana gelebilecek kazaları öngörebilmek, temel karakteristik yapılarını ortaya çıkarmak ve alınacak önlemlere katkı sunabilmek amaçlanmıştır. 2015-2020 yıllarında gerçekleşen ölümlü-yaralanmalı trafik kazalarının oluşmasında birlikte rolü olan etkenlerin belirlenmesi ve kaza şiddeti ile ilgili tahminler yapılması hedeflenmiştir.

Hedeflere dönük adımlar (iş akışı):

1. T.C. İçişleri Bakanlığı Emniyet Genel Müdürlüğü'nden alınan ölümlü-yaralanmalı trafik kazaları veri setine ait betimleyici istatistiklerin elde edilip, genel durumun incelenmesi,
2. Betimleyici istatistikleri elde ettikten sonra veri ön işleme sürecine geçip, veri temizleme, dönüştürme, birleştirme, gruplandırma işlemleri yapılarak verilerin analizler için uygun hale getirilmesi,
3. Tüm taşıtların yer aldığı veri setinde Birliktelik Kuralları oluşturularak, kazaya birlikte neden olan etkenlerin belirlenmesi, mevcut durumun saptanması,
4. Elde edilen bu örüntülerle ilgili durumların tartışılması, değerlendirilmesi
5. Veri setinde taşıt türleri içerisinde sıklığı (gerçekleşme sayısı) en fazla olan taşıt olan otomobillerin yer aldığı veri setinin oluşturulması,
6. Kazalara karışan otomobil sürücüleri ve kaza ile ilgili diğer özniteliklerden Karar Ağacı oluşturulması, bu şekilde ağaçtaki yaprağa ulaşana kadar gerçekleşen durumların belirlenmiş olması,
7. Yapılan iki analiz için de sonuçların yorumlanarak saptamaların yapılması ve önerilerin sunulmasıdır.

1.3. Dünya’da Trafik Kazaları

Dünya genelinde yaklaşık olarak 1,3 milyon insan trafik kazalarında hayatını kaybetmekte, 50 milyon kişi yaralanmaktadır. Trafik kazalarında hayatını kaybedenlerin yarısından fazlasını savunmasız yol kullanıcıları (yaya, bisiklet ve motosiklet kullanıcıları) oluşturmaktadır. Dünya genelinde bulunan taşıtların yaklaşık %60’ı düşük ve orta gelirli ülkelerde olmasına karşın, trafik kazalarında hayatını kaybedenlerin %93’ünü bu ülkelerdeki kazalar oluşturmaktadır. Ayrıca 5-29 yaş aralığında bulunan genç nüfusun ölüm nedenlerinin başında trafik kazaları gelmektedir (Dünya Sağlık Örgütü, 2022). Trafik kazaları ölümle sonuçlanmasa da bazen insanların yaralanmasına, ömür boyu kalıcı sakatlıklara yol açmaktadır. 2019’da yapılan bir çalışmada, trafik kazalarının, sakatlanmalara yol açan kazalar arasında 7. sırada olduğu belirtilmektedir (Vos vd., 2020). Avrupa ülkelerine ait trafik kaza bilgilerine göre; Avrupa Birliği’ne üye olan ülkelerde (AB-28) bin kişiye düşen otomobil sayısı 524, Türkiye’de 162 iken; Avrupa Birliği’ne üye olan ülkelerde (AB-28) bir milyon otomobile düşen ölü sayısı 93, Türkiye’de 391’dir (KGM_Trafik Kazaları Özeti, 2022). Avrupa Birliği’ndeki ülkelerde 2009-2018 yılları arasında yaralanmaların incelendiği raporda, Şekil 1.1’de görüldüğü üzere, 1000 kişide trafik kazalarında yaralanmaların ortalaması 5,74 olarak ifade edilirken, bu değer Türkiye’de yaklaşık olarak 9,5 olarak belirtilmiştir (Injuries in the European Union 2009-2018, 2021).



Şekil 1.1. 2009-2018 yılları arasında trafik kazalarında, 1000 kişide ortalama yaralanmalar (Injuries in the European Union 2009-2018, 2021).

Dünya Sağlık Örgütü (DSÖ) ve Birleşmiş Milletler Bölgesel Komisyonları tarafından geliştirilen, karayolu güvenliğine yönelik on yıllık küresel eylem planındaki hedeflerden biri, 2030 yılında yollardaki ölüm ve ciddi yaralanmaları en az %50

oranında azaltmaktır. Bu küresel eylem planını gerçekleştirmek için ülkelerin ve onların bu konudaki paydaşları olan tüm yönetimlerin “güvenli sistem yaklaşımı”nı benimsemeleri gerekir. Bu yaklaşım;

- İnsan hatalarını önceden tahmin edip uyum sağlama,
- Yollardaki ölüm veya ciddi yaralanmaları önlemek için çarpışma şiddetlerini azaltan yol ve taşıt tasarımlarını içerme,
- Yolların tasarımını ve bakımlarını yapan kurumlar, taşıt üreten firmalar ile yol kullanıcılarının kaza ile ilgili sorumluluklarının paylaşılması sağlanarak kaza sonrası sadece sürücü ya da diğer yol kullanıcılarının suçlanması yerine, kazaların önlenmesi için tüm sistemin birlikte düşünülmesi,
- Yolların ve taşıtların sürekli iyileştirilmesi ilkesi benimsenerek, maliyet düşürme ya da daha hızlı ulaşım gibi istekler uğruna “sıfır ölüm veya ciddi yaralanma” ilkesi göz ardı edilmeyerek, sadece kazaların en son gerçekleştiği yol kesimlerini iyileştirmek yerine tüm ulaşım ağının iyileştirilmesi odaklı çalışma, yöntem ve uygulamaları kapsamaktadır (Dünya Sağlık Örgütü, 2021).

Dünya’nın çeşitli bölgelerinde trafik kazalarının nedenlerini anlamak, kaza örüntülerini keşfetmek, kaza şiddetini tahmin etmek üzere çeşitli çalışmalar yapılmıştır.

Aşağıdaki bölüm bu çalışmalar hakkında detaylı bilgileri içermektedir.

1.3.1. Trafik güvenliği ve trafik kazaları ile ilgili çalışmalar

Bu bölümde, çeşitli analiz yöntemleri kullanılarak trafik güvenliği analizleri içerikli yapılan çalışmalara yer verilmiştir.

Augustynowicz, yaptığı çalışmada, sürücüleri en saldırgan (agresif) olandan en ılımlıya sınıflandırmaya çalışmıştır. Bu anlamda sürücülerin hız profilini (hızlanma, sabit hızla gitme, yavaşlama) incelemiştir. Çalışma sonuçlarına göre, sürüş tarzıyla ilgili yapılan çalışmalara dayalı olarak şimdiye kadar kurulmuş olan modellerin, çeşitli trafik koşullarında karmaşık insan davranışlarına bağlı olarak ortaya çıkması beklenen olası sürücü tepkilerini çözümleyerek tahmin edecek düzeyde yaygın ve tam olarak uygulanabilir olmadığı ortaya çıkmıştır (Augustynowicz, 2009).

Romanya-Bükreş’te yapılan bir çalışmada, şirket taşıt filosuna ait 23 araca GPS yerleştirilip hızlanma-frenleme, hız değerleri gibi verilerle sürüş davranışlarına (risk eğilimlerine) göre sürücüler “saldırgan” ve “sakin” taşıt sürücüleri olarak tanımlanmış

ve Hiyerarşik Kümleme ile Temel Bileşen analizleri kullanılarak sınıflandırılmıştır. Çevresel (hava durumu, yol durumu, gün ışığı, trafik koşulları), sürücü özel durumları (hasta, yorgun, uykusuz, dikkati dağılmış), sürücü kişisel özellikleri (cinsiyet, yaş, taşıt kullanım sıklığı, tecrübe, kişisel tutum) gibi birçok etkenin sürücü davranışını etkilediği vurgulanarak, süre ve sürücü durumlarına ait bilgi yetersizliği kısıtları bulunduğundan bu durumların incelenemediği; çalışmanın, kişisel özellikler de veri setine eklenerek geliştirilebileceği belirtilmiştir. Ayrıca, sürücülerin bilişsel durumlarının anlaşılmasının zor olabileceği ancak taşıtlara yerleştirilebilecek algılayıcılardan (sensörlerden) gelen video kayıtlarının bilgisayar görüntüleme analizleri ile gerçek çözümler üretilebileceği sonucuna varılmıştır (Constantinescu vd., 2010).

Amerika’da Ulusal Karayolu Departmanı’nın yaptığı çalışmada, 16 yaş ve üzeri 6000 kişiye telefonla anket yapılmış, sürüş tecrübesi, hızlanma davranışı ve güvensiz sürüş ile ilgili, agresif sürücülüğü belirlemeye yönelik sorular sorulmuştur. Katılımcıların %60’tan fazlası diğer sürücülerin güvensiz taşıt kullandıklarını, kendileri ve aileleri için tehdit oluşturduklarını beyan etmişlerdir. Trafikteki güvensiz sürüş ortamının ve kazaların oluşmasındaki ana nedenlerden birisi olarak, trafik denetimlerinin azalması trafik sıkışıklığının artmasıyla beraber insanların zaman konusunda baskı hissetmesi ve sürücülerin agresifleşmelerine bağlı olarak hızlarını arttırmaları olarak belirtilmiştir. 1980 yıllarında Amerika’da, alkolden dolayı sürücülere ağır cezalar verilmediğinden alkolü sürüşün çok yaygın bir trafik olgusu olarak ortaya çıkmasına bağlı olarak çok fazla motorlu taşıt kazası ölümle sonuçlanmıştır. Benzer şekilde günümüzde agresif sürüş ve aşırı hızdan dolayı çok fazla kaza ölümle sonuçlandığından cezaların caydırıcı olması gerekmektedir. Bu kurumun agresif sürücü tanımı; “Bireylerin diğer kişileri ve mülkleri tehlikeye atan sürüş davranışlarında bulunmasıdır” olarak ifade edilmiştir. Agresif sürüş ile ilişkili hareketler ise; hız sınırını aşma, takip aralığının çok kısa olması, kararsız ve güvensiz şekilde şerit değiştirme ile trafik işaret ve levhalarına (dur işareti, yol ver işareti, trafik ışıkları) uymama şeklinde belirtilmiştir. Bu kurumun çalışmalarına göre, birden fazla trafik kuralı ihlali, trafik kazası riskini arttırmaktadır (Aggressive driving enforcement, 2004).

Sürücü dikkat dağınıklığının da kazaları önemli ölçüde etkilediği bilinmektedir. Dikkat dağınıklığı nedenlerinden biri de cep telefonu kullanımıdır (Mobile Phone Use,

2023). Genç sürücülerde cep telefonu kullanımı daha yaygın olup direksiyon başında daha az tecrübeleri olduğu da göz önüne alındığında dikkatlerinin dağılma olasılığının yaşlı sürücülere göre daha yüksek olduğu ifade edilmektedir. Cep telefonu kullanımının kaza riskine etkisinin kesin olarak belirlenmesi zor olsa da yapılan araştırmalara göre, taşıt başında cep telefonu kullanan sürücülerin kazaya karışma riskinin diğerlerinden 4 kat fazla olduğu belirtilmektedir. Telefonla konuşurken taşıt sürmenin tehlikeleri, kapsamlı araştırmalarla ortaya çıkarılmıştır. Gerek gözlemsel gerekse epidemiyolojik çalışmalar telefonla konuşmanın sürüşte yarattığı eksikliğin, alkollü olmanın yarattığı eksikliğe yakın olduğunu göstermiştir (Clifasefi vd., 2006; Lipovac vd., 2017; Redelmeier & Tibshirani, 1997). Cep telefonu kullanımına ilişkin yasaların etkinliği konusunda bazı araştırmalar yapılmış olsa da yasaların cep telefonu kullanımını azaltıp azaltmadığına dair etkin sonuçların üretildiği çalışmalara olan ihtiyaç devam etmektedir. Ayrıca, elde telefonla konuşmayı yasaklayan yasaların eller serbest olarak bluetooth aracılığı ile cep telefonu kullanımını arttırıp arttırmayacağına, bu durumun aynı oranda güvenlik riski oluşturup oluşturmayacağına da belirlenmesi gerekmektedir (Christopher Chabris & Daniel Simons, 2018; T.C.İçişleri Bakanlığı Emniyet Genel Müdürlüğü-Cep Telefonu Kullanımı, 2023).

Avustralya’da yürütülen bir çalışmada, ağır yaralanmayla sonuçlanan kazalarda, hastaneye kaldırılan sürücüler tarafından beyan edilen dikkat dağınıklığının, bahsedilen çalışmadaki veri setinde yer alan kaza nedenlerinin %14’üne denk geldiği belirlenmiştir (T.C.İçişleri Bakanlığı Emniyet Genel Müdürlüğü-Cep Telefonu Kullanımı, 2023).

Avustralya’da yapılan başka bir çalışmada, ehliyet sisteminin (aşamalı) kazaları azaltmaya etkisini incelemek için Poisson Regresyon analizi kullanılmış, tüm acemi sürücü kazaları (tek taşıtlı, gece, alkollü, ölümlü ve ciddi yaralanmalı) için ayrı modeller kurulmuştur. Tek taşıtlı kazalarda, tecrübeli sürücüler ile acemi sürücüler arasında fark bulunamamış, gece meydana gelen kazalarda ise yeni ehliyet alanların kazaya karışma eğiliminin, tecrübeli sürücülere göre %6 daha az olduğu belirlenmiştir (Senserrick vd., 2021).

Yaşlı sürücüler için yapılan bir çalışmada, demans problemi olan veya olmayan yaşlı sürücülerin sürüş performansları kapsamlı bir şekilde incelenmiştir. Bilişsel bozukluğu

olan bireylerin, olmayanlara göre daha çok birbirine benzer sürücü davranışları sergiledikleri ortaya çıkmıştır (Feng vd., 2021).

Tayvan-Taoyuan'da yapılan bir çalışmada, trafik ihlalleri ile kazalar arasındaki ilişki, derin öğrenme ile ortaya konulmaya çalışılmıştır. İncelenen trafik ihlalleri, kazalara neden olan beş temel ihlal olan; yol vermeye uymama, yasal olmayan dönüş, yasal olmayan şerit değiştirme-kullanma, hızlı sürme, trafik ışığı ihlali ve diğer nedenler olarak ayrılmıştır. Çalışma sonucunda Evrişimli Sinir Ağları ve Geçitli Tekrarlayan Birim yöntemlerinin dikkat çekici düzeyde bir tahmin doğruluğu sağladığı, diğer yaklaşımlardan (Bütünleşik Otoregresif Hareketli Ortalama, Rassal Orman ve Gradyan Yükseltmeli Karar Ağaçları) daha iyi performans gösterdiği sonucu elde edilmiştir. Trafik güvenliği ile ilgili yapılacak çalışmalarda, trafik kaza ve ihlal verilerine uygulanacak olan makine öğrenmesi ve derin öğrenme yaklaşımlarının uygun olduğu bu çalışma ile gösterilmiştir (Y.-W. Wu & Hsu, 2021).

Alavi ve ark., İran'da, 800 ağır taşıt sürücüsü ile yaptıkları çalışmada, yaş arttıkça kaza riskinin azaldığı, depresyonun kaza olasılığını 2,4, takıntı rahatsızlığının ise 2,7 kat arttırdığı, ayrıca kaza yapan ve yapmayan sürücüler olarak ayrılan grupların sürücü davranışı, kasıtlı ihlal, dönme-manevra hatası, kasıtsız ihlal gibi davranışları açısından anlamlı farklılık gösterdiği bulgularına ulaşmıştır (Alavi vd., 2017).

Destek Vektör Makineleri ve Hidden Markov Model algoritmalarının kullanıldığı bir diğer çalışmada, kavşakta sürücü davranışları, çeşitli istatistiksel modellerle analiz edilmiştir. Kavşakta durma davranışı için; sürücü özelliklerinden yaş ve tepki süresi, sarı ışık yanma anı ve geometrik parametreler (taşıt hızı ve kavşağa uzaklık) incelenmiştir. Çalışmada diğer çalışmalardan farklı olarak aynı model içinde birçok parametrenin birleştirilme esnekliğinin bulunduğu beyan edilmiştir. Ayrıca sürücülerin sınıflandırılmasının, daha gelişmiş sürücü asistanı sistemlerinin temellerini oluşturacağı beyan edilmiştir (Aoude vd., 2012).

Sürücü dikkat dağınıklığı; görsel (taşıt içindeki diğer yolcu ya da telefona bakma), görsel ve fiziksel (bir şey yazma, radyo vs. ilgilenme), bilişsel (telefon ya da taşıt içindeki yolcularla konuşma) dağınıklık olmak üzere üç kategoriye ayrılabilir. Taşıttaki kişi sayısının ihlal ve hataları azaltmada etkili olduğu araştırmacılar tarafından tartışılmış olup bazı araştırmacılar taşıttaki kişilerin ihlal ve hataları arttırdığını savunmuştur. Genç sürücülerin kullandığı taşıtlarda akranları olması

durumunda, kaza riskinin önemli derecede arttığı, taşıtta kardeşlerin olmasının kaza riskini minimuma indiren etki yaptığı, çocuk yolcuların ise dikkat dağınıklığı yarattığı ancak kaza riskini düşük oranda etkilediği görülmüştür (Arthur Goodwin vd., 2010).

Bu tez çalışmasında; taşıtta başka şahısların olmasının kazaya neden olabileceğinin göz ardı edilmemesi gerektiğine vurguda bulunmakla beraber, bu etkinin mevcut veri setlerinden net olarak belirlenememesinin dikkate alınarak taşıtta başka kişinin veya kişilerin olduğu kazalar veri setinden çıkarılmış ve analizler buna göre yapılmıştır.

Yol sınıfı ve yolculuk süresinin de sürücü davranışını etkilediği, yolculuk süresi uzun olduğunda sürücülerin kaybedilen zamanı telafi etmek için hızlı taşıt kullanma eğiliminde oldukları belirlenmiştir. Ayrıca sigorta şirketlerinin, gidilen km "kullandığın kadar öde" ve güvenli sürüşüne bağlı olarak (sürücü geçmişine (kaza, ceza, tecrübe vs.)) "sürüş tarzına göre öde" politikalarını benimsedikleri ifade edilmiştir (Singh & Kathuria, 2021).

Sürücü ön-eğitim programlarının, beş psiko-sosyal etken üzerinde 12 ay boyunca etkisinin incelendiği bir çalışmada telefonla ve internet üzerinden anketler yapılmış ve iyimser tutuma karşı yönelimin zamanla arttığı (1-4. ve 9-12. aylar arasında), diğer bir deyişle bu konularda bilincin zamanla oturduğu görülmüştür. Katılımcılar, başlangıç aylarına göre 9-12 ay sonunda, sürücülüğün doğasında bulunan riskleri daha derinlikli olarak algılamışlardır. Bu tip çalışmalar, yetkililerin ehliyet verme ile ilgili uygun süre ölçekli eğitim modellerini geliştirmeleri gerektiğini göstermektedir. Bu eğitim modellerinin topluma sağlayacağı sosyal ve maddi getirilerin de dikkate alınması ile, konunun öneminin anlaşılacağı ifade edilmelidir (Bates vd., 2020).

Yapılan bir diğer çalışmada, önceki çalışmalarda incelenen trafik akım koşulları, sürücü takip davranışları, sürücü dikkati, sürücü karakterleri, sürüş ortamı, sürücülerin çevredeki taşıtlarla etkileşimi gibi ayrık etkenlerin birleştirilmiş etkisini ortaya çıkarmak adına Yapısal Eşitlik modeli kurulmuştur (K.-F. (Ken) Wu & Wang, 2021). (K.-F. (Ken) Wu & Wang, 2021).

Yapılan bir derleme çalışmasında, sürücü davranışları; “normal”, “madde etkisi altında” (alkol-uyuşturucu vs.) ve “trafik güvenliği açısından riskli” olarak sınıflandırılmıştır. İlgili çalışmada sürücü davranışının, trafik kazalarında en büyük etken olduğu vurgulanmış, sürücülerin hızlarını düşürmelerinin, sürüş esnasında dikkat dağıtıcı her şeyden uzak durmalarının ve olumsuz hava koşullarında

dikkatlerini arttırmalarının önemli olduğu belirtilmiştir. Sürücü davranışının geliştirilebilmesinin, sürücü davranışlarıyla ilgili aktif ve süreklilik arz eden bir geri bildirim sürecinin işletilmesi ile başarılabacağı ifade edilerek, bu anlamda cep telefonunun doğal sürüş verilerini toplamada önemli bir aygıt olduğu bildirilmiştir (Singh & Kathuria, 2021).

44 farklı çalışmanın incelenip değerlendirilmesi ile yapılan bir derleme çalışmasında uykusuzluğun ve yorgunluğun güvenli sürüşe olan etkisi analiz edilmiş, son 24 saatte 5 saatin altında uyuyan sürücülerin, tıpkı alkol ve uyuşturucu etkisi altında olan bozulmuş bir zihin yapısındaki gibi bir sürüş zihnine sahip olduklarının kabul edilebileceği belirtilmiştir (Dawson vd., 2021).

Işıklı kavşaklarda yayaya çarpma kazalarının incelendiği bir çalışmada, kamera kayıtlarından elde edilen yaya ve taşıt hareketleri Yapay Sinir Ağları ile modellenip hata matrisi oluşturulmuştur. Bu çalışmada önerilen modelin en etkin kullanım alanı, çarpışma uyarı sistemlerinin geliştirildiği modeller olarak öne çıkmaktadır (S. Zhang vd., 2020).

Fizyolojik sinyaller kullanılarak sürücü uyku durumunu algılama yöntemlerinin duyarlılığı ve belirleyiciliği ile ilgili yapılan bir çalışmada, sürücüyü belli anlarda izleyen ve uyaran bir sistemin kritik anlarda kazaları azaltmaya yardımcı olabileceği ancak çoklu sinyal yaklaşımlarının daha yüksek duyarlılığa sahip olduğu bu sayede daha etkin olabileceği belirtilmiştir (Watling vd., 2021).

Avrupa Birliği tarafından desteklenen “Trafik Güvenliği Kültürü” projesi kapsamında Estonya, Yunanistan, Kosova, Rusya ve Türkiye’de 743 sürücünün katıldığı bir çalışma gerçekleştirilmiştir. Çalışmanın amacı, agresif sürücü davranışının, hata, ihlal ve pozitif sürücü davranışına etkisini ortaya çıkarmaktır. Sürücü öfke ölçeği ve sürücü davranışları anketleriyle sürücülerin trafikteki davranışı hakkında bilgi toplanmak istenmiş, anket için çevrimiçi anket programı kullanılmış, farklı diller için ayrı ayrı linkler oluşturulmuş, katılımcılara, alınan cevapların anonim olacağının garantisi verilerek güven duymaları sağlanmıştır. Çalışmada farklı ülkelerde, sürücü davranışlarının sürüş ortamından etkilenme durumu farklı şekillerde sonuçlanmıştır. Örneğin, Yunan vatandaşlarda cinsiyetin önemli etkilerinin olduğu, Kosovalı vatandaşlarda seyahat sıklığının önemli olduğu, trafikteki diğer sürücülerin agresif davranışlarının tüm ülkeler için, sürücülerin anormal ve pozitif sürücü davranışlarında

etkili olduđu belirlenmiřtir. Ayrıca dűřmanca lekte agresif davranan sűrűcűlerin agresif uyarılar yaparak daha ok hata yaptığı, agresif uyarı hareketlerinin ılımlılıktaki rolű, alıřmadaki 3 űlke dıřındaki űlkelerde, dűřmanca saldırganlık ve intikam ile iliřkili bulunmuřtur. Estonya, Yunanistan ve Tűrkiye’de uyarılara maruz kaldıklarında sűrűcűlerin diđer űlkelere gre daha ok řiddete eđilimli davrandıkları belirlenmiřtir. Tűrkiye zelinde insanların, kiřiler arası atıřmalara ve fke űretmeye daha yatkın oldukları belirlenmiřtir. Polis denetimleri ile agresif davranıřlardan ziyade sadece kural ihlalleri belirlenmeye alıřıldıđı iin, diđer sűrűcűlere agresif tepki gstermenin, Tűrk trafik normlarına uygun olduđu dűřűncesinin Tűrk sűrűcűler arasında yaygın olduđu sonucuna varılmıřtır. Yapılan alıřma sonucunda, kiřisel deđerlerin ve kűltűrel seviyelerin trafik gűvenliđi analizlerinde dikkatli bir řekilde gz nűne alınması gerektiđi, her űlkenin kendi trafik kűltűr normlarına sahip olduđu, bu yűzden anketi oluřturan maddelerin her űlke iin farklı tanımlanması gerektiđi ifade edilmiřtir (T. zkan vd., 2011).

1.3.1.1. Birliktelik Kuralları kullanılarak yapılan alıřmalar

Bu alıřmada analiz sűrelerinin en temel kavramlarından biri olan ve ieriđi tezin ilgili blűműnde detaylı olarak aıklanacak olan Birliktelik Kuralları ile ilgili literatűr alıřmaları ařađıda aıklanmıřtır.

Cai (2020), kaza tűrű ve yeri, mevsim, gűnűn saati, hava durumu, aydınlatma durumu, grűř mesafesi, bariyer konumları, yol kaplamasının tűrű ve kavřak tűrű znelikleri arasındaki Birliktelik Kurallarını ortaya ıkarmıřtır. Kurallarda, lűmlű kazaların "sűrűř gűvenliđini zayıflatan diđer davranıřlarla birlikte tařıt kullanıldıđında" meydana geldiđi ifade edilmiřtir. Ayrıca, "hareket halindeki tařıtlarla arpıřma", mevsimin yaz ve gűn saatlerinin gece olması durumunda daha fazla lűmlű sonulanmaktadır. alıřmanın bir diđer dikkat ekici sonucu ise havanın gűneřli, arazinin dűz, saatin akřam ve caddede aydınlatma olmadıđı durumlarda kazalardaki lűm oranının arttıđına iřaret edilmesidir (Cai, 2020).

Yu ve ark., Wisconsin’de (ABD) yaptıkları alıřmada, veri madenciliđi algoritmalarından olan apriori ile trafik kaza oluř tűrű ve řiddetine etki eden etmenler arasında Birliktelik Kuralları ıkarmıřlardır. Sonulara gre; 16-29 yař aralıđındaki erkek sűrűcűlerin blűnmemiř yollarda meydana gelecek kazalara karıřma eđiliminde oldukları belirlenmiř, lűmlű kazaların daha ok řehir dıřı, maddi hasarlı kazaların ise řehir merkezlerinde meydana geldiđi ortaya ıkarılmıřtır (Yu vd., 2019).

Das ve ark., 2014-2018 yılları arasındaki yaya ölümlerini, en yaygın ölümlü kaza senaryolarından oluşan 4 gruba uygulanan Birliktelik Kuralları aracılığıyla incelemiştir. Bu çalışmada, apriori kullanılarak ve bir performans ölçütü olarak lift değeri alınarak 4 alt grup için en iyi yirmi kural elde edilmiştir. Bazı temel öznitelik kategorileri; aydınlatma, düz giden taşıtlar, dönüş yapan taşıtlar, caddeler ve yaya yaş aralığı olarak ele alınmıştır. Kural kalıpları, yayanın yaya geçidinin içinde ve dışında olmasına göre farklılık göstermiştir. Yaya, yaya geçidi dışında bir yerde olduğunda, gece aydınlatmanın olmaması birçok kaza ile ilişkili olarak Birliktelik Kuralları ile ortaya çıkmıştır (Das vd., 2021).

Kong ve ark., doğal sürüş veri seti ile hızlı sürüş davranışı ve yolculuk/sürüş/yol özellikleri arasındaki ilişkileri Birliktelik Kuralları algoritmaları kullanarak araştırmıştır. Çalışmanın bulgularından biri, uzun yolculuklarda ortaya çıkan yorgunluk ve dikkat dağınıklığının yüksek standartlı yollarda hızlı sürüşle beraber gerçekleşmesi durumunda şiddetli kazalara neden olabileceğidir (Kong vd., 2020).

Zhu, bisiklete çarpıp kaçma türündeki kazaların gerçekleşme olasılığını arttıran etkenleri belirlemek için Birliktelik Kurallarından yararlanmıştır. Elde edilen kurallarda, bisiklete çarpıp kaçma kazalarının sabah ve zirve saatlerde meydana gelme olasılığının düşük olduğu bulunmuştur (Zhu, 2020).

Hong ve ark., yük taşıtlarının karıştığı kaza verilerinden apriori algoritması kullanarak Birliktelik Kuralları elde etmiştir. Ortaya çıkan kurallarda, 20'li ve 30'lu yaşlardaki genç yük taşıtı sürücülerinin karıştığı kazaların, düşey ve yatay dönemeç olmayan, iyi tasarlanmış yol geometrisine sahip bölgelerle bağlantılı olduğu bulunmuştur. Ayrıca, çalışmanın dikkat çektiği diğer nokta; iyi yol koşulları ve sürücülerin trafik kurallarına uymaması arasındaki bağlantıdır (Hong vd., 2020).

Das ve ark., yağışlı havalarda meydana gelen kazaları açıklayan önemli Birliktelik Kuralları elde etmiştir. Dikkat çekici bulgulardan biri, 15 ile 44 yaş grubundaki sürücülerin çoğunlukla, aydınlatmanın zayıf, yolun yatay dönemeçli olduğu, yağmurlu havalarda kazalara karıştığıdır (Das vd., 2020).

Xu ve ark. tarafından yapılan çalışmada, 2009-2013 yılları arasında, Çin'de gerçekleşen şiddetli yaralanmalı trafik kazalarında karşılıklı bağıllığı olan etkenler Birliktelik Kuralları ile ortaya çıkarılmıştır. Sonuçlar, şiddetli yaralanmalı kazaların yol kullanıcısı davranışı, taşıt, yol geometrik özellikleri ve çevresel etkenler arasındaki

karmaşık etkileşimlerinin bir sonucu olduğunu göstermiştir. Taşıt ile ilgili durumlardan olan; lastik patlaması, fren tutmaması ve direksiyon arızaları gibi etkenler, görece yüksek standartlı yollarda bile şiddetli kazalara neden olabilmektedir. Edinilen kurallara göre, düşük standartlı ve yatay dönemeç bulunan yol kesimlerinde meydana gelen şiddetli kazaların tek taşıtlı kaza olma olasılığı yüksektir. Ayrıca, kötü hava koşullarında meydana gelen kazaların hızlı sürüş ve diğer sürücü hataları ile ilişkili olduğu bulunmuştur (Xu vd., 2018).

1.3.1.2. Karar Ağaçları kullanılarak yapılan sınıflandırma çalışmaları

Kategorik veri içeren durumlarda veri analizlerinde kullanılan Karar Ağaçlarının, trafik kazaları ile ilgili olanlarından bazılarına ait literatür çalışmaları ve bulguları aşağıda sunulmuştur.

Shahverdy ve ark., sürüş asistan sistemlerinden ve K-ortalamlar Kümeleme tekniğinden faydalanarak sürücüler; normal-güvenli, sinirli, dikkati dağınık, uykulu ve alkollü sürücü olarak beş sınıfa ayırmıştır. Daha sonra Yapay Sinir Ağları algoritması uygulanarak dikkat dağınıklığı, uyuklama ve direksiyon tutma davranışı tahmininde başarılı olunabileceği gösterilmiştir (Shahverdy vd., 2020).

Çin’de 2006-2013 yılları arasındaki kazalarını incelemek için, veri madenciliği yöntemlerinden biri olan Karar Ağaçlarından C4.5 algoritması kullanılarak model kurulmuştur. Her değişkenin modelin sınıflandırma etkisine katkısının derecesine göre, kazanın türünü etkileyen çeşitli modeller elde edilmiş ve model tahmin doğruluğu yaklaşık %80 olarak bulunmuştur. Değişkenler; mevsim, haftanın günü, günün saati, taşıt tipi, yol kesimi (köprü, tünel,), dönemeç yarıçapı ve kaza nedenidir. Çalışma sonucunda, bu modelin, eğitim ve test verilerinin iyi bir genel tahmini sağladığı, bu yüzden de C4.5 Karar Ağacının kaza oluş türü dağılımını analiz etmek için uygun bir yöntem olduğu görülmüştür. 2019’da yapılan bu çalışmada Çin'deki bilgi toplama sisteminin hala nispeten geri düzeyde olduğu ve gelecekte daha kapsamlı ve doğru bilgiye ihtiyaç duyulduğu belirtilmiştir (J. Li vd., 2019).

ABD eyaletlerinden Kuzey Dakota’da yapılan bir doktora tezinde, yüksek ölüm oranlarına sahip olduğu bilinen iki kaza tipi (karayolu-demiryolu hemzemin geçit kazaları ve ticari tır kazaları) araştırılmıştır. Bu çalışmada, Karar Ağaçları, Gradyan Yükseltme (gradient boosting) ve Yapay Sinir Ağları modelleri kullanılmıştır. Karar Ağaçları modellerinin; doğrusal olmayan ilişkilerin belirlenmesi, görsel olarak

açıklanabilir olması ve ağaç tabanlı modeller (bir veya daha fazla Karar Ağacından tahminler oluşturmak için bir dizi eğer-ise kuralı kullanılarak oluşturulan modeller) üretmesi dolayısı ile trafik kaza analiz çalışmalarında en çok tercih edilen modellerden biri olduğu ifade edilmiştir. Basit Karar Ağaçları modeli kurulduktan sonra Gradyan Yükseltme modeli geliştirilmiştir. Ayrıca belli bir varsayıma dayanma ihtiyacı olmaması açısından Yapay Sinir Ağlarının, genelleştirilmiş doğrusal modellemeden daha cazip olduğu belirtilmiştir (Zijian Zheng, 2018).

ABD eyaletlerinden Colorado'da yapılan bir çalışmada, yayaya çarpma kaza şiddetine etki eden etmenler 2006-2016 yılları arasındaki verilerle incelenmiştir. Değişkenler; kazanın kavşağa yakınlığı, gün ışığı durumu, taşıt tipi ve taşıt hızı olarak alınmış, olasılık oranları hesaplanmış ve Lojistik Regresyon modeli kurulmuştur. Yol kullanıcısı özelliklerinden yaya yaşı, yaya ve sürücünün bilişsel bozukluğu olup olmadığı (alkol-uyuşturucu kullanmış) dikkate alınmıştır. Çalışmada, kazalarda yayaların ölümü ya da yaralanmasıyla sonuçlanma oranları; konuma, hıza, sürücü yaşına göre ayrı ayrı analiz edilmiştir (Batouli vd., 2020).

Şerit değiştirme manevralarının yaralanma şiddetine etkisinin incelendiği bir çalışmada, bu manevraların birçok sürücü için (özellikle genç, tecrübesiz ve yaşlı sürücüler) zor olduğu belirtilmiştir. Bu yapıda gerçekleşen kaza çıktıları ile kazaya neden olan önemli etkenlerin ilişkisini anlamak için bir Lojistik Regresyon (Lojit) Model geliştirilmiştir. Modelde kazalar ağır yaralanmalı, hafif yaralanmalı ve yaralanma olmayan şeklinde sınıflandırılmıştır. Model çıktıları genç erkek sürücülerin, şiddetli yaralanmayla sonuçlanan kazalara karışma olasılığının yaşlı erkek sürücülerden fazla olduğunu ortaya koymaktadır (Adanu, 2021).

Zhang ve ark., trafik kurallarına uymamayı ve kaza şiddetini bağımlı değişkenler olarak alarak, bu durumlara neden olan etkenleri belirlemek için aşamalı regresyon modeli kurmuştur. Çalışmada, iki yıldan az deneyime sahip olan sürücülerin diğer sürücülere göre daha fazla trafik kurallarına uymama eğiliminde olmalarına rağmen şiddetli kazalara karışma risklerinin daha az olduğu belirtilmiştir. Çalışmanın bir diğer önemli bulgusu da sabah zirve saatlerinin kaza şiddeti üzerinde önemsiz bir etkisi varken bu saatlerde daha fazla trafik kurallarına uyulmamasıdır (G. Zhang vd., 2013).

Ma ve ark., farklı yaş gruplarındaki sürücülerde, yayaya çarpma türündeki kazalarda yayanın kaza şiddetini etkileyen unsurları araştırmıştır. Yaya yaşı, taşıt tipi ve hava

koşulları, yayaların karıştığı kazalarda kaza şiddetinde etkili olan önemli değişkenler olarak ortaya çıkmıştır. Bir diğer bulgu ise, bağımsız değişkenlerden olan taşıt sayısı ve trafik türünün genç ve orta yaşlı sürücülerin kavşaklardaki davranışlarını etkilediğidir. Ayrıca, genç ve yaşlı (65+) sürücülerin neden olduğu yayaya çarpıp kaçma kazaları, yolun bölünmüş olup olmaması gibi önemli bir değişkenle ilişkilidir. Geliştirilen Probit Regresyon modelleri ile sürücülerin eğitim seviyesinin, kavşak kontrol sistemleri ve tasarımları ile birlikte, kavşaklarda geçişe karar verme süreçlerinde yayaların güvenliğini artırmaya yönelik önemli etkiye sahip oldukları bulunmuştur (Ma vd., 2018).

Adanu, trafik akışı ve güvenliği üzerinde oldukça etkili olan sık ve uygunsuz şerit değiştirme davranışlarını araştırmıştır. Genç ve yaşlı sürücülerin karıştığı şerit değiştirmeye ilgili kazalardan kaynaklanan kaza şiddeti iki yaş grubu için iki Lojit Model kurularak analiz edilmiştir. Model tahminlerine göre, genç erkek sürücülerin şerit değiştirme kazalarında yaralanma olasılığı, yaşlı erkek sürücülere göre daha yüksektir. Bunun temel nedeni, yaşlı sürücülerin şerit değiştirirken daha dikkatli ve temkinli olmalarını kazaya karışma olasılıklarını azaltmalarıdır. Bir diğer bulgu ise eyaletler arası otoyollarda yaşlı sürücülerin şiddetli yaralanma ile karşılaşma olasılığının yüksek olmasıdır (Adanu, 2021).

Chiou ve ark., trafik kaza şiddetini tahmin etmek için iki taşıtın karıştığı kazalarda, her iki taşıtın da kazaya neden olan etkenlerini genelleştirilmiş tahmin denklemleri yöntemiyle incelemiştir. Analiz edilen trafik kazaları veri setinde, her iki taşıt içinde; sürücü özellikleri, taşıt türleri, yol aydınlatma koşulu ve çarpışma türü yer almaktadır. Çalışma sonuçlarına göre, kaza şiddetini en çok etkileyen etmenler; taşıt türü, hızlı sürüş davranışı, yandan çarpma türünde kazalar ve alkollü sürüş olmuştur (Chiou vd., 2020).

Alkheder ve ark., kaza şiddeti tahmini için veri madenciliği yöntemlerinden olan; Karar Ağaçları, Bayes Ağları ve Destek Vektör Makineleri (SVM) yöntemlerini kullanarak analizler yapmış ve bunlar arasında Karar Ağaçlarının, geleneksel istatistikî yöntemlere göre çok daha fazla üstün özelliklere sahip olduklarını ifade etmişlerdir. Bu yöntem, sürekli, ikili, kategorik, sıralı veri tiplerine uygulanabilmektedir. Analizlerde değişken olarak yaş, emniyet kemeri kullanımı, kazada zarar gören kişinin bulunduğu yer (sürücü koltuğu, arka koltuk, yaya), yol sınıfı, hız limiti, şerit sayısı, kaza oluş türü alınarak, belli ölçütlere göre yöntemlerin başarıları karşılaştırılmıştır.

Analiz bulgularına göre Bayes Ağları en iyi, Destek Vektör Makineleri en kötü tahmin sonucunu vermiştir. Çalışmada yayaların, sürücü ve yolculara kıyasla en savunmasız yol kullanıcısı oldukları ifade edilmiştir. Bir başka bulgu ise, erkek sürücülerin ve ön koltukta oturan yolcuların kaza sonucunda daha şiddetli yaralanma ve ölümlerle karşılaştığıdır (AlKheder vd., 2020).

Batouli ve ark., ABD'nin Colorado eyaletinde yaptıkları çalışmada, yayaya çarpma kazalarının şiddetini etkileyen etmenleri 2006-2016 yılları arasındaki veriler üzerinden incelemiştir. Lojistik Regresyon modeli aracılığıyla, kazanın kavşağa yakınlığı, gün ışığı durumu, taşıt tipi ve hızı değişkenleri göz önünde bulundurularak ölümlü ve ölümlü olmayan yaya kazalarının meydana gelme olasılıkları hesaplanmıştır. Çalışmanın sonuçlarına göre, yaşlı sürücülerin taşıt içi teknolojik aygıtları ve cep telefonlarını genç sürücülere kıyasla daha az kullanmaları, yaşlı sürücülerin dikkatlerinin daha az dağılmasını, tepki sürelerinin ve görüş yeteneklerinin artmasını ve dolayısıyla daha az ciddi kazalara karışmalarını sağlamaktadır. Araştırmanın bir diğer bulgusu ise, yaya şiddetinin hem yayanın hem de sürücünün kendi dikkat bozukluğundan etkilendiğidir (Batouli vd., 2020).

1.4. Türkiye’de Trafik Kazaları

2011 yılında hazırlanan “2011-2020 Karayolu Trafik Güvenliği Stratejisi ve Eylem Planı” ile ilk hedef olarak trafik kazalarındaki can kayıplarının %50 azaltılması planlanmıştır. 2017 yılında “Trafik Güvenliği Uygulama Politika Belgesi”nin yayınlanması ile, trafikte kısa, orta ve uzun vadede alınması gereken önlemler planlanmıştır. Tablo 1.1’de 2015’ten 2022’ye kadar (2022 dahil) trafiğe kayıtlı taşıt sayıları ve kaza sayıları verilmiştir.

Tablo 1.1. Trafiğe kayıtlı taşıt, kaza sayısı, 2015-2022 (Türkiye İstatistik Kurumu, 2023).

Yıl	Trafiğe kayıtlı taşıt sayısı	Toplam kaza sayısı	Ölümlü-yaralanmalı kaza sayısı	Maddi hasarlı kaza sayısı
2015	19 994 472	1 313 359	183 011	1 130 348
2016	21 090 424	1 182 491	185 128	997 363
2017	22 218 945	1 202 716	182 669	1 020 047
2018	22 865 921	1 229 364	186 532	1 042 832
2019	23 156 975	1 168 144	174 896	993 248
2020	24 144 857	983 808	150 275	833 533
2021	25 249 119	1 186 353	187 963	998 390
2022	26 482 847	1 232 957	197 261	1 035 696

Tablo 1.1’de trafiğe kayıtlı taşıt sayısının her yıl bir milyon civarında arttığı görülmektedir. Kaza sayısının ise COVID-19 salgınından dolayı 2020’de azaldığı, diğer yıllarda ortalama bir milyon iki yüz bin civarında olduğu görülmektedir.

Tablo 1.2’de aynı yıllarda gerçekleşen kazalardaki ölü, yaralı sayıları verilmiştir.

Tablo 1.2. Trafik kazalarına bağlı olarak ölen ve yaralananların sayısı, 2015-2022 (Türkiye İstatistik Kurumu, 2023).

Yıl	Ölü sayısı			Yaralı sayısı
	Toplam	Kaza yerinde	Kaza sonrası (1)	
2015	7 530	3 831	3 699	304 421
2016	7 300	3 493	3 807	303 812
2017	7 427	3 534	3 893	300 383
2018	6 675	3 368	3 307	307 071
2019	5 473	2 524	2 949	283 234
2020	4 866	2 197	2 669	226 266
2021	5 362	2 421	2 941	274 615
2022	5 229	2 282	2 947	288 696

(1) Trafik kazasında yaralanıp sağlık kuruluşuna sevk edilenlerden kazanın neden ve etkisiyle 30 gün içinde ölenleri kapsamaktadır.

2015 yılı itibariyle, trafik kazalarına karışan ve yaralanan kişilerin kazadan 30 gün sonrasına kadar takibi yapılmakta, bu kişilerin ilgili süre içinde hayatlarını kaybetmeleri söz konusu olursa, bu kazazedeler kaza sonrası ölü kayıtlarında yer almaktadır.

Tablo 1.3’te ölümlü yaralanmalı kazalardaki kusur dağılımı verilmiştir. Sürücü kusurlarının %85 üzerinde olduğu ve yaya kusurlarının 2020 yılında %9,5’e yükseldiği görülmektedir.

Tablo 1.3. Ölümlü yaralanmalı trafik kazasına neden olan kusur sayısı, 2015-2022 (Türkiye İstatistik Kurumu, 2023).

Yıl	Toplam kusur	Sürücü kusuru	Yolcu kusuru	Yaya kusuru	Yol kusuru	Taşıt kusuru
	Sayı	(%)	(%)	(%)	(%)	(%)
2015	210 498	89,3	0,4	8,8	0,9	0,6
2016	213 149	89,6	0,4	8,7	0,8	0,5
2017	213 325	89,9	0,4	8,5	0,7	0,5
2018	217 898	89,5	0,9	8,4	0,6	0,6
2019	204 538	88,0	1,3	8,2	0,5	2,0
2020	177 867	88,3	1,4	7,0	0,5	2,7
2021	224 418	87,1	1,8	8,2	0,4	2,6
2022	235 176	86,8	1,2	9,5	0,4	2,1

Tablo 1.3’te 2018 yılından sonra taşıt kusurlarının da yükseldiği görülmektedir. Bu tablodan, trafik denetimleri, polislere verilen eğitimlerin kaza tutanaklarının daha doğru doldurulmasına katkısı olduğu çıkarımı yapılabilir.

Tablo 1.4'te 2022 yılında ölümlü yaralanmalı kazalarda etkilenen kazazedelerin sayıları yer almaktadır.

Tablo 1.4. Ölümlü yaralanmalı trafik kazasından etkilenen kazazedeler, 2022 (Türkiye İstatistik Kurumu, 2023).

Kazazede	Ölü sayısı			Yaralı sayısı
	Kaza yerinde	Kaza sonrası	Toplam	
Sürücü	1084	1265	2349	141795
Yolcu	830	832	1662	109790
Yaya	368	850	1218	37111
Toplam	2282	2947		288696

1.4.1. Türkiye’de trafik güvenliği alanında yapılan çalışmalar

Türkiye’de 1997-2006 yıllarında şehir dışında gerçekleşen trafik kazalarının bölgesel ve zamansal değişimi 20 değişken (sosyoekonomik ve ulaştırma) kullanılarak, toplulaştırılmış verilerle (koordinat bilgileri olmadığı için şehir bazında) mekânsal ve zamansal kaza analizi yapılmıştır. Kaza sayısına ve şebeke uzunluğuna göre ölüm ve yaralanma oranları hesaplanmış, bu oranlar kullanılarak k-ortalamlar yöntemi ile kümeleme analizi yapılmıştır. Zamansal analizde ise, toplam trafik kaza sayıları kullanılarak zaman serileri yöntemiyle kaza tahmin modeli kurulmuştur (Atalay, 2010).

Erzurum’daki bölünmüş yollarda, 2005-2010 yıllarında gerçekleşen kaza verileriyle Yapay Sinir Ağları ve Genelleştirilmiş Lineer Regresyon yöntemleri kullanılarak trafik kaza tahmin modelleri kurulmuştur. Kazaya karışan ağır taşıt yüzdesi, yaz mevsiminde olan kazaların yüzdesi ve yatay dönemeç sayısı artışının kaza sayısını azalttığı; aksine Yıllık Ortalama Günlük Trafığın (YOGT) artışı, düşey dönemeç sayısı artışı ve yol kesim km’sinin artışının kaza sayısını arttırdığı bulunmuştur (Çodur, 2012).

Konumsal kaza verilerinin coğrafi bilgi sistemleri kullanılarak analizinin yapıldığı bir çalışmada, Kocaeli iline ait 2013 yılı yaralanmalı ve ölümlü kazalar incelenerek, kaza alan analizi ve kaza nokta belirlenmiştir. Kaza noktası belirleme yöntemleri; kaza sayısı, Kaza Tekrar Oranı, Tablo, Eşdeğer Ağırlık, Oran-Kalite Kontrol’dür. Çekirdek Yoğunluk Tahmin modeli kurularak istatistiksel analizlerle sonuçlara ulaşılmıştır. Bu anlamda noktasal kazaların yoğun yaşandığı yerlerdeki kazalar incelendiğinde aynı yerlerde benzer kazaların aynı sürücü hatalarına bağlı olarak gerçekleştiği belirlenmiştir. Örneğin, bir kavşakta gerçekleşen kazalar yoğunlukla iki otomobilin

kavşak noktasında dik açılı çarpışması şeklinde olmuş, ayrıca açık hava koşullarında yolun da kuru olması nedeniyle taşıt hızlarının artışına bağlı olarak kaza sayıları da artmıştır (Çelebi, 2014).

Alver ve ark., 2057 genç katılımcıya uyguladıkları yüz yüze anket verilerine bağlı olarak, sosyo-demografik özellikler ile ilgili katılımcıların belirlenen 4 adet ana sürücü kusuruna dayalı olarak trafik kazası yapmalarına neden olan davranışlarının arasındaki ilişkiyi incelemek için İkili Lojit (Binary Logistic Regression) yöntemini kullanmışlardır. Ankete göre, katılımcıların %23,9'unun son üç yıl içinde en az bir kazaya karıştığı, bunların %47,4'ünün emniyet kemeri kullanmama yüzünden trafik cezası aldığı belirlenmiştir. Çalışmadaki bu bulgu ile, emniyet kemeri kullanmama ile trafik kazasına karışma arasında bir ilişki olduğu ifade edilmiştir (Alver vd., 2014).

Türkiye'de yük taşımacılığının %86'sının karayolu ile yapıldığından yollardaki kamyon sayısının fazlalığı, kamyonların trafik kazalarında da büyük bir paya sahip olmasına neden olmaktadır. Kamyonların karıştığı kazalarda çarpışma şiddetinden dolayı kaza sonucundaki ölüm ve yaralanma sayısı fazla olduğundan bu kazaların incelenmesine karar verilmiştir. Ankara-Aksaray-Ereğli bölünmüş yollarında 2008-2011 yılları arasında gerçekleşen kazalar, YOGT, kamyon yüzdesi, şerit genişliği, ortalama kamyon hızı ve kavşak sayısı değişkenleri kullanılarak, Poisson Regresyon, Negatif Binom Regresyon, Sıfır Değer Ağırlıklı Negatif Binom Regresyon ve Yapay Sinir Ağları yöntemleri kullanılarak analiz edilmiştir. Çalışmada, trafik hacmi ve kamyon yüzdesinin, şeritlerin dar olmasının kazaları arttırdığı gibi sonuçlar elde edilmiştir (Funda TÜRE KİBAR, 2015).

Kadılar, sürücü profillerinin trafik kaza şiddeti üzerindeki etkilerini Sıralı Lojistik Regresyon (Ordinal Logistic Regression) modeli ile incelemiştir. Çalışmada, sürücülerin hayatını kaybettiği kazalar modellenmiştir. Buna göre, sürücünün yaşı ve eğitim düzeyi, alkol-uyuşturucu kullanımı, emniyet kemeri kullanımı, yol tipi ve yol koşulları, çarpışmanın yeri ve türü kaza şiddetini etkileyen önemli etkenler arasında yer almaktadır. Öte yandan, sürücü cinsiyeti, yol yüzeyi, hava durumu ve taşıt yaşının kaza şiddeti üzerinde istatistiksel olarak anlamlı bir etkiye sahip olmadıkları bulgularına ulaşılmıştır (Kadılar, 2016). Bu tez çalışmasında ise, otomobil yaşı gruplanarak öznitelik olarak veri setinde yer almış ve kaza şiddetine etkisi olduğu belirlenmiştir.

Yapılan bir yüksek lisans tezinde, tramvay kaza şiddetine etki eden etkenler Çok Terimli Lojit (Multi nomial logistic regression) modelle incelenmiştir. Çeşitli şehirlere ait veriler ve kaza literatürü incelendiğinde, kaza şiddetinin genellikle üç farklı sınıfta değerlendirildiği görülmüştür. Şiddetine göre ölümlü, yaralanmalı ve maddi hasarlı olarak yapılan bu sınıflandırma, bu çalışma kapsamında da benimsenmiştir. Kesikli tercih analizinde, bağımlı değişken sürekli değerler almamakta, iki veya daha fazla seçenek sınıfı arasından birine düşmektedir. Bu model sınıflandırma sayısının ikiden çok olduğu durumlarda kullanılır. Model sonuçlarına göre, tramvay hat genişliğinin normal veya yolun bölünmüş olması durumunda ölümlü ve yaralanmalı kazaların gerçekleşme olasılığı artmakta, maddi hasarlı kaza olma olasılığı ise azalmaktadır. Analiz sonuçları ayrıca gece meydana gelen kazaların ölümle sonuçlanma olasılığının arttığı, kazaya karışan kişi sayısı arttıkça da kazanın maddi hasarlı olma olasılığının azaldığı bulgularını ortaya koymaktadır (Akalin, 2016).

Ülkemizde yapılan bir diğer çalışmada, trafik kazalarını etkileyen etmenler Yapısal Eşitlik modeliyle incelenmiştir. Bu çalışmada üretilen modelde bağımlı değişken kaza boyutu, bağımsız değişkenler ise yer-konum, kaza türü, hava şartları, çevre, sürücü etkenleri olarak alınmıştır. Veri setini Ankara'da 2010-2015 yılları arasındaki trafik kazaları oluşturmaktadır. Kaza tutanaklarından elde edilen veri setinde uygulanacak güvenilirlik ve geçerlilik analizleri için Cronbach Alfa istatistiği ve Doğrulayıcı Faktör analizinden faydalanılmıştır. Çalışmada yapılan analizler sonucunda, kazanın yer ve konumunun, kaza oluş ve taşıt türünün, yol özelliklerinin, çevresel etkenlerin, hava şartlarının, sürücü yaşı, cinsiyeti ve eğitim durumunun kaza sonucunu (ölümlü, yaralanmalı, maddi hasarlı) etkileyen ana etmenler olduğu belirlenmiştir (Eygü, 2018).

Özkan ve ark. yaptığı çalışmada, sürücü beceri ve davranışlarına bağlı olan kaçınma davranışının trafik güvenliğindeki etkisi incelenmiştir. Bu çalışmada sürücülerin demografik özelliklerinin kaçınma davranışıyla ilişkilerinin belirlenmesi amaçlanmıştır. Çalışma sonucunda; yaş ve cinsiyetin trafikten kaçınma ile negatif, ölümlü/yaralanmalı kaza geçmişinin ise pozitif ilişkili olduğu, kadın sürücülerin trafikte daha fazla kaçınma davranışı gösterdiği belirlenmiştir (Ö. Özkan & Öz, 2019).

Yapılan bir diğer çalışmada, sürücü demografik özelliklerinin trafikte kaçınma davranışı ile ilişkisi üzerinde durulmuştur. Bu çalışmada, yaş ve cinsiyetin doğrudan sürücü davranışları, performansı ve kazaya yatkın olma durumu ile ilişkili olduğuna çeşitli çalışmalardan atıf yapılmıştır (Huang & Winston, 2011; Leonard Evans, 1991).

Trafikle ilgili demografik özelliklerin ehliyet sahibi olma süresi ve kaza sayısı temel değişkenlerinin trafik ortamında incelenmesinin, alınacak önlemler planlamalar açısından önemli bir etkiye sahip olduğunun açık olduğu belirtilmiştir (Ö. Özkan & Öz, 2019).

Kent içi trafik kazalarının Kümeleme ve Lojistik Regresyonla incelendiği İstanbul’da belirli ilçeler için 2015 yılı verileri ile yapılan bir çalışmada, ölümlü ve yaralanmalı kazalar, kaza türü, gün durumu, taşıt sayısı, yolun tipi ve geçit geometrisine göre sınıflandırılmıştır. Nüfusu büyük olan ilçelerde, kümelerin ayrışmasında en etkili olan etkenler yolun yüzeyi ve kaplaması, hava durumu ve ilçe adı iken, nüfusu küçük olan ilçelerde ise taşıt sayısı, kaza oluş türü ve yaralı sayısı öne çıkmaktadır (Karakuş, 2020).

Katanalp ve Eren (2020), Karar Ağacı tabanlı Hibrit-Bulanık modeller kullanarak 2013-2017 yılları arasında bisiklet kazalarından oluşan bir veri seti ile analizler yapmıştır. Bu kazalarda yaralanma şiddetini sınıflandırmak için C4.5 Karar Ağacı algoritması hem yaralanma-şiddet seviyesini sınıflandırmak hem de Karar Ağacı – Dönüştürülmüş Bulanık Mantık (DT-CFL) modeli kurmak için tüm bölünme (dallanma) koşullarını elde etmek için kullanılmıştır. Analizin başarısı, yaralanma-şiddet sınıflandırmasının doğruluk, kesinlik, geri çağırma ve F-ölçütü değerleri ile gösterilmiştir. (Katanalp & Eren, 2020).

1.5. Sakarya’da Trafik Kazaları

Tablo 1.5’te Sakarya’da 2015-2022 yıllarında gerçekleşen kaza sayıları verilmiştir. Tüm Türkiye’de olduğu gibi 2020 yılında kaza sayısının düşük olmasının nedeni COVID-19 salgınında sokağa çıkma yasakları ile trafikteki taşıt sayısının normal durumdan oldukça az olmasıdır.

Tablo 1.5. Sakarya’da 2015-2022 yıllarındaki taşıt ve kaza sayıları (Türkiye İstatistik Kurumu, 2023).

Yıl	Trafiğe kayıtlı taşıt sayısı	Toplam kaza sayısı	Ölümlü-yaralanmalı kaza sayısı
2015	245537	15798	3056
2016	262926	14649	3244
2017	279655	15736	3199
2018	285960	15686	3025
2019	287793	14742	2619
2020	299175	12696	2174
2021	312552	15758	2727
2022	327954	15990	2789

Tablo 1.6’da aynı yıllarda kazalarda ölen ve yaralananların sayıları verilmiştir.

Tablo 1.6. Sakarya’da 2015-2022 yıllarında kazalarda ölen ve yaralanan sayıları (Türkiye İstatistik Kurumu, 2023).

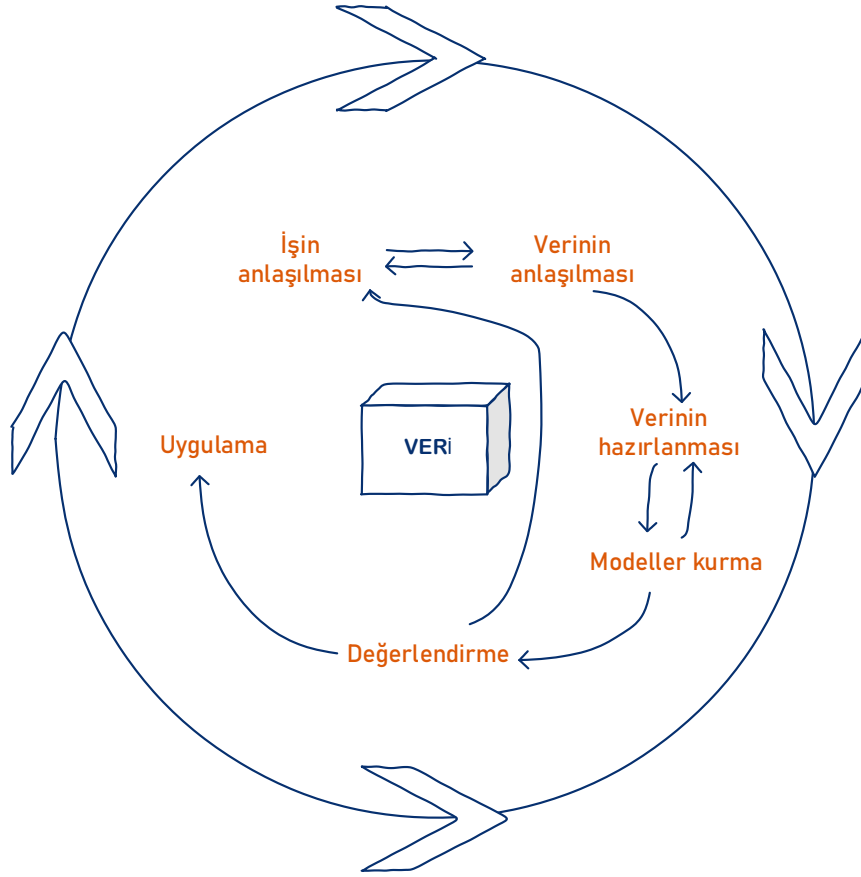
Yıl	Ölü sayısı			Yaralı sayısı
	Toplam	Kaza yerinde	Kaza sonrası ⁽¹⁾	
2015	108	69	39	5 232
2016	77	49	28	5 270
2017	64	40	24	5 091
2018	54	31	23	4 856
2019	54	29	25	4 110
2020	39	19	20	3 124
2021	59	33	26	3 972
2022	66	30	36	4 055

(1) Trafik kazasında yaralanıp sağlık kuruluşuna sevk edilenlerden kazanın neden ve etkisiyle 30 gün içinde ölenleri kapsamaktadır.

Tablo 1.6’da trafik kazalarında ölen ve yaralananların sayısının 2015’ten 2022’ye kadar geçen süreçte azaldığı görülmektedir. Bu duruma “trafikte sıfır can kaybı” hedefi için yapılan çalışmaların (denetim, yol geometrilerinde iyileştirmeler, trafik güvenliği için işaretlemelerde iyileştirmeler) etkisi olduğu ifade edilmelidir.

1.6. Çalışmanın Crisp-Dm Süreci ile Yönetimi

Veri madenciliğinin daha iyi ve hızlı sonuçlanması için, 200’den fazla veri madenciliği kullanıcısından oluşan endüstri liderleri tarafından geliştirilip ortaya çıkarılan Crisp-dm (Cross industry standard process for data mining - Endüstriler arası standart veri madenciliği süreci), (Eckerson vd., 2000) yaklaşımı analiz sürecinde uygulanmıştır. Bu yöntem Şekil 1.2’de görüldüğü gibi; işi anlama, veriyi anlama, veri hazırlama, modelleme, değerlendirme ve uygulama olmak üzere altı aşamadan oluşmakta, sözü geçen aşamaların çalışma boyunca tekrarlanması durumu bulunduğu bu sürecin bir döngü olduğu bilinmektedir.



Şekil 1.2. CRISP-DM sürecinin yaşam döngüsü, Eckerson ve ark. (2000)'dan uyarlanmıştır.

1.6.1. Amacın anlaşılması

Veri madenciliğinin kullanıldığı çalışmalarda en önemli aşama, yapılacak olan çalışma ya da işin amacının anlaşılmasıdır. Amacın yanlış tanımlanması verinin yanlış süreçlerden geçirilmesine neden olabilir. Bu yüzden işi ve amacı anlamak için çeşitli sorular sorularak ulaşılmaya çalışılan hedef belirlenmeye çalışılmalıdır. Hedef belirledikten sonra, çalışmada üzerinde durulan konu ya da sorunun veri madenciliği problemine dönüştürülmesi sağlanabilir. Bu çalışmada EGM'den alınan oldukça geniş kapsamlı ve içerikli ölümlü-yaralanmalı kaza veri seti analizleri ile, kazaların oluşmasında birlikte etkili olan öznitelikler belirlenmiş, sürücü kaza şiddeti sınıflandırma modeli geliştirilmiştir.

1.6.2. Verinin anlaşılması

Verinin içeriğini anlamak için titiz bir çalışma süreci gerekmektedir. Veri setini tanıma sürecinde verinin kalitesi, doğruluğu ve eksik kalan kısımları keşfedilir. Veri setinin analizi sonucunda ortaya çıkarılabilecek saklı bilgilerin kalitesi bu aşamanın özenli bir

şekilde geçilmesine bağlıdır. Veri temininden sonra, verinin anlaşılması aşamasında tüm özniteliklere ait, Bölüm 2.3'te verilen tanımlayıcı tablo ve grafikler elde edilmiştir.

Trafik kaza veri tiplerinin bazı özellikleri, onlar üzerinde istatistiksel yöntemlerin uygun bir şekilde analiz yapabilmesini zorlaştırmaktadır. Bu özelliklerden bazıları şunlardır:

- Kesin olmama durumu: Olay yerinde kaza tutanakları tutulurken, yazım yanlışlarının, eksik değerlerin ve verilerin sonradan işlenmesi işlemlerine insanlar dahil olduğundan, yapılan hatalar sonucunda gürültülü veri (hatalı veri) ortaya çıkabilir. Gürültülü verileri belirlemek ve düzeltmek için çeşitli algoritmalar mevcut olmasına rağmen, bu algoritmaların çoğu sürekli (continuous) sayısal verilerle ilgilenmektedir. Bu yüzden kategorik veri kümelerindeki hatalı değerleri belirlemek ve düzeltmek oldukça zordur. Bu çalışmada kullanılan veri setindeki tüm öznitelikler için geçerli olmak üzere, programlardaki (SPSS, MS Excel, Rapid Miner) filtre özelliği ile, yanlış olduğu düşünülen ve eksik olan veriler veri setinden çıkarılmıştır.
- Taraflı veri: Kaza tutanaklarına kaydedilen verilerin, tutanağı dolduran memurun öznel ya da hatalı yorumunu da içerebileceği göz ardı edilmemelidir. Polis memurlarının kaza araştırmaları konusundaki bilgi birikimlerine bağlı olarak verinin kalitesi değişebilmektedir. Bu durumların üstesinden gelebilmek için veri setinde yer alan özniteliklerde kendi aralarında doğrulama yapılarak yanlış olma ihtimali olan veriler temizlenmeye çalışılmıştır. Bazı özniteliklerin içindeki farklı sınıflar çok yakın anlamlara geldiğinden bu sınıflar tek bir sınıf altında toplanarak aynı durumun farklı ifade edilmesi sonucunda farklı veri olarak analize girme sorununun önüne geçilmesi sağlanmıştır.
- Seyrek gerçekleşen olaylar: Ölümlü trafik kazaları, yaralanmalı ve maddi hasarlı kazalara göre daha nadir, seyrek gerçekleşen olaylardır. Kaza tutanaklarında yer ve zaman bilgisi olsa da nadiren gerçekleştiklerinden, kazanın gerçekleştiği yerdeki verileri belirli bir dağılımla modellemek ve orada gerçekleşebilecek kazaları öngörmek çok zordur. Bu çalışmada, bir bölge ya da kaza kara noktaları yerine 2015-2021 yıllarında toplam 7 yılı kapsayacak şekilde, Sakarya'da polis sorumluluğunda bulunan tüm yollarda meydana

gelen kazalarla ilgili veri seti kullanılmış olup bu anlamda analizde kullanılan veri sayısı, modelleme için yeterli olmuştur.

- Verinin farklı değerlendirilebilir yapısı: Trafik kaza veri seti tekdüze bir yapıya sahip değildir. Kazaya etki eden özniteliklerin kaza oluşumunda birlikte yer alması ya da özel bir kaza nedeninin tüm veri seti için önemli olmaması durumu veri setinde düzenlemeler yapmayı gerektirebilir. Farklı özniteliklerden farklı veri setleri oluşturulması sonucunda ise çeşitli sonuçlar elde edilebilmektedir. Bu çalışma kapsamında, taraflı olmamasına dikkat edilerek yüzlerce veri seti oluşturularak analizler yapılmıştır. Literatürdeki çalışmalardan edinilen bilgiler doğrultusunda, benzeri özellikte ya da içlerinden birkaçının bileşimi olacak şekilde (verideki özniteliklerin seçimi, içlerindeki sürekli verilerin sınıflandırılma şekli) veri setleri üretilmiştir.
- Veri aralığının çok geniş, aşırı yayılmış olması: Birlikte Kuralları ve CHAID Karar Ağacı kurulmasında karşılaşılan bir zorluktur. Bu çalışmadaki veri setinde de bazı öznitelikler içerisinde yer alan verilerin bu tipte olduğu görülmüştür. Bu sorunu gidermek için sayısal veriler, belirli sınıflar altında gruplandırılıp kategorik veri tipine dönüştürülmüştür.

1.6.3. Verinin hazırlanması

Bu aşamada, veri setinin son halini oluşturmak için yapılan tüm işlemler yer almaktadır. Veri seçimi, verinin temizlenmesi, veri setindeki özniteliklerin birleştirilmesi veya dönüştürülmesi, istenen veri formatının oluşturulması işlemleri, bu süreçte yapılan işlemlerdir. Bölüm 2.4'te veri hazırlama aşaması olan veri ön işleme süreçleri açıklanmıştır.

1.6.4. Modelleme

Bu aşamada, çalışma için uygun olacak veri madenciliği yöntemleri seçimi, seçilen yöntemler için test uygulamaları, modellerin oluşturulması işlemleri yer alır. Verinin ön işleme süreçlerinden sonra birçok veri seti oluşturularak çok fazla model kurulmuştur. Bu modelleri kurarken literatürdeki çalışmalar örnek alınmıştır. Bu çalışmada, tez amacına uygun olarak çalışmaya özgü veri setleri elde edilmiş olup Bölüm 3'te açıklandığı şekilde modelleme aşamasında birliktelik analizleri yapıp, Karar Ağacı modelleri kurulmuştur.

1.6.5. Değerlendirme

Bu aşamada, kurulan modellerin çalışma amacına uygunluğu değerlendirilir. Bu aşamanın sonunda veri madenciliği çıktılarının nasıl kullanılacağına karar verilmektedir. Çalışmanın bitirilip uygulamaya geçilmesi burada alınacak karara bağlıdır. Çıktılar amaca ulaştıracak yeterlilikteyse uygulamaya geçilir, değilse daha fazla yineleme yapılması ya da yeni modeller kurulması uygun olacaktır. Bu çalışmanın 4 ve 5. bölümlerinde model sonuçlarının yorumlanması ve değerlendirilmesi yapılarak, trafik kazalarını azaltmak için üzerinde durulması gereken etkenler ve kurallar ortaya çıkarılmıştır.

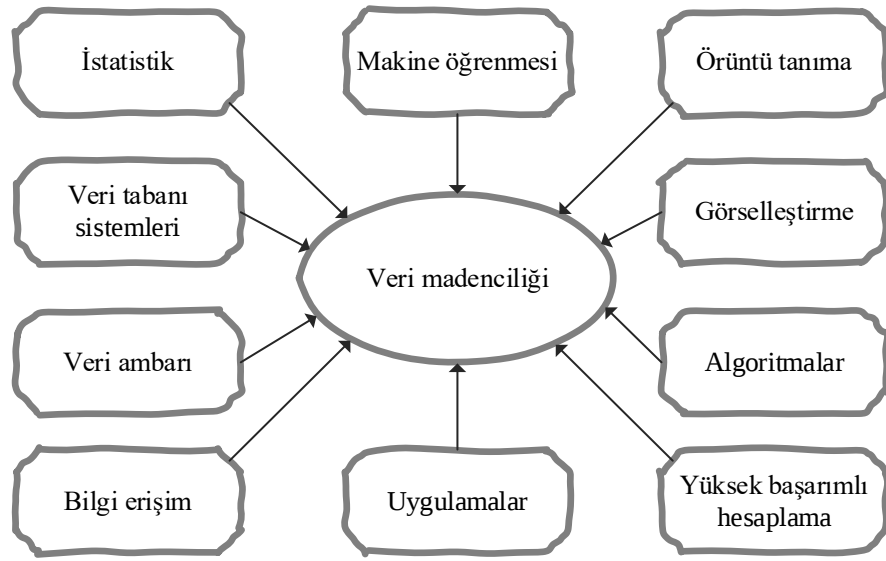
1.6.6. Uygulama

Bu kısım CRISP-DM sürecinde çalışmanın sona erdiği bölümdür. Bu bölümde şimdiye kadar gerçekleşen tüm aşamaların sonucunda uygulanan ve değerlendirme aşamasından geçen modellerin sonuçları raporlanarak, amaçlanan bilgi kazanımı elde edilmiş olur (Eckerson vd., 2000).

2. YÖNTEM VE VERİ

Literatürde, dünyada ve ülkemizde trafik kazaları ile ilgili çok fazla çalışma yapılmış, çeşitli yöntemlerle veriler toplanmış, trafik kazaları analiz edilmiştir. Bu çalışmada, kullanımı giderek yaygınlaşan veri madenciliği yöntemleri kullanılmıştır.

Uygulama alanı oldukça geniş olan veri madenciliği; istatistik, makine öğrenmesi, örüntü tanıma, veri tabanı ve veri depolama sistemleri, bilgi erişim, görselleştirme, algoritmalar, yüksek başarımlı hesaplama ve birçok uygulama alanından farklı teknikleri içermektedir (Şekil 2.1). Veri madenciliğinin disiplinler arası doğası, araştırma ve geliştirmenin başarısına önemli ölçüde katkıda bulunmaktadır (Han vd., 2012).

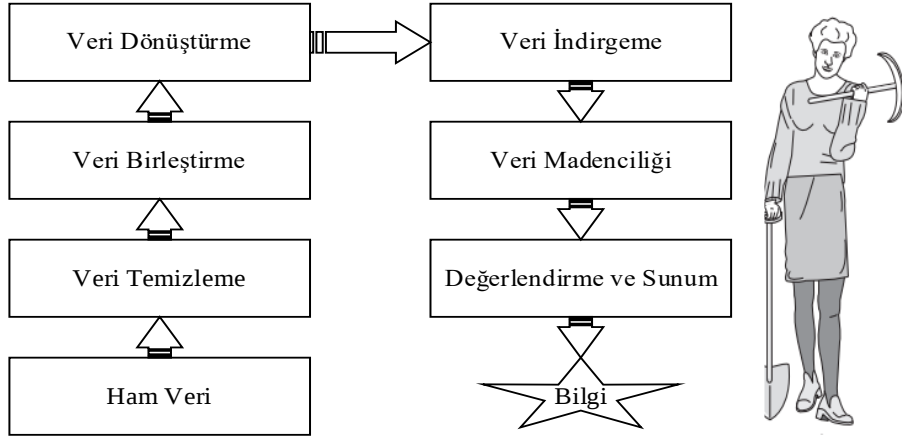


Şekil 2.1. Veri madenciliğini besleyen alanlar, Han ve ark. (2012)'den uyarlanmıştır.

2.1. Veri Madenciliği

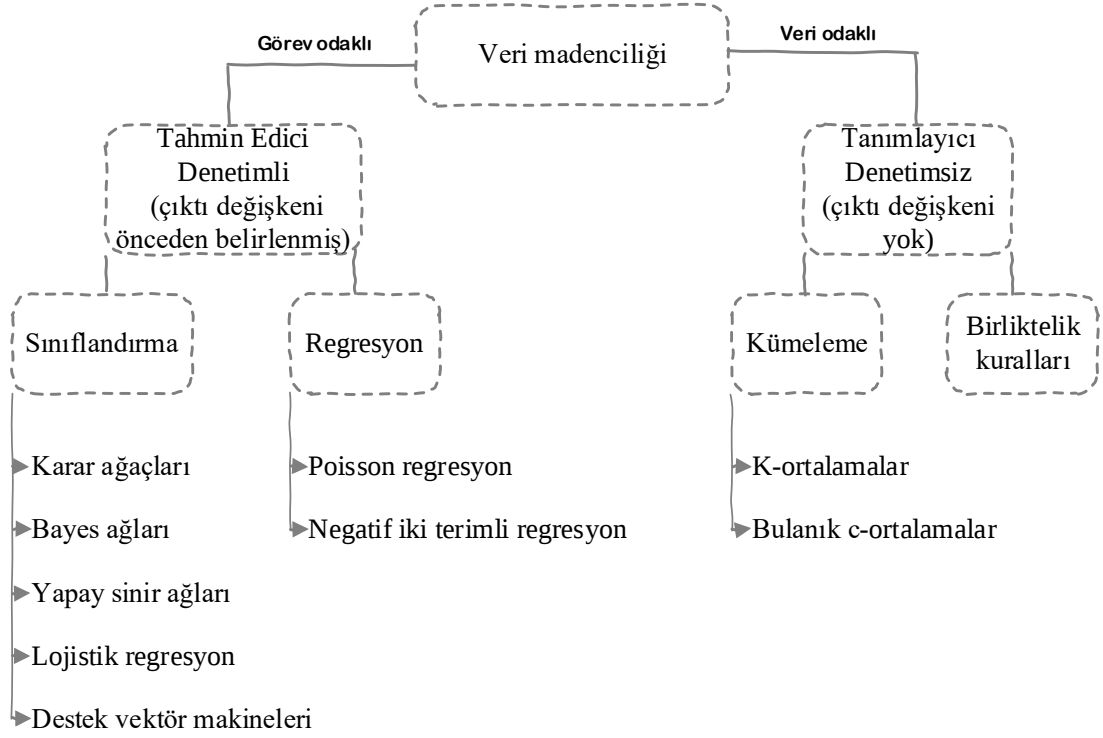
Veri madenciliği, büyük veri kümelerindeki gizli, ilgi çekici ve değerli örüntülerin keşfedilmesidir. İstatistiksel yöntemlerde olduğu gibi veri madenciliğinde de etkili, anlamlı sonuçlar üretebilmek için verinin yapısı, niteliği önemlidir. Ayrıca, veri madenciliğinin yaygın kullanımı sayesinde, veri toplama süreçlerinin daha anlamlı, etkin ve nitelikli içerikte olması söz konusu olabilmektedir (Hand, 2007). Nitelikli bilgiler üretebilmek için veri setinin analizlere uygun, temiz, anlamsız ve gereksiz

bilgilerden arındırılmış olması gerekmektedir. Ayrıca analiz süreçlerinde kullanılmak üzere elde edilmiş veri setlerinde öz niteliklerin dönüştürülmüş ya da birleştirilmiş hale getirilmesi çok önemlidir. Bunun için uygulanan veri madenciliği aşamaları Şekil 2.2’de verilmiştir.



Şekil 2.2. Veri madenciliği aşamaları, Han ve ark. (2012) ve Yu ve ark. (2019)’dan uyarlanmıştır.

Veri madenciliği yöntemlerini başlangıçta denetimli ve denetimsiz olarak Şekil 2.3’te de görüldüğü gibi iki gruba ayırmak gerekir. Denetimli veri madenciliği yöntemlerinde çıktı olan bir bağımlı değişkeni tahmin etme durumu vardır. Denetimsiz veri madenciliği yöntemlerinde ise çıktı değişkeni olmamakla beraber burada amaç durum belirlemesi, durum tanımlama ya da örüntü bulma şeklinde gerçekleşmektedir.



Şekil 2.3. Veri madenciliği grupları, Fian Tabea (2021) ve Maimon ve Rokach (2005)'ten uyarlanmıştır (Fian Tabea, 2021; Maimon & Rokach, 2005).

Denetimli makine öğrenmesi yöntemlerinde, belirli olan çıktı değişkenini, girdi değişkenlerinden öğrenerek tahmin etme söz konusudur. Amaç, çıktı değişkeninin bilinmediği veriler için, bilinen girdi (bağımsız) ve çıktı (bağımlı) değişkenleri arasındaki ilişkiden bilinmeyen çıktı verilerini belirlemektir. Sınıflandırma ve regresyon bu grubun içinde yer alır (Stanford Machine Learning, 2023). Bu tez çalışmasında, çıktı değişkeni sürücü kaza sonucu (yaralı, sağlam, ölü) alınarak yalnızca otomobil kazalarında, sınıflandırma yöntemlerinden biri olan Karar Ağaçları ile tahmin modeli kurulmuştur.

Denetimsiz makine öğrenmesi grubu, sınıflandırılacak ya da tahmin edilecek bir çıktı (bağımlı) değişkeninin bulunmadığı, veri kümesindeki değişkenler arasındaki örüntüleri bulmayı amaçlayan yöntemlerin bulunduğu gruptur. Kümeleme ve Birliktelik Kuralları yöntemleri bu gruptadır (Stanford Machine Learning, 2023). Bu tez çalışmasında, Birliktelik Kuralları kullanılarak trafik kazalarının gerçekleşmesinde ortak etkiye sahip öğeler ortaya çıkarılmıştır.

2.1.1. Birliktelik Kuralları

Birliktelik Kuralları, veri madenciliğinin tanımlayıcı modellerinden biri olup veri setinde bulunan veriler arasındaki ilişkileri tanımlamaktadır.

Birliktelik Kuralları, 1993'te Agrawal ve ark. tarafından geliştirilen bir yöntem olup yöntemin çıkış amacı pazarlama sektöründe, birlikte en fazla satılan ürünleri keşfetmek olmuştur. Bu yöntemin yıllar içerisinde sağlık, mühendislik gibi alanlarda da kullanımı oldukça artmıştır (Agrawal vd., 1993). Agrawal ve ark. (1994), apriori algoritması geliştirmiş olup bu algoritma, büyük veri kümelerindeki gizli ve önemli bağlantıları keşfetmek için eldeki tüm verileri tarayıp, başa dönüp tekrar verilen Destek (Support), Güven (Confidence) eşik değerini değerlendirerek, yeni bağlantılar bulunamayana kadar devam eden bir süreci kapsar (Agrawal vd., 1994).

Birliktelik Kuralları, büyük verinin, araştırma için hipotez çıkarmayı zorlaştırdığı durumlarda, veri setindeki öznitelikler arasında örüntüler kurarak anlamlı bilgiler üretmek için kullanılan bir yöntemdir (Srikant & Agrawal, 1995).

Bu algoritma, öncelikle veri kümesi içerisinde sıklıkla geçen öğeleri, Destek eşik değerine göre seçerek, birlikte sık geçen öğeleri bulmakta ve birlikte sık geçmeyen öğeleri elemektedir. Bu algoritmada kullanılan Destek ve Güven değerleri, aranan bağlantıların veri kümesi içerisinde ne kadar yaygın olduğunu belirleyen değerlerdir. Destek değeri, bir ya da birden fazla ögenin, bütün veri seti içerisindeki payı, güven değeri ise en az iki nesnenin birlikte olduğu durumların, ilk nesnenin olduğu durumların içerisindeki payıdır. Bu iki değer dışında bir de Lift (kaldıraç) ölçütü vardır. Bu ölçüt, güven değerleri yaklaşık olarak aynı olan iki Birliktelik Kuralının, hangisinin daha anlamlı olduğunu (bütün içerisinde) anlamaya yardımcı olmaktadır (Agrawal vd., 1994). Bu anlamda kullanılan Lift değeri Destek ve Güven eşik değerlerine ek olarak geliştirilen bir ölçüt olarak, ilgili veri setinde Destek ve Güven değerleri yakın olan kuralların hangilerinin daha büyük öneme sahip olduklarının belirlenmesi amacıyla kullanılmaktadır. Bir diğer tanımlama ile Lift, bir kuralda geçen veri gruplarının bir bütün (veri seti) içindeki büyüklüğünün ifadesidir (Han vd., 2012; Kotu & Deshpande, 2014; Pande & Abdel-Aty, 2009). Bu değer 1'den uzaklaştıkça elde edilen kuralın verdiği bilginin değeri artmaktadır.

Destek, Güven ve Lift değerlerini açıklamak için Tablo 2.1'de bu çalışmada kullanılan kaza veri setinin küçük bir bölümü verilmiştir.

Tablo 2.1. Trafik kaza veri setinin bir bölümü.

Kaza ID	Öznitelik-1	Öznitelik-2	Öznitelik-3
	Sürücü kaza sonucu	Sürücü kusuru	Kaza oluş türü
1	Sağlam	Yakın takip	Arkadan çarpma
2	Sağlam	Yol vermeme	Yandan çarpma
3	Yaralı	Yol vermeme	Yandan çarpma
4	Yaralı	Yol vermeme	Yandan çarpma
5	Sağlam	Hatalı dönüş	Yandan çarpma
6	Sağlam	Hızlı sürüş	Yandan çarpma
7	Yaralı	Yol vermeme	Arkadan çarpma
8	Yaralı	Hızlı sürüş	Yayaya çarpma

Tablo 2.1'deki öznitelik ve ögeler kullanılarak yapılan Destek, Güven ve Lift hesaplamaları aşağıdaki denklem 2.1, denklem 2.2, denklem 2.3, denklem 2.4'teki gibidir.

$$\text{Destek (Sağlam)} = \frac{\text{Sağlam ögesinin sıklığı}}{\text{Toplam veri sayısı}} = \frac{4}{8} = 0,50 \quad (2.1)$$

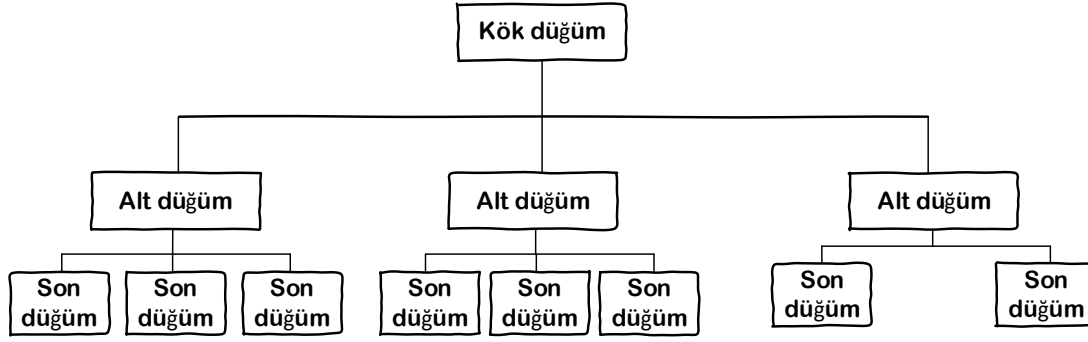
$$\text{Destek (Yaralı, Yol vermeme)} = \frac{\text{Yaralı ve yol vermeme ö. s.}}{\text{Toplam veri sayısı}} = \frac{3}{8} = 0,37 \quad (2.2)$$

$$\text{Güven (Yaralı, Hızlı sürüş)} = \frac{\text{Yaralı ve hızlı sürüş ö. s.}}{\text{Yaralı ögesinin sıklığı}} = \frac{1}{4} = 0,25 \quad (2.3)$$

$$\text{Lift (Yaralı, Yol vermeme)} = \frac{\text{Güven (Yaralı, Yol vermeme)}}{\text{Destek (Yol vermeme)}} = \frac{0,75}{0,5} = 1,50 \quad (2.4)$$

2.1.2. Karar Ağaçları

Karar Ağaçları, sınıflandırma ve tahmin etmede güçlü ve yaygın kullanılan yöntemlerdendir. Ağaç tabanlı yöntemlerin çekiciliği büyük ölçüde sinir ağlarının aksine Karar Ağaçlarının kurallar ortaya çıkarmasından kaynaklanmaktadır. Bir Karar Ağacı, bir dizi soru ve cevabı temsil etmektedir. Karar Ağaçları en üstte kök ve en altta yapraklar olacak şekilde çizilmektedir. Kökte yer alan öznitelik, çıktı (hedef) değişkeni ile en fazla ilişkisi olan öznitelik olmaktadır. Karar Ağaçlarında veri setindeki bir kayıt ağaca kök düğümünden girmekte, sonraki aşamada hangi alt düğümle (node) karşılaşacağını belirlemek için ise bir test uygulanmaktadır. Bu test için farklı algoritmalar vardır. Uygun algoritmaları seçerken amaç, sınıflandırmayı (sınıflar arası ayrımı) en iyi yapan algoritmayı seçmektir. Bu test ve alt düğüme geçme işlemi, yaprak, son (leaf, terminal) düğüme gelene kadar tekrarlanmaktadır (Linoff & Berry, 2011). Şekil 2.4'te Karar Ağacının şematik hali verilmektedir.



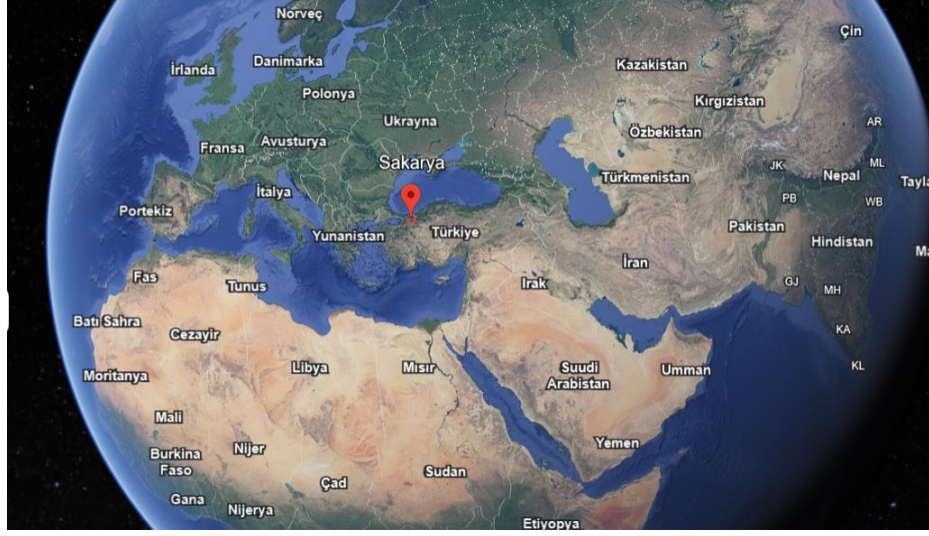
Şekil 2.4. Karar Ağacı gösterimi. Anvari ve ark. (2017)'den uyarlanmıştır (Anvari vd., 2017).

Karar Ağaçları algoritmalarından bazıları; CHAID (Ki-kare Otomatik Etkileşim Algılama/Chi Square Automatic Interaction Detection), CART (Sınıflandırma ve Regresyon Ağaçları/Classification and Regression Trees), C4.5, ID3 ve Rassal Orman (Random Forest)'dir. Bu algoritmalar kategorik verilerde sınıflandırma tahminleri yapmak için çok uygundur (Ijaz vd., 2021). Bu çalışmadaki tüm değişkenler (bağımlı ve bağımsız) kategorik veri tipinde olduğundan (sürekli olan veriler de gruplandırılarak kategorik veri tipine dönüştürülmüştür), analiz yöntemi olarak Karar Ağaçlarından biri olan Ki-kare Otomatik Etkileşim Algılama algoritması (CHAID) seçilmiştir. Bu yöntem Kass tarafından geliştirilmiştir (Kass, 1980). Ki-kare Otomatik Etkileşim Algılama (CHAID) yöntemi, Morgan ve Sonquist tarafından 1963'te geliştirilen Otomatik Etkileşim Algılama (AID)'den sonra geliştirilmiştir (Morgan, J.N. & Sonquist, J.A., 1963; Sonquist & Morgan, J.N., 1964). Oldukça yaygın uygulama alanı olduğuna dair değerlendirmelere Orr'un çalışmasında değinilmiştir (Orr, 1972). CHAID yönteminde Ki-kare değerleri kullanılarak istatistiksel ilişkiler belirlenmekte, bu da oluşturulan Karar Ağaçları sayesinde sınıflandırma yöntemi olarak bu yöntemin uygulama alanı bulmasını sağlamaktadır. Bu yöntemin diğer Karar Ağaçlarından üstünlüğü, ezber (overfitting) yapmadan önce ağacın büyümesini durdurmasıdır.

2.2. Çalışma Alanı

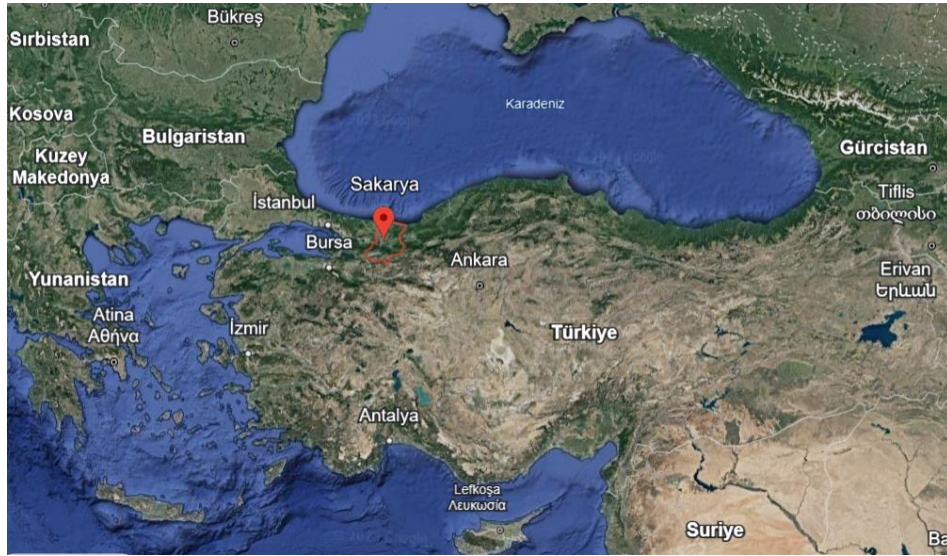
Sakarya, 4817 km² 'lik bir alanı kaplayan ve üzerinde 1,06 milyon kişinin yaşadığı, Türkiye'nin kuzeybatısında yer alan bir büyükşehirdir (T.C. Milli Savunma Bakanlığı Harita Genel Müdürlüğü, 2023; Türkiye İstatistik Kurumu, 2023). Türkiye'de 2021 yılında 1000 kişi başına düşen otomobil sahipliği oranı 162 iken, Sakarya'da 151'dir ve şehir bu açıdan Türkiye ortalamasını temsil edebilmektedir (TÜİK-Trafik kaza

istatistikleri, 2021). Şekil 2.5'te çalışmanın yapıldığı ilin Dünya'daki konumu görülmektedir.



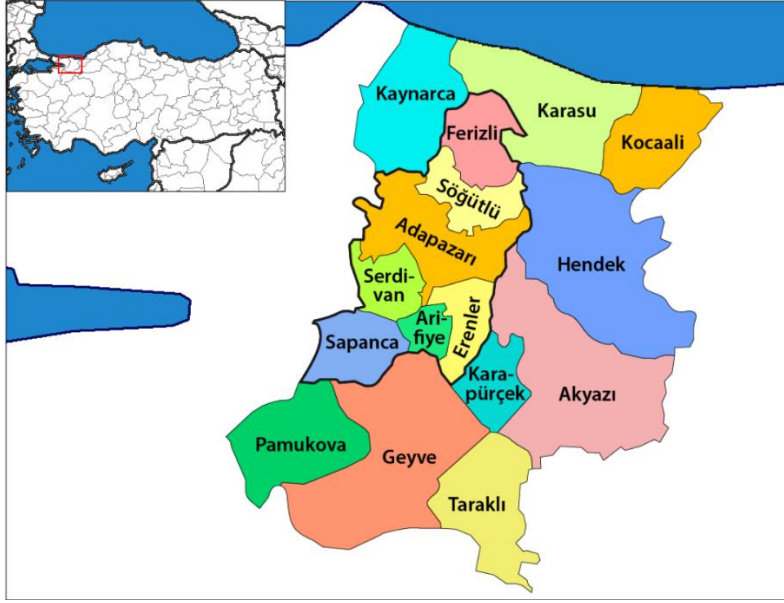
Şekil 2.5. Sakarya şehrinin Dünya'daki konumu (Google Earth, 2023).

Şekil 2.6'da ise Sakarya'nın Türkiye'deki konumu ve çevre iller görülmektedir.



Şekil 2.6. Sakarya'nın Türkiye'deki konumu (Google Earth, 2023).

Şekil 2.7'de Sakarya ilinde bulunan 16 ilçenin konumları verilmiştir. Merkez ilçesi Adapazarı olup nüfus açısından ikinci büyük ilçesi Serdivan'dır.



Şekil 2.7. Sakarya'nın ilçeleri (T.C. Milli Savunma Bakanlığı Harita Genel Müdürlüğü, 2023).

Sakarya'nın karayolu ağı şehir içi yollar, köy yolları, otoyollar, devlet yolları, il yolları ve otoyollardan oluşmaktadır. Sakarya'dan geçen otoyollar O-4 (TEM) ve O-7 (KMO-NMM), devlet yolları ise D-100, D-650 ve D-020'dir. Belirtilmesi gereken bir diğer önemli husus ise Sakarya'nın Türkiye'nin beş büyük şehri olan Kocaeli (37 km), İstanbul (148 km), Bursa (158 km), Bolu (114 km) ve Ankara'nın (306 km) geçiş noktası olmasıdır. Sakarya, İstanbul Kurtköy Sabiha Gökçen Havalimanı'na 110 km, İstanbul Havalimanı'na ise 176 km uzaklıktadır (T.C. Milli Savunma Bakanlığı Harita Genel Müdürlüğü, 2023). 2021 yılında Türkiye genelinde iyileştirilmesi planlanan 20 kaza kara noktasından biri Sakarya'dadır (KGM-Kaza kara noktaları, 2022). Şehirdeki toplam kayıtlı motorlu taşıt sayısı 348.935 olup bunun 175.120 'si otomobildir (Türkiye İstatistik Kurumu, 2023). Toplam kaza sayısı 15.990, ölümlü-yaralanmalı kaza sayısı 2789, ölen kişi sayısı ve yaralanan kişi sayısı 2022 yılında sırasıyla 66, (30'u kaza yerinde, 36'sı kaza sonrası) ve 4055 olarak verilmiştir (Türkiye İstatistik Kurumu, 2023).

Şehrin konumu, trafik kaza ve taşıt sayısı ile ilgili oldukça kapsamlı verilerin elde bulunması dolayısı ile analiz için Sakarya ili seçilmiştir. Çalışma alanı, şehrin tüm ulaşım ağını kapsamakta olup ilde bulunan mevcut yollar, sokaklar, caddeler, devlet ve otoyollarının tümüne ait veri setleri kullanarak analiz ve incelenmeler yapılmıştır.

CRISP DM sürecinde yer alan iş ve verinin anlaşılması kısımları analiz öncesinde dikkate alınmıştır. Literatür kapsamında incelenen trafik kaza analizlerinin yapıldığı

alışmalar doęrultusunda, Sakarya’da gerekleşen ölümlü-yaralanmalı trafik kazalarının veri madencilięi yöntemlerinden olan tanımlayıcı yöntemlerden Birliktelik Kuralları, tahmin edici yöntemlerden de Karar Aęaçları ile analiz edilmesinin, veri setindeki özniteliklerin büyük çoęunluęunun kategorik yapıda olması ayrıca alışmanın kapsamı ile uyumlu olması dolayısı ile uygun olduęu belirtilmelidir.

2.3. Veri Seti

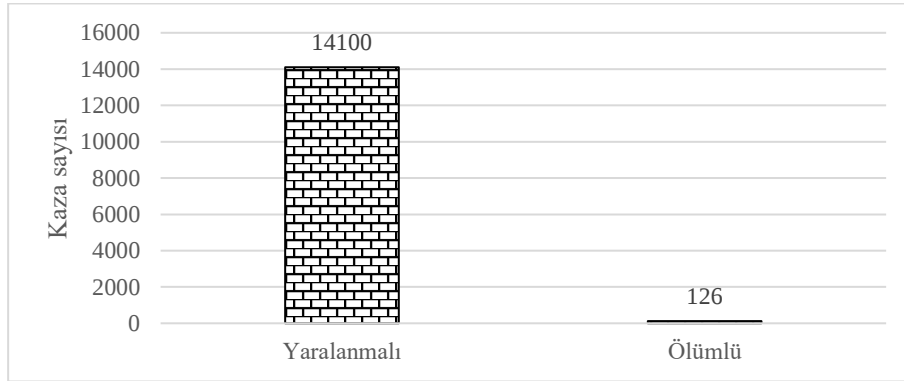
Bu tezde, Sakarya’da 2015 yılının başından 2020 yılının sonuna kadar gerekleşen ölümlü ve yaralanmalı kazalara ait T.C. İçişleri Bakanlığı Emniyet Genel Müdürlüęü’nden alınan kaza tutanaklarından elde edilen veriler ile alışılmıştır. Türkiye de 2015 yılından itibaren kazadan sonra 30 güne kadar hasta takipleri yapıp tutanaklar tutulmaya başlandıęı için, bu sürenin de analizlerde yer alabilmesi adına başlangı yılı olarak 2015 belirlenmiştir. Veri setinde Jandarma Genel Komutanlıęı’nın kayıt altına aldıęı veriler olmadığından nceki bölümde verilen tablodaki sayılar arasında farklılıklar bulunmaktadır. Veri setinde toplam 14.226 ölümlü-yaralanmalı kaza bulunmakta olup bunların %84’ü (11.886) yerleşim yerinde %16’sı ise (2.340) yerleşim yeri dışında gerekleşmiştir. Veri seti, Emniyet Genel Müdürlüęü’nden üç ayrı dosya şeklinde alınmış olup 1. dosyada kazaya, 2. dosyada taşıt ve sürücüye, 3. dosyada yolcu ve yayalara ait bilgiler bulunmaktadır. 1. dosyada bulunan kazaya ait bilgilerden bazıları; kaza oluş türü, kazanın konumu, kazanın tarihi, kaza saati, kazanın olduęu yol kesimine ait geometrik özellikler ve yol yüzeyi ile ilgili açıklayıcı ve tanımlayıcı bilgiler olarak ifade edilebilir. Sürücü ve taşıtlara ait bilgilerin sunulduęu 2. dosyada sürücü kusuru, sürücüye ait yaş-eęitim durumu, sürücünün kaza sonucu durumu, taşıt modeli (yaşı) ve cinsi gibi bilgiler bulunmaktadır. Yolcu ve yayaya ait bilgilerden oluşan 3. dosya, yolcu-yaya yaşı, yolcu-yaya kaza sonucu durumu, yolcu-yaya kusuru içerikli verileri bulundurmaktadır. Bu dosyalar, kaza ve taşıtlara ait tanımlayıcı ve ayırt edici (ID) bilgilere dayalı olarak bir araya getirilmiş ve tek bir dosyaya indirgenmiştir.

2.3.1. Veriye ait tanımlayıcı istatistikler

Bu tezde analiz edilen ham veri dosyasında 2015-2020 yıllarında gerekleşen ölümlü-yaralanmalı trafik kazaları yer almaktadır. T.C. Emniyet Genel Müdürlüęü Trafik Başkanlıęı’ndan alınan Excel dosyasında 3 sayfa bulunmaktadır. İlk sayfada kazaya (14.226 kaza), 2. sayfada sürücülere ve taşıtlara (23.551 sürücü ve taşıt), 3. sayfada

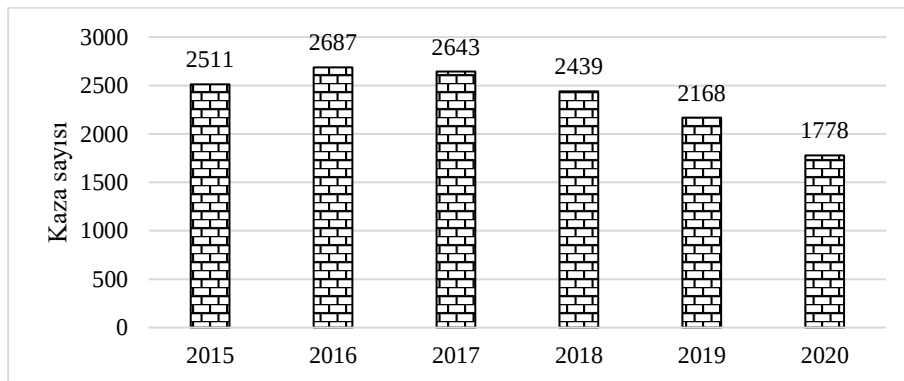
yolcu ve yayalara (12.517 yolcu ve yaya toplamı) ait bilgiler yer almaktadır. Bu tez çalışmasında, çalışmanın odak noktasının farklı oluşu nedeniyle 3. sayfadaki veriler kullanılmamıştır. Ham veriye ait histogramlar aşağıdaki şekillerde verilerek, verinin ilk durumundaki yapısı gösterilmektedir.

Şekil 2.8’de trafik kazalarının 14.100’ünün (%99) yaralanma, 126’sının (%0,9) ölümlle sonuçlandığı görülmektedir.



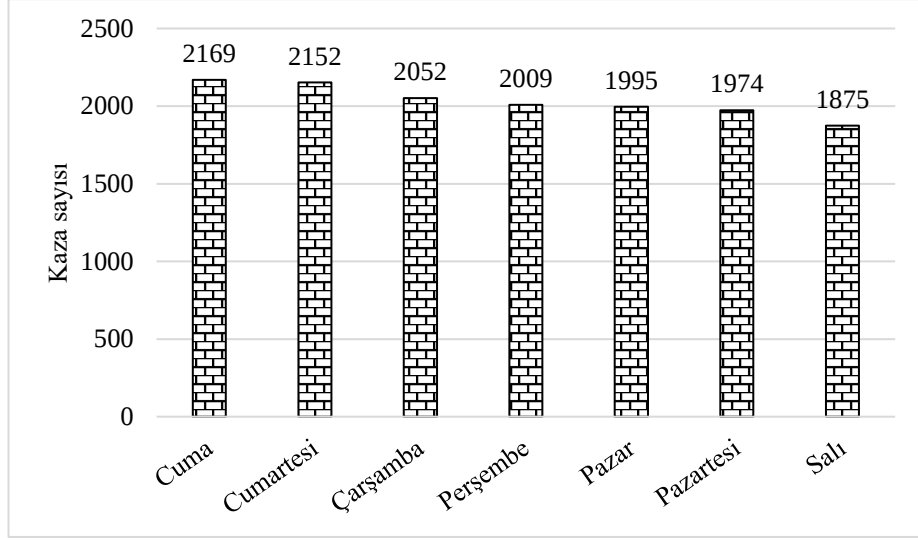
Şekil 2.8. Trafik kazalarının sonuçlarına göre histogramı.

Şekil 2.9’da 2015, 2016, 2017, 2018, 2019 ve 2020 yıllarında gerçekleşen ölümlü-yaralanmalı kaza sayıları yıllara göre sıralanarak verilmiştir. 2020 yılında sayılardaki belirgin düşüşün nedeni COVID-19 salgını nedeniyle sokağa çıkma yasakları ve trafikteki taşıt sayısının azalmasına bağlı olarak trafik kaza sayılarının da azalmasıdır.



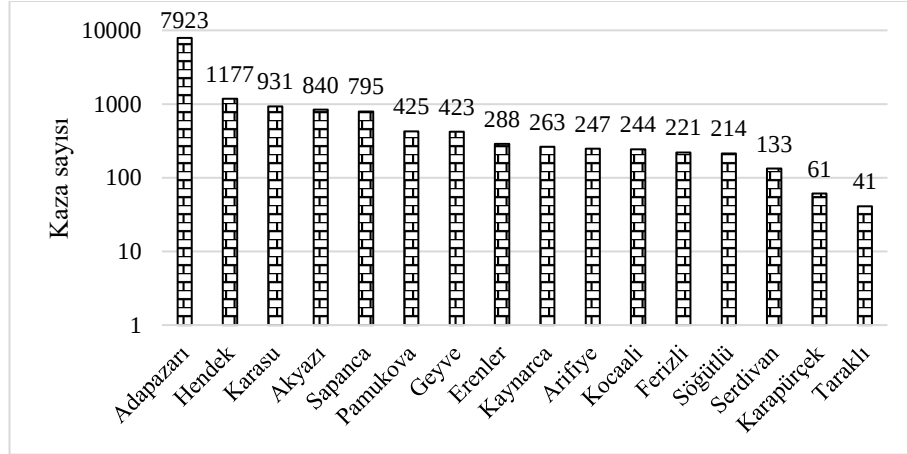
Şekil 2.9. 2015-2020 yıllarında Sakarya’da gerçekleşen trafik kazalarının yıllara göre histogramı.

Şekil 2.10’da trafik kazalarının günlere göre histogramı verilmiştir. Cuma ve cumartesi günlerinde göreceli yüksek yolculuk değerlerinin söz konusu olmasına bağlı olarak daha fazla trafik kazası gerçekleşmektedir.



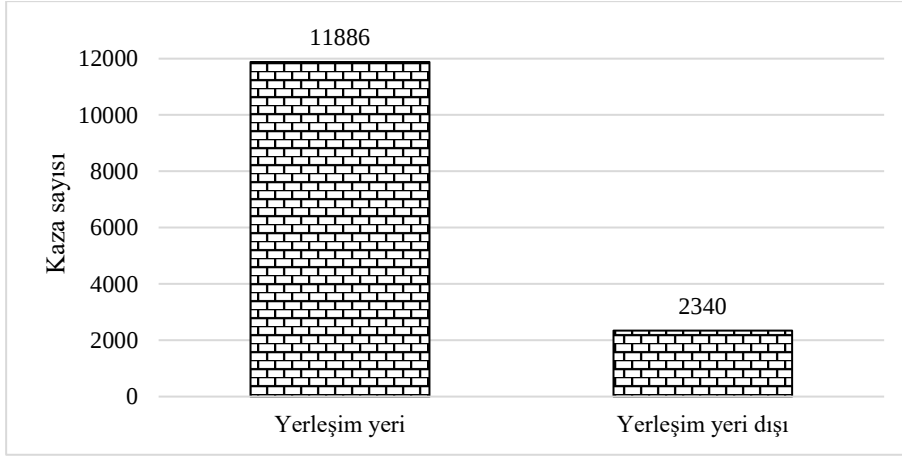
Şekil 2.10. Trafik kazalarının haftanın günlerine göre histogramı.

Şekil 2.11’de kazaların Sakarya’da ilçelere göre histogramı verilmektedir. Tüm kazaların %56’sının merkez ilçe olan Adapazarı’nda gerçekleştiği görülmektedir.



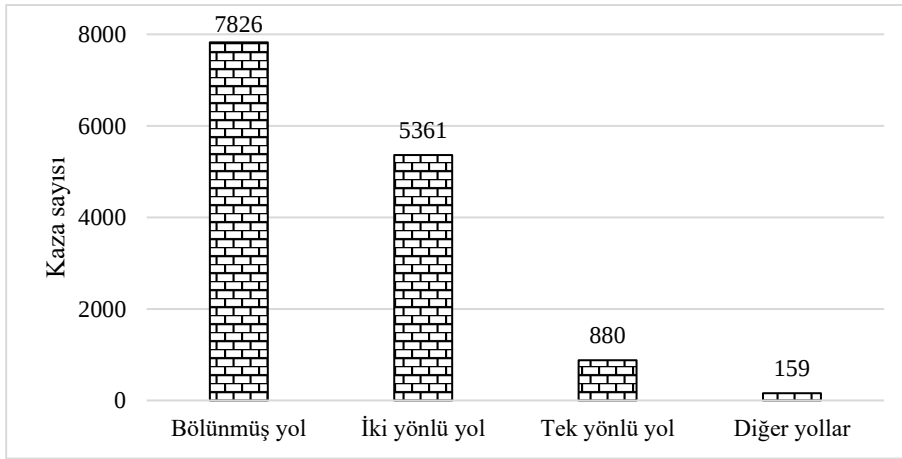
Şekil 2.11. Trafik kazalarının ilçelere göre histogramı.

Şekil 2.12’de kazaların gerçekleştiği yol kesiminin yerleşim yeri içi ya da dışında olması ile ilgili histogram yer almaktadır. Kazaların %84’ünün yerleşim yeri içinde gerçekleştiği görülmektedir.



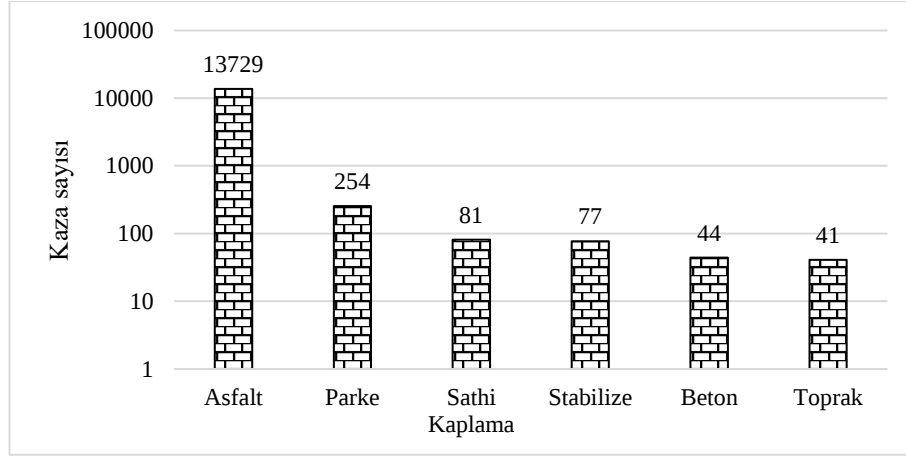
Şekil 2.12. Trafik kazalarının yerleşim yerlerine göre histogramı.

Şekil 2.13'te kazaların yol tipine göre histogramı verilmektedir. %55'inin bölünmüş yollarda, %38'inin iki yönlü yollarda gerçekleştiği görülmektedir.



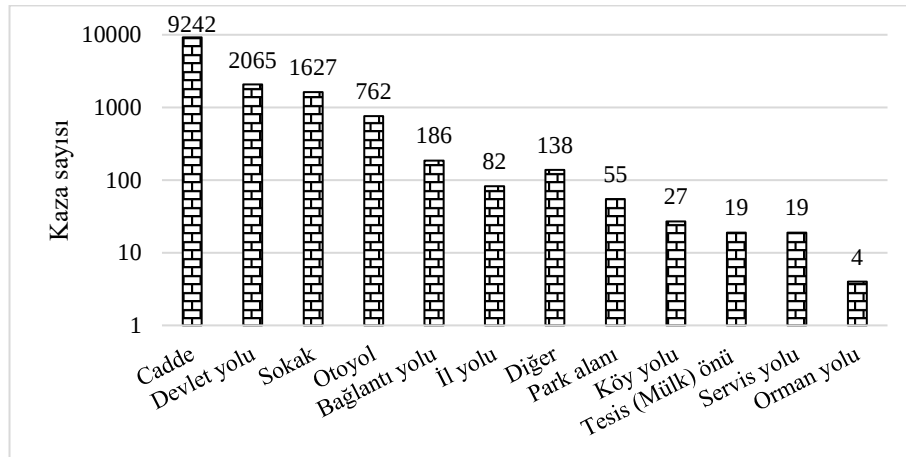
Şekil 2.13. Trafik kazalarının yol tipine göre histogramı.

Şekil 2.14'te kazaların gerçekleştiği yolların kaplama çeşidine göre histogramı verilmektedir. Kazaların %97'sinin asfalt yollarda gerçekleştiği görülmektedir. Diğer yol kaplamalarında meydana gelen kazaların oranları yaklaşık binde 2-3 civarlarında olup çok düşüktür.



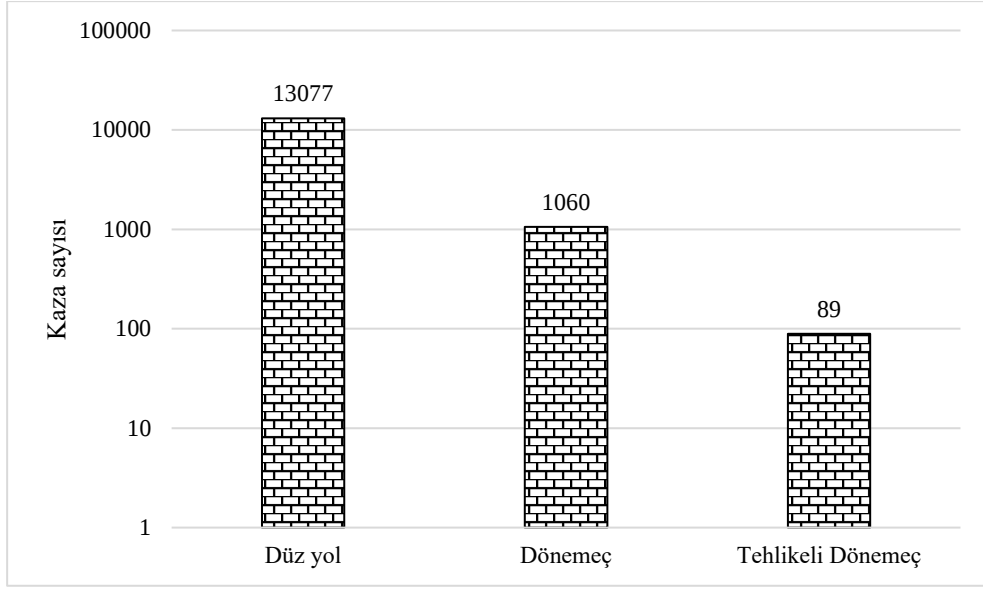
Şekil 2.14. Trafik kazalarının yol kaplamasına göre histogramı.

Şekil 2.15'te gerçekleşen trafik kazalarının yol sınıflarına göre histogramı verilmiştir. Meydana gelen trafik kazalarının %68'i caddelerde, %14'ü illeri birbirine bağlayan devlet yollarında, %11'i sokaklarda, %5'i otoyollarda, geri kalan %2 si de grafikte görülen diğer yollarda olmuştur.



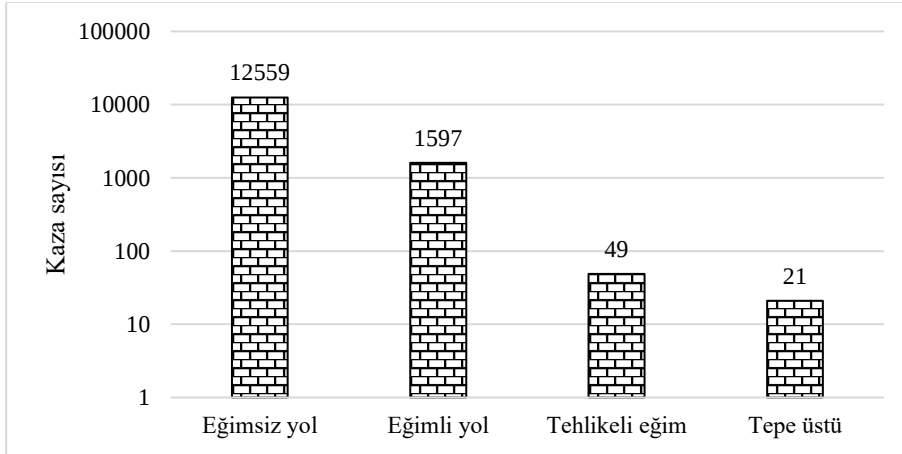
Şekil 2.15. Trafik kazalarının yol sınıfına göre histogramı.

Kazaların gerçekleştiği yol kesiminde yatay yol geometrisinde dönemeç ve tehlikeli dönemeç olup olmaması ile ilgili trafik kaza tutanaklarına veri girişi yapılmıştır. Şekil 2.16'da verilen histogramlara göre kazaların %92'si dönemeç olmayan düz yol kesimlerinde meydana gelmiştir.



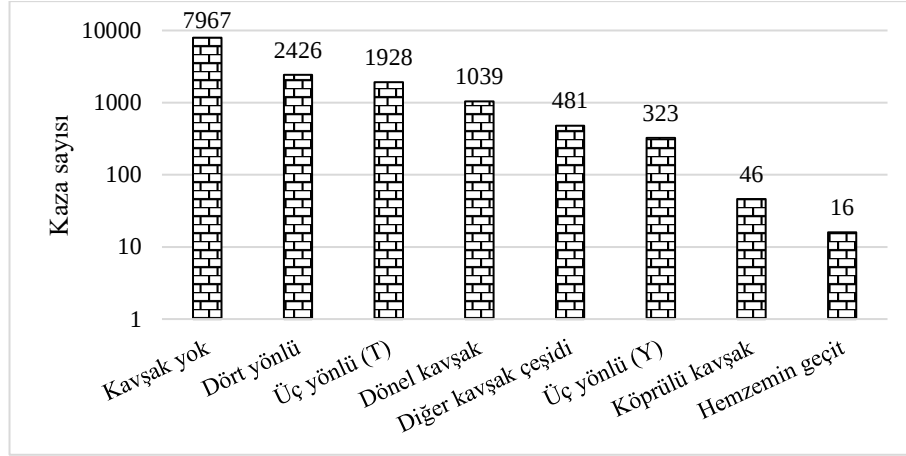
Şekil 2.16. Trafik kazalarının yatay yol geometrisine göre histogramı.

Şekil 2.17’de düşey yol geometrisi ile ilgili histogramlar yer almakta olup kazaların %88’i eğimsiz yol kesimlerinde meydana gelmiştir.



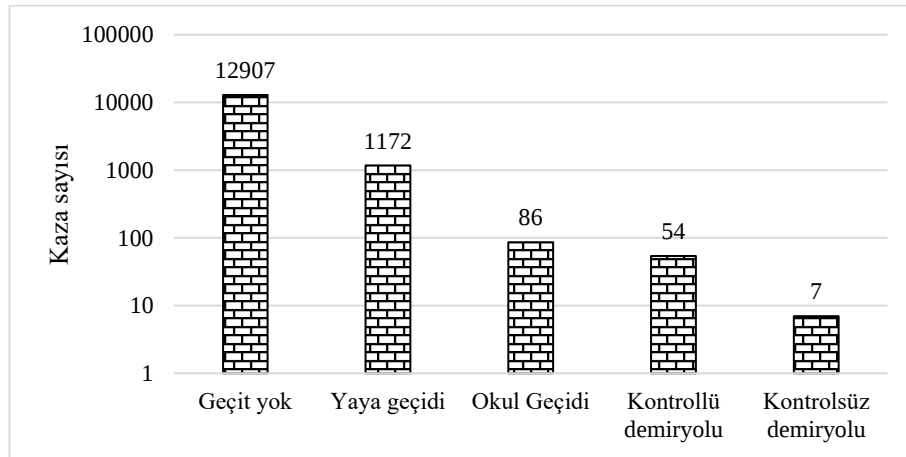
Şekil 2.17. Trafik kazalarının düşey yol geometrisine göre histogramı.

Şekil 2.18’de kazaların gerçekleştiği yol kesimlerinde kavşak varlığı, varsa çeşidinin histogramları verilmiştir. Kazaların %56’sının kavşak olmayan yol kesimlerinde gerçekleştiği görülmektedir. Dört kollu, 3 kollu T ve dönel kavşaklarda gerçekleşen kaza yüzdeleri sırasıyla %17, %14 ve %7 olmuştur.



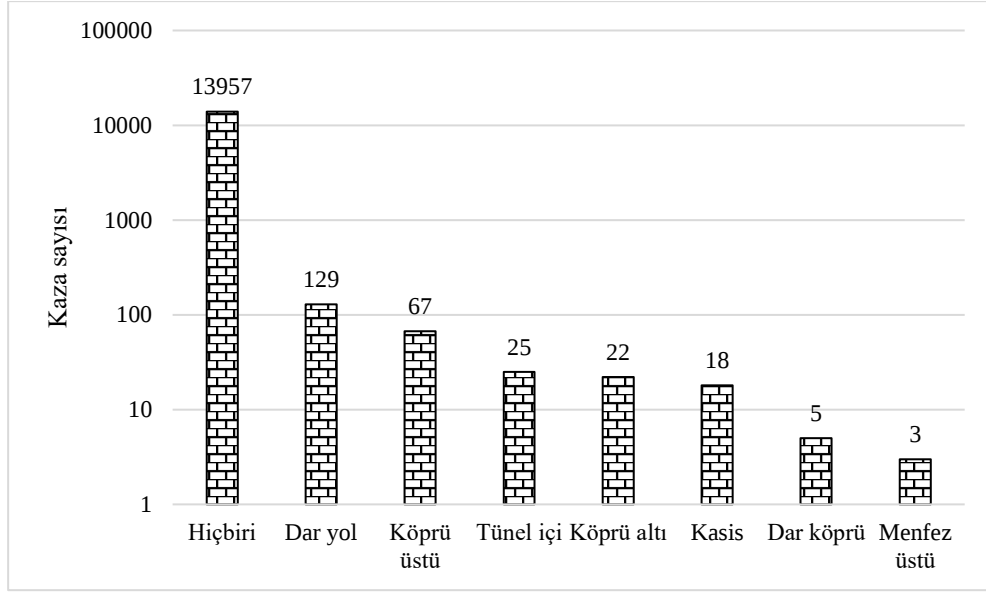
Şekil 2.18. Trafik kazalarının kavşak türüne göre histogramı.

Şekil 2.19’da trafik kazaların gerçekleştiği yol kesiminde geçit varlığı ile ilgili histogramlar verilmiştir. Kazaların gerçekleştiği yol kesimlerinin %91’inde geçit olmadığı görülmektedir. Yaya geçidinde gerçekleşen kazalar ise tüm kazaların %8’ine denk gelmektedir.



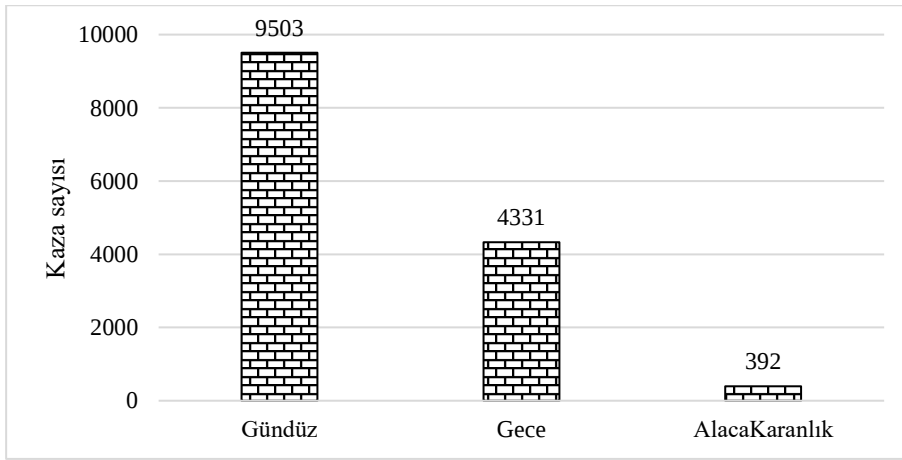
Şekil 2.19. Trafik kazalarının geçit türüne göre histogramı.

Şekil 2.20’de kazaların gerçekleştiği yol kesiminde yola ait başka bir unsurun varlığı ile ilgili bilgilerin histogramı verilmiştir. Kazaların gerçekleştiği yol kesimlerinin %98’inde hiçbir farklı unsur olmadığı görülmektedir.



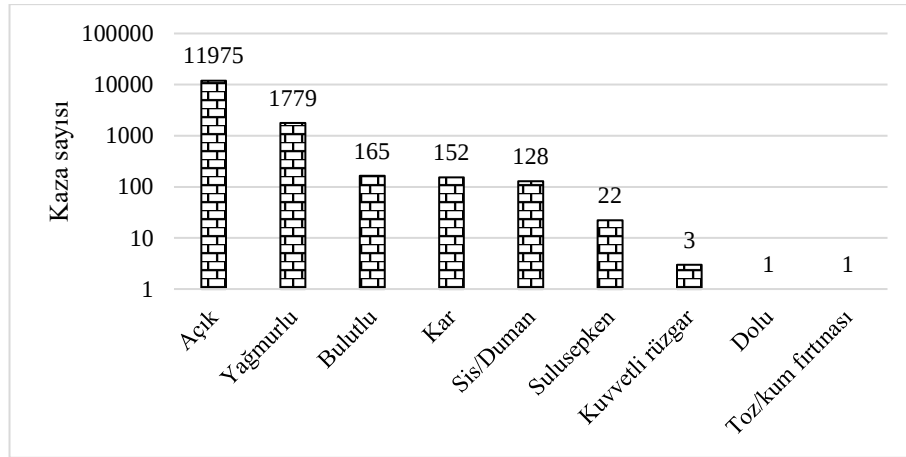
Şekil 2.20. Trafik kazalarının diğer yol unsurlarına göre histogramı.

Şekil 2.21’de kazaların gerçekleştiği anda gün ışığı histogramları verilmiştir. Kazaların %67’sinin gündüz, %30’unun gece ve %3’ünün alaca karanlıkta gerçekleştiği görülmektedir.



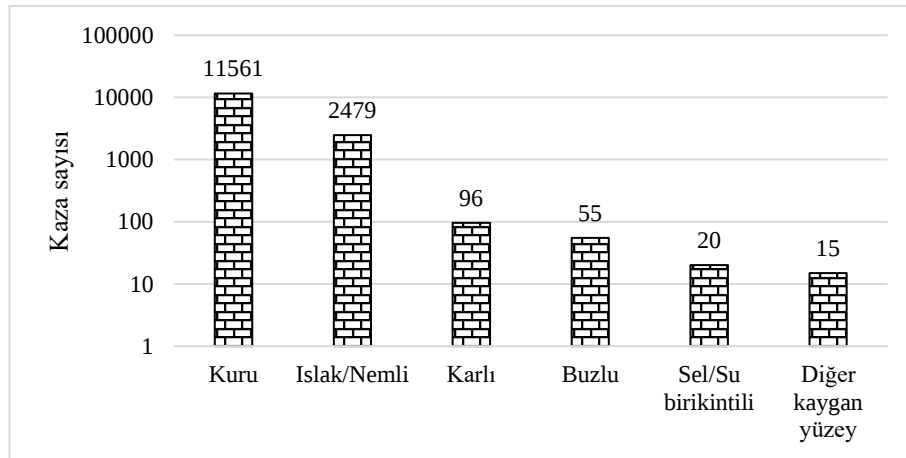
Şekil 2.21. Trafik kazalarının gün ışığına göre histogramı.

Şekil 2.22’de kazaların gerçekleştiği andaki hava durumu bilgilerine ait histogram verilmiştir. Kazaların %84’ünün açık, %12’sinin yağmurlu, geri kalanında bulutlu, karlı, sis/dumanlı gibi hava koşullarında gerçekleştiği görülmektedir.



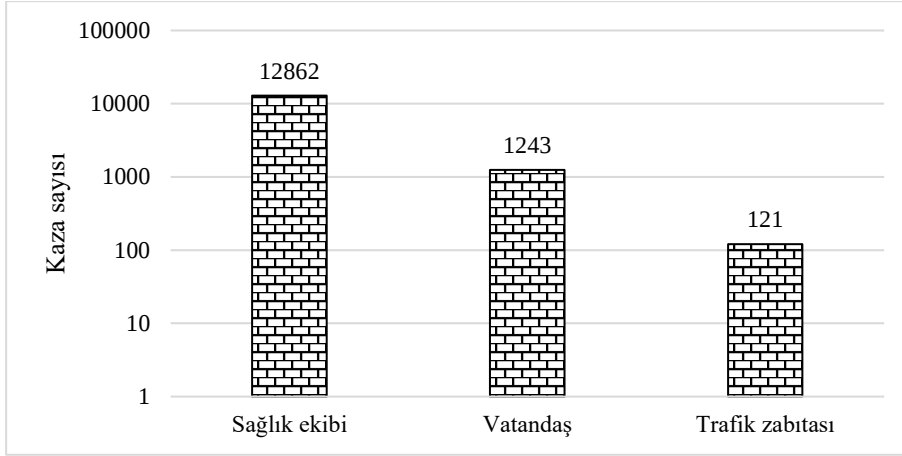
Şekil 2.22. Trafik kazalarının hava durumuna göre histogramı.

Şekil 2.23'te kaza anında yol yüzeyinin kuru ya da diğer (ıslak, karlı, buzlu, su birikintili, diğer kaygan yüzey) durumları ile ilgili histogramları verilmektedir. Trafik kazalarının %81'inin kuru, %17'sinin ıslak/nemli yollarda gerçekleştiği görülmektedir. Karlı, buzlu, sel/su birikintisi olan ve diğer kaygan yüzeyler de ıslak/nemli sınıfında yer alabilirken, tutanaklarda bu şekilde yazıldığından ham veride de ayrı sınıflarda yer almıştır.



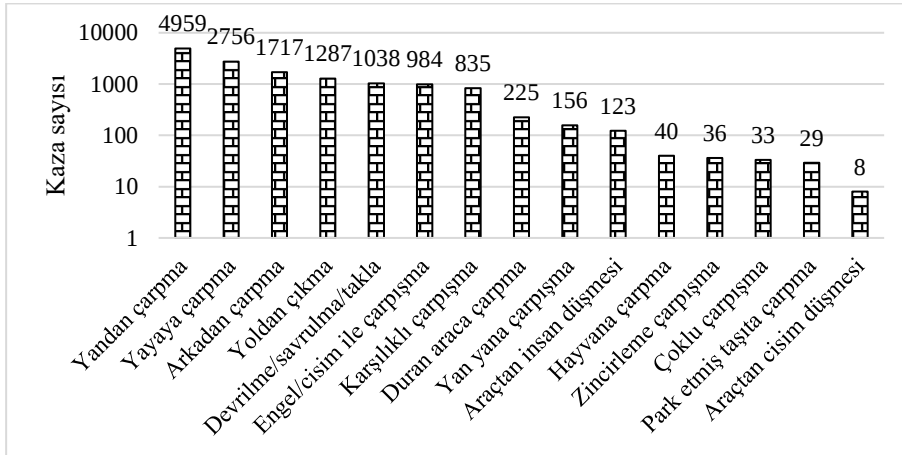
Şekil 2.23. Trafik kazalarının yol yüzeyine göre histogramı.

Trafik kazalarında kazazedelere ilk yardımın kimler tarafından yapıldığı ile ilgili histogram Şekil 2.24'te verilmiştir. Tüm kazaların %90'ında sağlık ekibi, %9'unda kazanın çevresinde o anda olan vatandaşlar, %1'inde ise trafik zabıtası ilk yardım gerçekleştirmiştir.



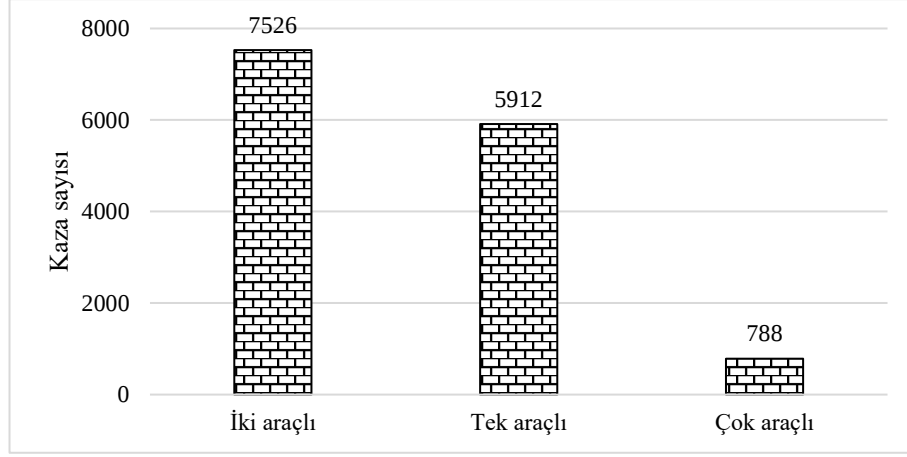
Şekil 2.24. Trafik kazalarında ilk yardım histogramı.

Şekil 2.25’te kaza oluş türüne göre histogramlar verilmiştir. Kazaların %35’i yandan çarpma, %19’u yayaya çarpma, %12’si arkadan çarpma, %9’u yoldan çıkma, %7’si devrilme/savrulma/takla, %7’si engel/cisim ile çarpışma kalan %11’i histogramda belirtilen diğer şekillerde gerçekleşmiştir.



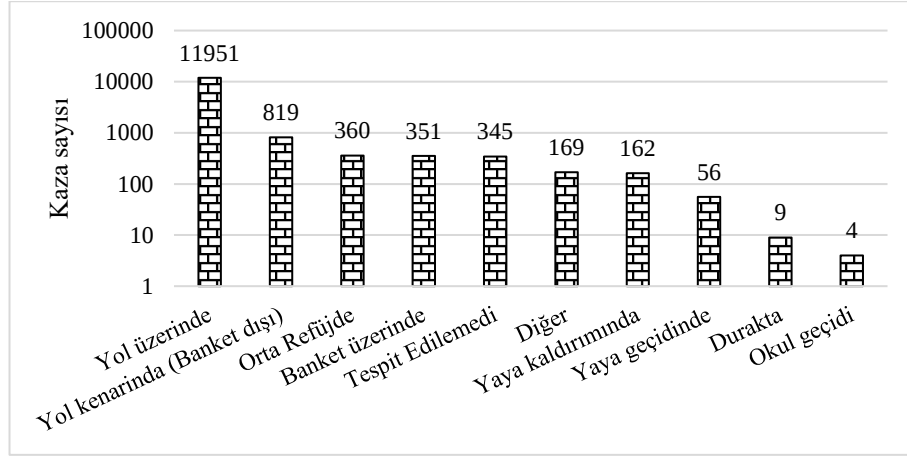
Şekil 2.25. Trafik kazalarının oluş türüne göre histogramı.

Şekil 2.26’da kazalara karışan taşıt sayısına göre histogramlar verilmiştir. Kazaların %53’ü iki taşıtın yer aldığı, %42’si tek taşıt, %5’i ise ikiden fazla taşıtın yer aldığı kazalardır.



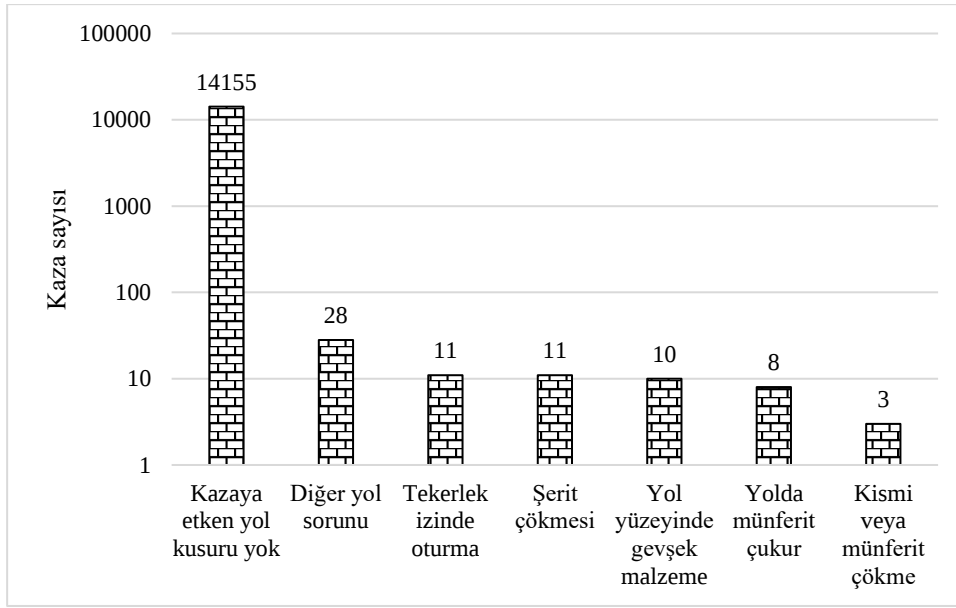
Şekil 2.26. Trafik kazalarına karışan taşıt sayısına göre histogramı.

Şekil 2.27’de kazaların gerçekleştiği çarpışma yeri histogramı verilmiştir. Kazaların %84’ü yol üzerinde gerçekleşirken, geri kalanları ise banket, banket dışında, orta refüjde, yaya kaldırımı gibi yerlerde gerçekleşmiştir.



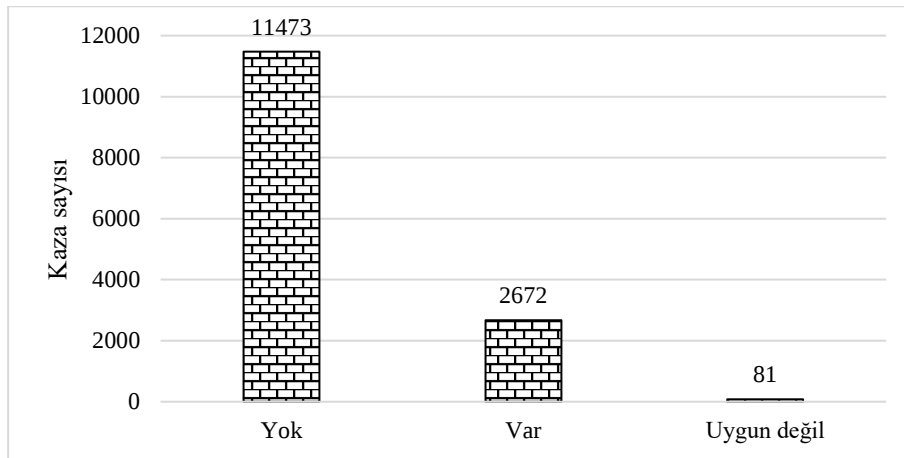
Şekil 2.27. Trafik kazalarının gerçekleştiği çarpışma yerine göre histogramı.

Şekil 2.28’de kazaların gerçekleştiği yollardaki kusur histogramı verilmiştir. Buna göre %99,5’inde yol kusuru olmadığı görülmektedir.



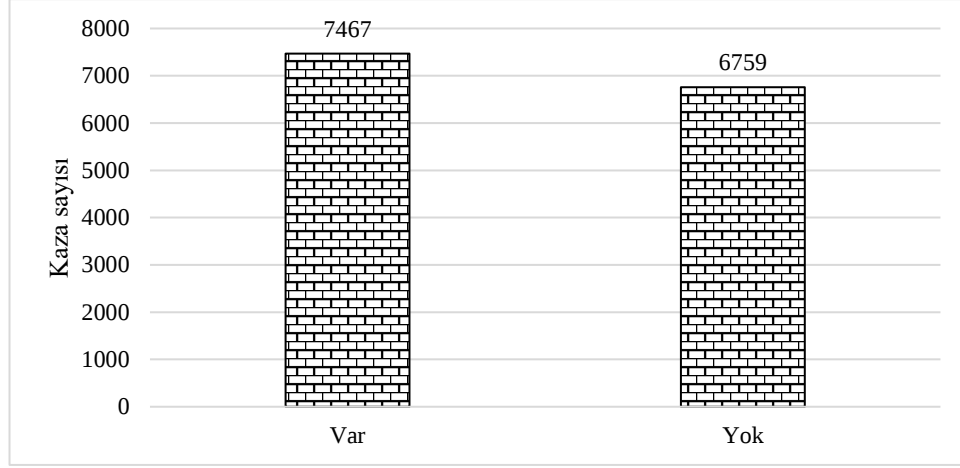
Şekil 2.28. Trafik kazalarının gerçekleştiği yerlerde yol kusurları histogramı.

Şekil 2.29’da kazaların gerçekleştiği yerlerdeki otokorkuluk varlığına ait koşullar verilmiş olup %80,7’sinde otokorkuluğun olmadığı, %18,7’sinde var olduğu, %0,5’inde var ama uygun olmadığı görülmektedir.



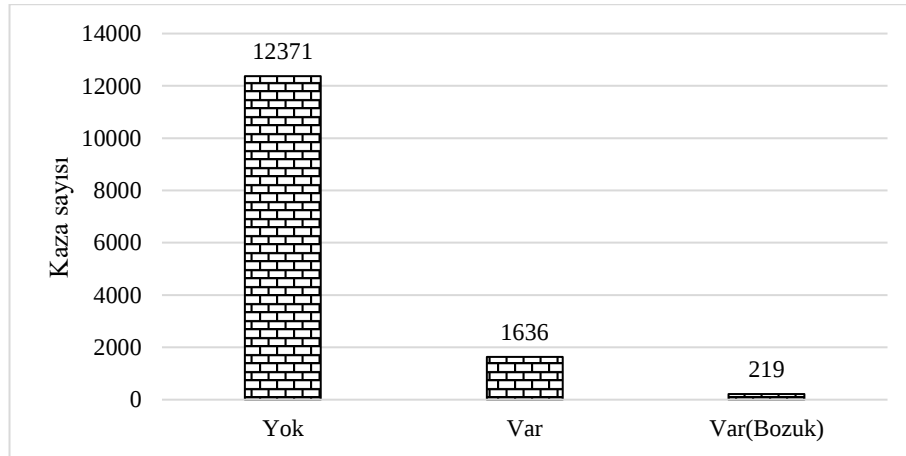
Şekil 2.29. Trafik kazalarının gerçekleştiği yerlerde otokorkuluk varlığı histogramı.

Şekil 2.30’da trafik kazalarının olduğu yerlerdeki yaya kaldırımı varlığına dönük durumlar verilmiş olup %52’sinde yaya kaldırımı olduğu, %48’inde yaya kaldırımı olmadığı görülmektedir.



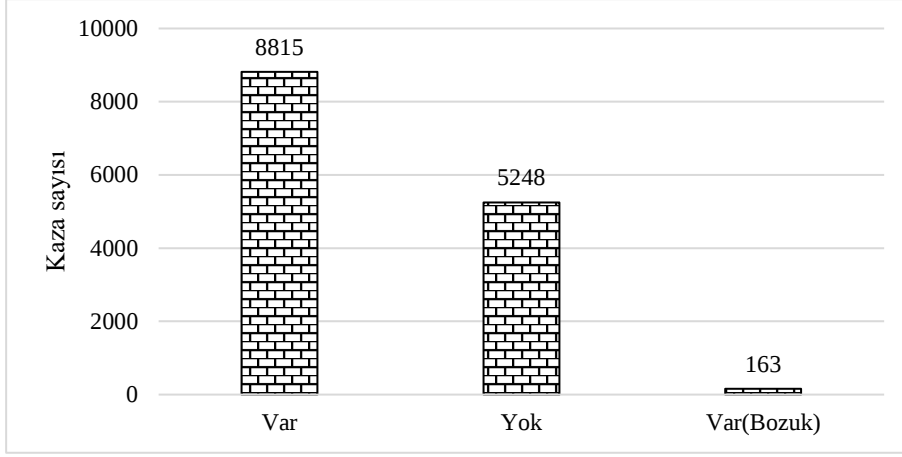
Şekil 2.30. Trafik kazalarının gerçekleştiği yerlerde yaya kaldırımı varlığı histogramı.

Şekil 2.31’de kazaların gerçekleştiği yerlerdeki trafik ışığı varlığı ile ilgili histogramlar verilmiş olup trafik ışığının, kazaların %87’sinde olmayıp, %12’sinde var ve %1’inde bozuk halde olduğu görülmektedir.



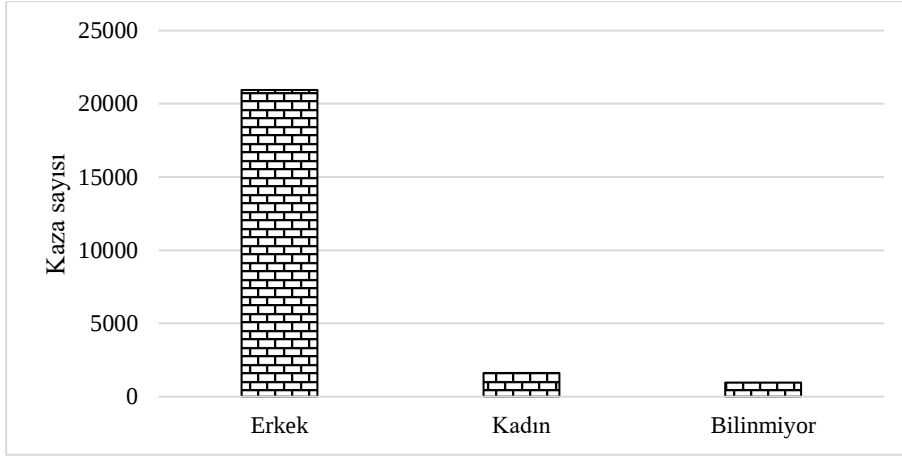
Şekil 2.31. Trafik kazalarının gerçekleştiği yerlerde trafik ışığı varlığı histogramı.

Şekil 2.32’de kazaların gerçekleştiği yerlerde aydınlatma varlığı ile ilgili histogram verilmiştir. Şekilden, kaza konumlarının %62’sinde aydınlatma var iken, %37’sinde olmadığı, %1’inde ise var ama bozuk olduğu görülmektedir.



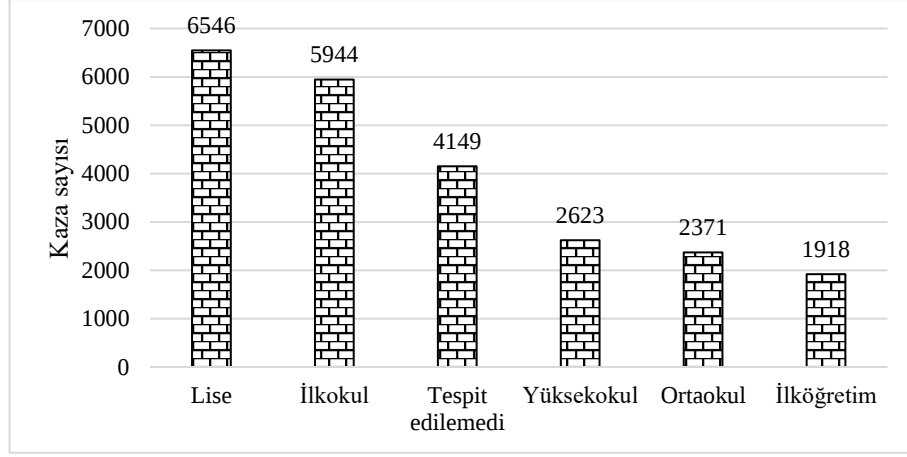
Şekil 2.32. Trafik kazalarının gerçekleştiği yerlerde aydınlatma varlığı histogramı.

Şekil 2.33’te trafik kazası geçiren sürücülerin cinsiyetlerinin histogramı verilmiştir. Sürücülerin %89’unun erkek, %7’sinin kadın olduğu %4’ünün cinsiyet bilgisinin kayıt altına alınamadığı görülmektedir.



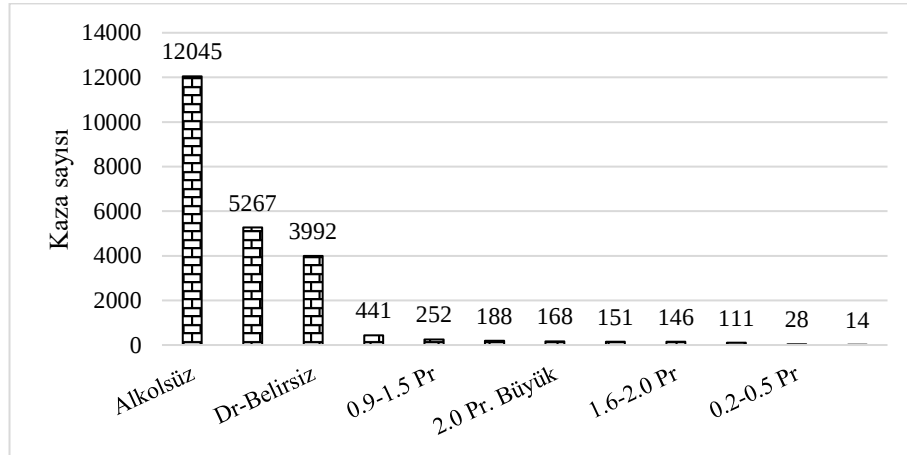
Şekil 2.33. Trafik kazası geçiren sürücülerin cinsiyet histogramı.

Şekil 2.34’te trafik kazası geçiren sürücülerin eğitim durumu histogramı verilmiştir. Sürücülerin %28’inin lise, %25’inin ilkokul, %11’inin yüksek okul, %10’unun ortaokul, %8’inin ilköğretim mezunu olduğu görülmektedir. Sürücülerin %14’ünün eğitim durumu ile ilgili bilgi bulunmamaktadır. İlkokul ve ilköğretim sınıfları, çalışmanın analiz kısımlarında ilköğretim sınıfı olarak birleştirilmiştir.



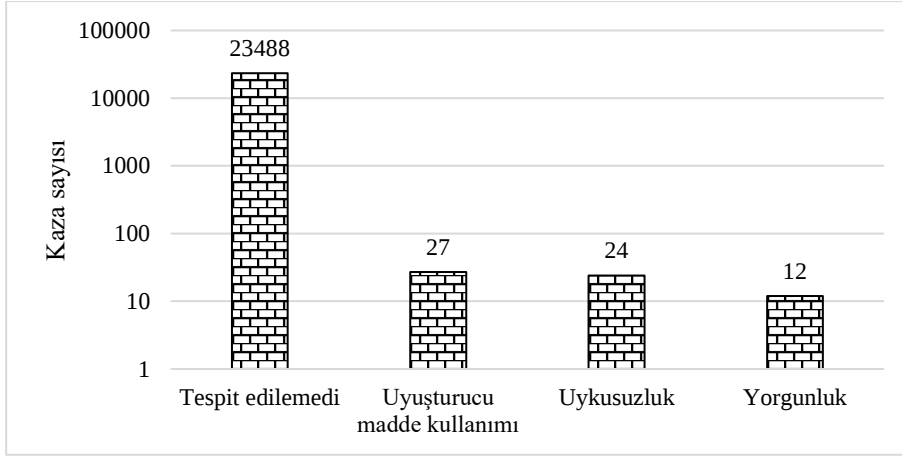
Şekil 2.34. Trafik kazası geçiren sürücülerin eğitim durumuna göre histogramı.

Şekil 2.35'te sürücülerin alkol kullanımına ait histogram verilmiştir. Sürücülerin %51'inin alkolsüz olduğu, %22'sinin doktor kontrolü ile alkolsüz olduğuna karar verildiği, %20'sinin belli olmadığı ve kalan sürücülerin de şekilde görüldüğü gibi kanlarındaki alkol seviyesine göre oranları verilmiştir.



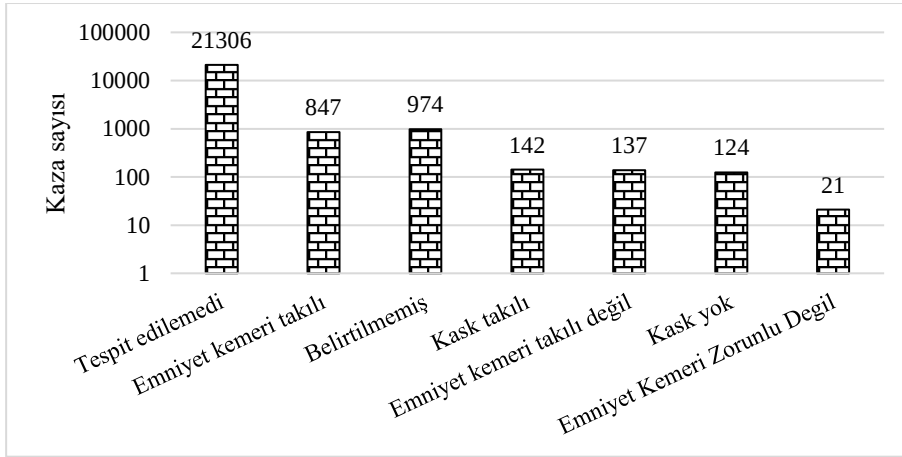
Şekil 2.35. Trafik kazası geçiren sürücülerin alkol kullanımına göre histogramı.

Şekil 2.36'da trafik kazası geçiren sürücülerin psikolojik ve fiziksel durumuna göre verilen bilgilerin histogramı vardır. Sürücülerin yalnızca binde 2'si ile ilgili uyuşturucu madde kullandığı, uykusuz ya da yorgun olduğu bilgisi verilmiştir.



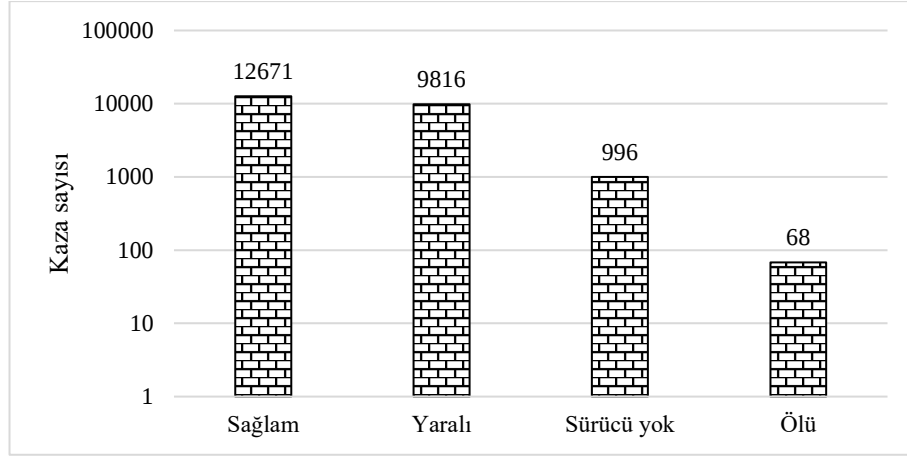
Şekil 2.36. Trafik kazası geçiren sürücülerin psikolojik/fiziksel durumuna göre histogramı.

Şekil 2.37’de sürücülerin kaza anında emniyet kemeri/kask kullanımı ile ilgili verilere ait histogram verilmiştir. Sürücülerin %94’ünde bu kullanımlar belirlenememiş olup bu durum histogramda görülmektedir.



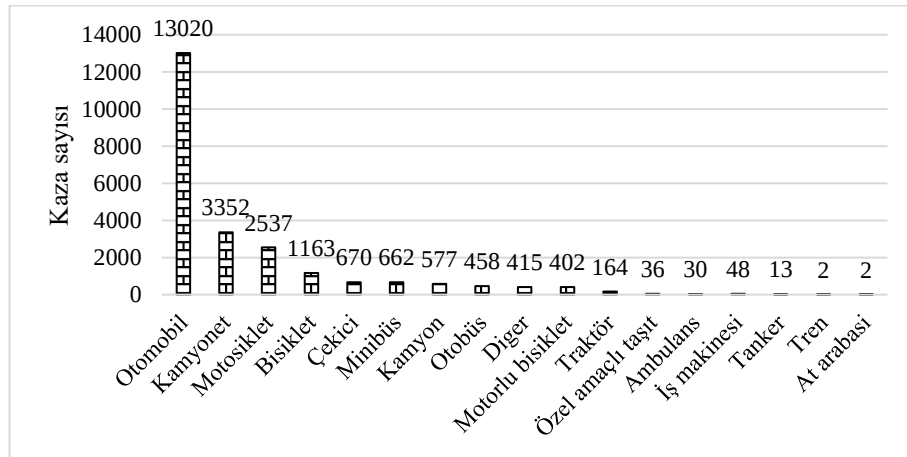
Şekil 2.37. Trafik kazası geçiren sürücülerin emniyet kemeri/kask kullanımına göre histogramı.

Şekil 2.38’de sürücülerin kaza sonucuna bağlı olarak karşılaştıkları sonuçlara ait histogram verilmiştir. Buna göre sürücülerin %53,8’inin sağlam, %41,7’sinin yaralı, %4’ünün taşıt içinde olmadığı, %0,3’ünün öldüğü görülmektedir.



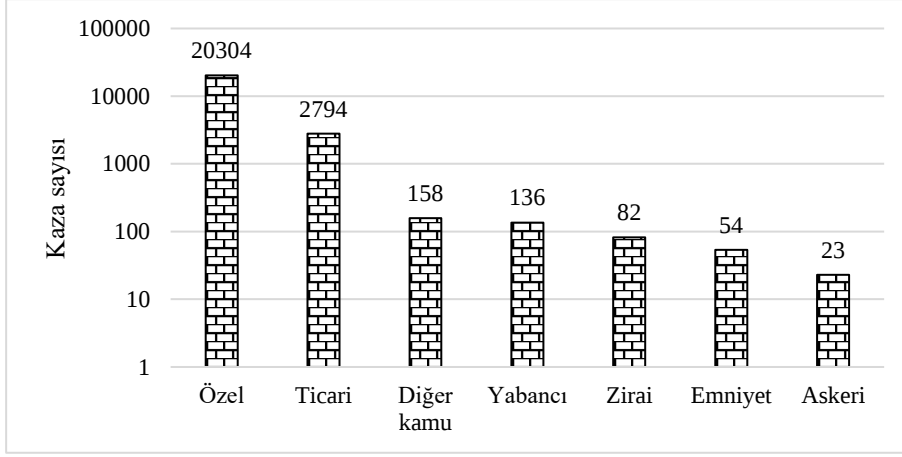
Şekil 2.38. Trafik kazası geçiren sürücülerin kaza sonucuna göre histogramı.

Şekil 2.39’da trafik kazasına karışan taşıtların türlerinin verildiği histogram yer almaktadır. Buna göre, kazalara karışan taşıtların %55’i otomobil, %14’ü kamyonet, %11’i motosiklet olup diğer taşıt türlerine ait oransal değerler de histogramda sunulmaktadır.



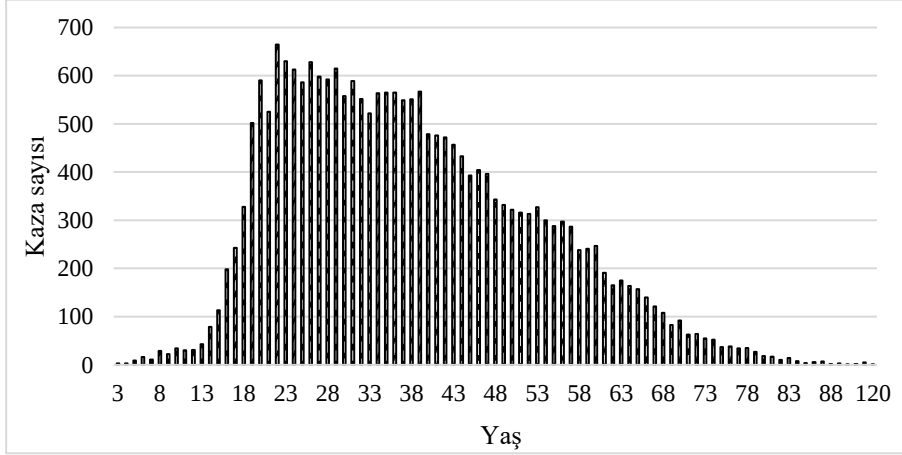
Şekil 2.39. Trafik kazasına karışan taşıtların türel histogramı.

Şekil 2.40’ta kazalara karışan taşıtların kullanım tipine göre histogramı verilmiş olup %86’sının özel, %12’sinin ticari, kalanlarının da kamu, yabancı, ziraî, emniyet, askerî amaçla kullanıldıkları görülmektedir.



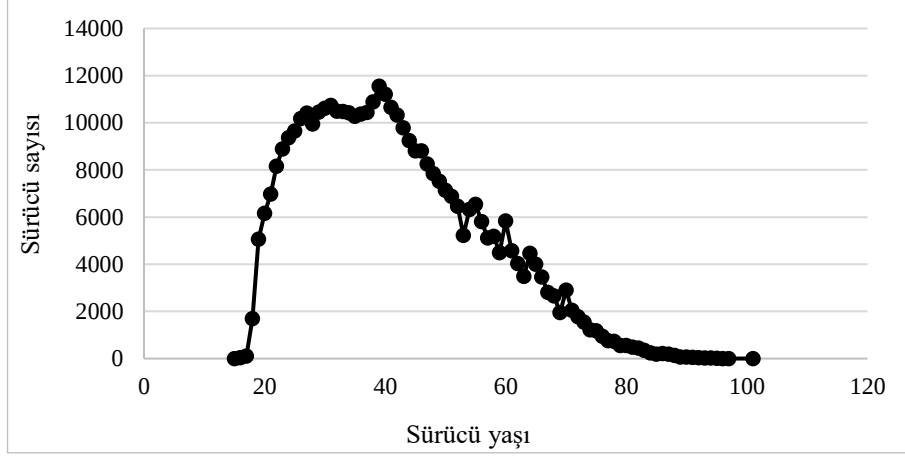
Şekil 2.40. Trafik kazasına karışan taşıtların kullanım tipine göre histogramı.

Şekil 2.41’de kaza geçiren sürücülerin yaşlarının histogramı verilmiştir. Bu anlamda kazaya karışan sürücü yaş aralığının 18-43 yaş arasında en yüksek değerde olduğu görülmektedir. Şekilden de görülebileceği gibi sürücü yaş verileri ile ilgili olarak yanlış veri girişi ya da kaza tutanak kaydı olduğu görülmektedir. Bu aykırı ve yanlış olduğu düşünülen veriler veri ön işleme sürecinde veri setinden çıkarılmıştır.



Şekil 2.41. Trafik kazası geçiren sürücülerin yaş histogramı.

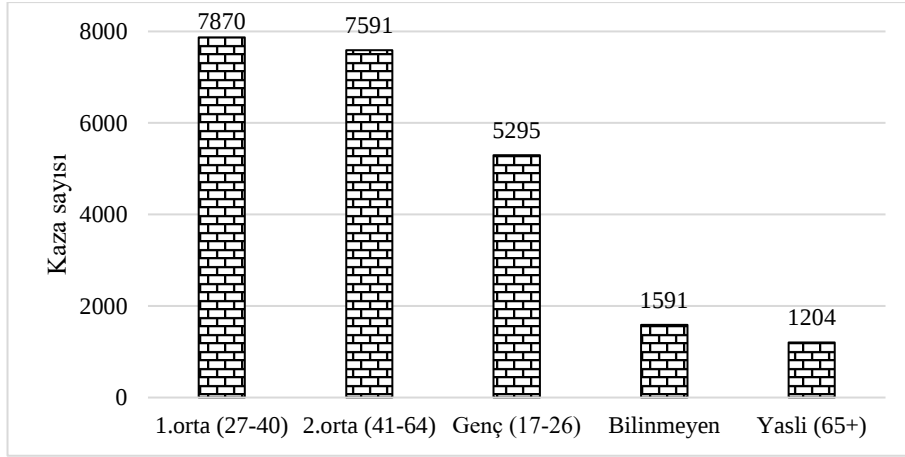
Ayrıca, T.C. İçişleri Bakanlığı Nüfus ve Vatandaşlık İşleri Genel Müdürlüğü’nden Sakarya’daki ehliyet sahiplerinin yaş bilgisi istenmiştir. Şekil 2.42’de ehliyet sahiplerinin yaş dağılımı verilmektedir.



Şekil 2.42. Sakarya'da ehliyet sahibi olanların yaş dağılımı – 2020 yılı.

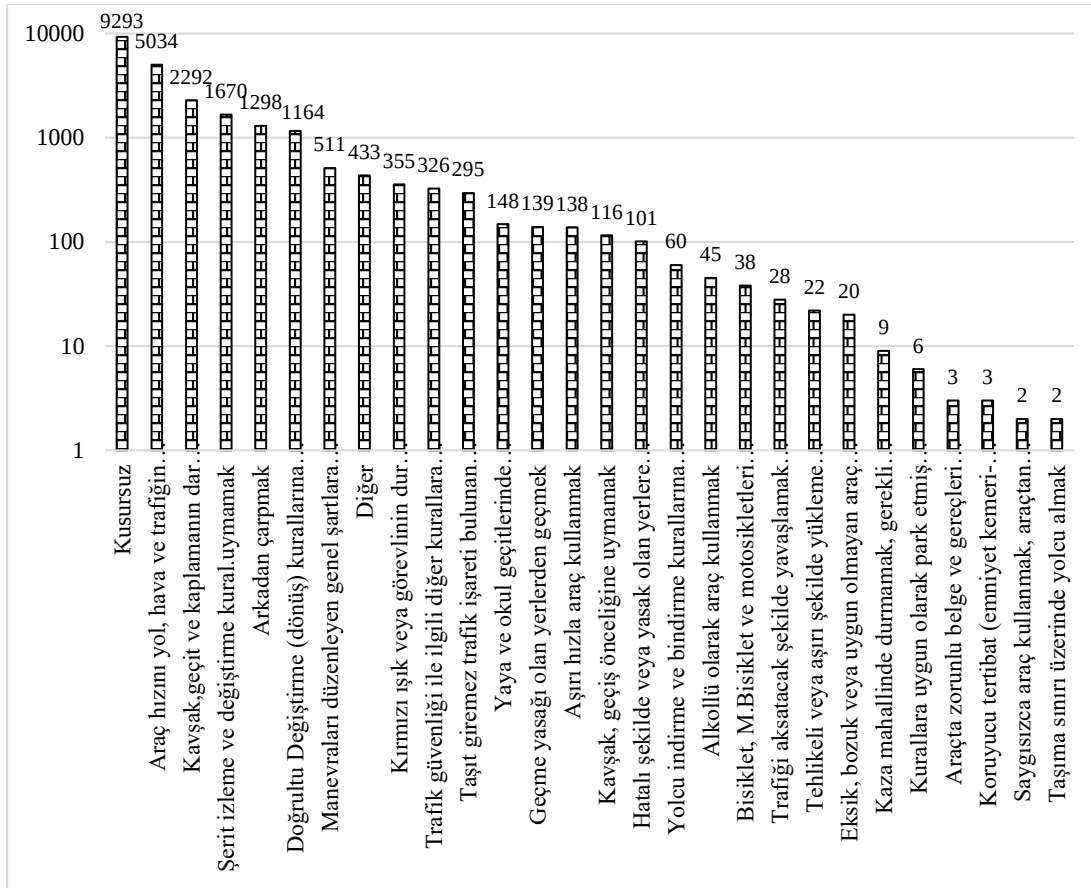
Sakarya'daki sürücülerin yaş dağılımı ile kazaya karışan sürücülerin yaş dağılımı benzer şekilde olmuştur. Sürücü sayısının fazla olduğu yaş gruplarında trafik kaza sayısının da fazla olduğu görülmektedir.

Sınıflara ayrılan sürücü yaşları ile ilgili histogram Şekil 2.43'te verilmiştir. Buna göre, sürücülerin %33'ünün 27-40 yaş aralığında, %32'sinin 41-64 yaş aralığında, %22'sinin 17-26 yaş aralığında, %5'inin 65 yaş ve üzerinde olduğu, %7'sine ait bilginin ise kayıt altında olmadığı görülmektedir.



Şekil 2.43. Trafik kazası geçiren sürücülerin yaş grubuna göre histogramı.

Şekil 2.44'te kazaya karışan sürücülerin kazaya neden olan kusurlarının sıklıkları verilmiştir.



Şekil 2.44. Trafik kazası geçiren sürücülerin kusurlarına göre histogramı.

Şekil 2.44'teki grafiğe göre, sürücü kusurları kusursuz sınıfı dışında 27 sınıfa ayrılmış durumdadır. En fazla yapılan kusurun taşıt hızını yol ve hava koşullarına uydurmama olduğu, bazı kusurlarının sayısının oldukça düşük olduğu, bazı kusurların da birbirleriyle aynı anlamı taşıdığı fark edilmiştir. İleride verinin hazırlanması bölümünde bu sınıflarla ilgili gerekli işlemler (birleştirme, temizleme gibi.) yapılmıştır.

2.4. Veri Önileme Süreci

Veri ön işleminin ilk aşamasında, EGM'den alınan trafik kaza veri setinde farklı dosyalar halinde bulunan veriler, ortak olan "kaza numaraları" kullanılarak birleştirilmiş ve tek bir dosya elde edilmiştir. Birleştirilen tek dosyada kaza, sürücü ve taşıta ait tüm bilgiler yer almaktadır.

Tablo 2.2'de EGM'den alınan dosyadaki ilk dosya olan kaza bilgileri öznitelikleri verilmektedir.

Tablo 2.2. Emniyetten alınan veri dosyalarının ilkinde yer alan öznitelikler.

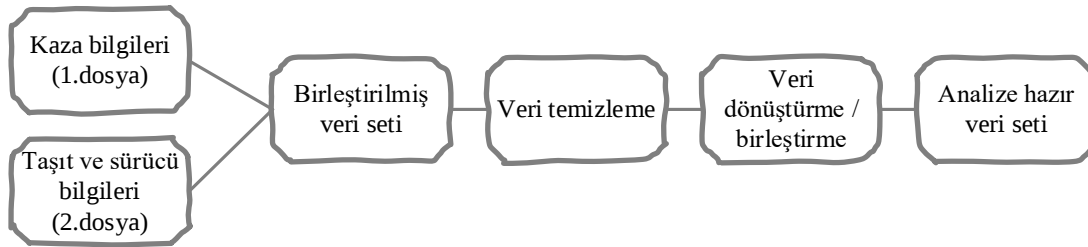
Öznitelik	Veri türü	Kayıp veri	En küçük değer	En büyük değer
Kaza yılı	Tam sayı	-	2015	2020
Kaza saati	Kategorik	-	-	-
Kaza ili plaka kodu	Tam sayı	-	54	54
X koordinat	Reel sayı	2	-	-
Y Koordinat	Reel sayı	2	-	-
Kaza Tarihi	Kategorik	-	-	-
Kaza İlçesi	Kategorik	-	-	-
Kaza yer yeri içi/dışı	Kategorik	-	-	-
Yolun tipi	Kategorik	-	-	-
Yol kaplaması	Kategorik	-	-	-
Yolun sınıfı	Kategorik	-	-	-
Yol adı kesim no	Kategorik	28	-	-
Yatay geometri	Kategorik	-	-	-
Düşey geometri	Kategorik	-	-	-
Kavşak geometrisi	Kategorik	-	-	-
Geçit geometrisi	Kategorik	-	-	-
Diğer geometrik öz.	Kategorik	-	-	-
Gün durumu	Kategorik	-	-	-
Hava durumu	Kategorik	-	-	-
Yolun yüzeyi	Kategorik	-	-	-
İlk yardım	Kategorik	-	-	-
Kaza oluş türü	Kategorik	-	-	-
Taşıt sayısı	Kategorik	-	-	-
Çarpışma yeri	Kategorik	-	-	-
Kaza yol kusuru	Kategorik	-	-	-
Yolda çalışma	Kategorik	-	-	-
Yol genişliği	Kategorik	-	-	-
Otokorkuluk	Kategorik	-	-	-
Yaya yolu	Kategorik	-	-	-
Emniyet şeridi/banket	Kategorik	-	-	-
Yol şerit çizgisi	Kategorik	-	-	-
Trafik işaret levhası	Kategorik	-	-	-
Trafik lambası	Kategorik	-	-	-
Aydınlatma	Kategorik	-	-	-
Trafik görevlisi	Kategorik	-	-	-
Görüş engel cisim	Kategorik	-	-	-
Hasar gören unsur	Kategorik	-	-	-
Yolda çalışma	Kategorik	14159	-	-
Toplam ölü sayısı	Tam sayı	-	-	7
Toplam yaralı sayısı	Tam sayı	-	-	35
Kaza şiddeti (ölümlü)	Tam sayı	-	-	1
Kaza şiddeti (yaralanmalı)	Tam sayı	-	-	1
Kaza numarası	Tam sayı	-	9	156820

Tablo 2.3'te EGM'den alınan ikinci dosyada bulunan taşıt ve sürücü bilgilerindeki öznitelikler verilmektedir.

Tablo 2.3. Emniyetten alınan veri dosyalarının ikincisinde yer alan öznitelikler.

Öznitelik	Veri türü	Kayıp veri	En küçük değer	En büyük değer
Kaza yılı	Tam sayı	-	2015	2020
Kaza ilçesi	Kategorik	-	-	-
Kaza numarası	Tam sayı	-	9	156820
Taşıt modeli	Tam sayı	2062	0	2020
Taşıt cinsi	Kategorik	415	-	-
Taşıt kaza sonrası konumu	Kategorik	-	-	-
Taşıt kullanım amacı	Kategorik	-	-	-
Taşıt hareketi	Kategorik	486	-	-
Taşıt darbe bölümü	Kategorik	-	-	-
Taşıt kusuru	Kategorik	23399	-	-
Taşıt hasar derecesi	Kategorik	-	-	-
Sürücü yaşı	Tam sayı	727	-	120
Sürücü cinsiyeti	Kategorik	104	-	-
Sürücü belgesi sahipliği	Kategorik	246	-	-
Sürücü belge sınıfı	Kategorik	2042	-	-
Sürücü öğrenim durumu	Kategorik	3538	-	-
Sürücü alkol ölçümü	Kategorik	458	-	-
Sürücü alkol sonucu	Kategorik	4449	-	-
Sürücü psiko/fiziksel durumu	Kategorik	22762	-	-
Sürücü emniyet kemeri	Kategorik	21587	-	-
Sürücü kaza sonucu	Kategorik	-	-	-
Taşıt numarası	Tam sayı	-	10	261427
Sürücü kusuru	Kategorik	166	-	-

Şekil 2.45’te bu iki dosyanın birleştirme süreci ve sonrasındaki birleştirilmiş dosyada veri setine uygulanan işlemler görülmektedir.



Şekil 2.45. Veri ön işleme süreci

2.4.1. Veri temizleme

Çalışmada kullanılan veri setinde yer alan özniteliklerden; kaza numarası, taşıt numarası, kaza koordinatları, yol adı (kesim numarası vb.) yalnızca tek bir kazaya aittir. Bu değerler, öznitelikler içerisinde sınıf oluşturamadığından veri setinden çıkarılmıştır. Bazı verilerin yanlış girildiği (bazı yaş bilgileri gibi) düşünülmektedir. Doğruluğu olası olmayan bu durumdaki veriler de veri setinden çıkarılmıştır.

Birliktelik Kuralları elde edilirken kayıp veriler önemli olmamıştır. Bu anlamda bakıldığında, “Bilinmiyor” gibi sınıfların yer aldığı kurallar, çalışma sonuçlarında değerlendirmeye alınmamıştır. Dolayısı ile ağaçta “bilinmiyor” şeklinde bir sınıfın yer alması, ağaçtaki kuralları yorumlarken belirsizliğe neden olacağı için Karar Ağacı oluşturulmadan önce eksik veriler veri setinden çıkarılarak temizlenmiştir. Bir diğer ifade ile, Karar Ağaçlarında sonuca olumsuz etkide bulunacak olan ve veri setinde bilinmeyen, eksik olan verilerin çıkarılması gerekmektedir. Aynı şekilde aykırı değerler için de bu durum geçerlidir. Birliktelik Kurallarında bir kuralın elde edilebilmesi için veri seti içerisinde belirli sıklıkta yer alması gerektiğinden (çalışmanın ilerleyen kısımlarında anlatılan eşik Destek ve Güven değerlerini aşması gerektiğinden) veri setinde yer alması sorun oluşturmazken, Karar Ağaçları oluşturulurken bu verilerin varlığı analiz sürecini olumsuz etkileyebilmektedir.

Bu durumda her öznitelik ayrı ayrı değerlendirilerek, gerekmesi durumunda veri setinden çıkarılması, birleştirilmesi, dönüştürülmesi işlemlerinden biri yapılmıştır.

Veri setinde yer alan verilerin çoğu kategorik olduğu için aykırı değerlerin belirlenmesinde çeyreklik gibi değerler kullanılamamış, literatürdeki çalışmalar incelenerek, doğru olmadığı kesin olan (yaş bilgisinin 120 ya da 3 gibi değerlere sahip olması gibi) veriler belirlenerek veri ön işleme gerçekleştirilmiştir.

2.4.2. Veri birleştirme

Ham veride yer alan özniteliklerden olan “yol tipi” ve “kavşak tipi” ayrı ayrı analize girmek durumunda kaldığında çoğu kazanın “kavşak tipi” sınıflarından olan “kavşak yok” özelliğinde olduğu görülmektedir. Bunun yerine “yol tipi” ve “kavşak tipi” özniteliklerinden alınan bilgilerle “yol bölümü” olarak yeni bir öznitelik oluşturulmuştur. Bu sayede kazanın hangi yol kesiminde (bölünmüş, tek yön, çift yön, dönel, 4 kollu kavşak, diğer kavşak çeşitleri) gerçekleştiğine dair daha detay bilgiler içeren öznitelik elde edilerek, analiz sürecinde kullanılmıştır.

Tarih verileri, kaza analizi için bir anlam ifade etmediğinden (1 gün öncesi ve 1 gün sonrası arasında büyük bir fark olmadığından) bu öznitelik içerisinden edinilen gün ve mevsim bilgileri için iki ayrı öznitelik oluşturulmuştur. Böylece verilerin günlere ve mevsimlere göre gruplanması sağlanmıştır. Analiz sonucunda yapılacak olan çıkarımlarda gün ve mevsime göre kazaların yorumlanması için böyle bir işleme gereksinim duyulmuştur.

“Gün durumu” ve “aydınlatma” özniteliklerindeki bilgilere ait içerikler bazı kaza durumları için anlamsız bir bilgi oluşturduğundan (gündüz olan bir kazada aynı zamanda aydınlatma olması gibi), “gün ışığı/aydınlatma” özneliği; “gün durumu” ve “aydınlatma” içinde yer alan bilgilerden oluşturularak (kaza anında ortamın aydınlık olup olmadığı bilgisini elde edebilmek için) analiz sürecine dahil edilmiştir.

Kaza oluş türünde devrilme-savrulma ile yoldan çıkma sınıfları, tutanağı hazırlayan trafik polisinin yorumuna açık olabileceği düşüncesi ile bu farklılığı ortadan kaldırabilmek için birleştirilmiştir.

Sürücü kusurları incelendiğinde, “manevraları düzenleyen genel şartlara uymamak” ve “şerit izleme ve değiştirme kurallarına uymamak” kusurlarının bazı yıllarda aynı oranlarda, bazı yıllarda ise birbiri yerine yazılmış olduğu fark edildiğinden, bu iki kusur “hatalı şerit değiştirme” sınıfı altında birleştirilmiştir.

Taşıt hasar derecesinde, bazı kavramların yakın anlamlar içerdiği dikkate alındığında, “ağır hasarlı” ve “hareket edemez” sınıfları birleştirilerek “ağır hasarlı”; “hafif hasarlı” ve “fonksiyonel hasar” sınıfları birleştirilerek “hafif hasarlı”; “hasarsız” ve “hasar bilgisi” olmayan sınıflar ise “hasarsız” olarak gruplandırılmıştır.

“Trafik ışığı varlığı” özneliğinin içinde yalnızca “var” ve “yok” sınıfları yer almakta olduğundan herhangi bir kazada trafik ışığı yok bilgisi, kavşak olan durum ve olmayan durumu aynı anda ifade etmektedir. Bunun yerine “kavşak tipi” özneliğinden de yararlanarak, daha fazla bilgi sunan yeni bir öznitelik olan “yoldaki trafik ışığı” türetilerek bir kazanın “ışıklı kavşak”, “ışiksiz kavşak” ve “kavşak yok” bölgelerinden hangisinde gerçekleştiğine dair bilginin analizde yer alması sağlanmıştır.

2.4.3. Veri dönüştürme

Analizde kullanılacak veriler işlenmek üzere aynı ya da çıktı olarak daha işlevsel olabilecek veri tiplerine dönüştürülmelidir. Bu çalışmadaki veri setinde yer alan “Sürücü yaşı”, “taşıt modeli”, “alkol seviyesi” gibi sürekli veri içeren öznitelikler, yeni sınıflar oluşturularak kategorik veriye dönüştürülmüştür. Veri dönüştürme ile yapılan bu işlem, çıktı olarak sürücü yaş grubu, taşıt model grubu ve alkol varlığı durumları hakkında çıktılar üretilmesine ve bu çıktıların yorumlanarak kazalar için alınabilecek önlemler konusunda bilgi sağlamıştır.

Bu ilk veri ön işleme sürecinden sonra elde edilen veri setinden iki kopya oluşturularak Birliktelik Kuralları elde etme ve Karar Ağacı modeli için ayrı ön işleme süreci uygulanmıştır.

2.5. Birliktelik Kuralları İçin Veri Ön İşleme Süreci

Birliktelik Kurallarında, veri setinde en çok bulunan öğelerin beraber oldukları yapılar elde edileceği için kullanılacak veri setindeki tüm verilerin kategorik olması sağlanmıştır. Ayrıca çalışmanın amacı sürücülerin kusurlu olduğu kazaları incelemek olduğundan, sürücülerin kusursuz olduğu kazalar veri setinden çıkarılmıştır. Öte yandan, sürücü cinsiyeti, hava durumu, taşıt cinsi, aydınlatma, emniyet kemeri/kask, sürücü belgesi, yol yüzeyi, kazanın yerleşim yeri içi/dışı olma bilgisi, yolun yatay ve düşey geometrisi, yaya geçitlerinin ve trafik işaretlerinin varlığı, yol kusuru, ilk yardım, kazaya karışan taşıt sayısı, yol bakım çalışması, bariyer, yol çizgisi, banket, trafik polisi varlığı öznitelikleri, sahip oldukları dengesiz sınıf dağılımı nedeniyle veri setinden çıkarılmıştır. Alkol seviyesi, sürücünün fiziksel ve psikolojik durumu, yaya ve yolculara ait özniteliklerde kayıp veri çok fazla olduğundan bu öznitelikler veri setinden çıkarılmıştır.

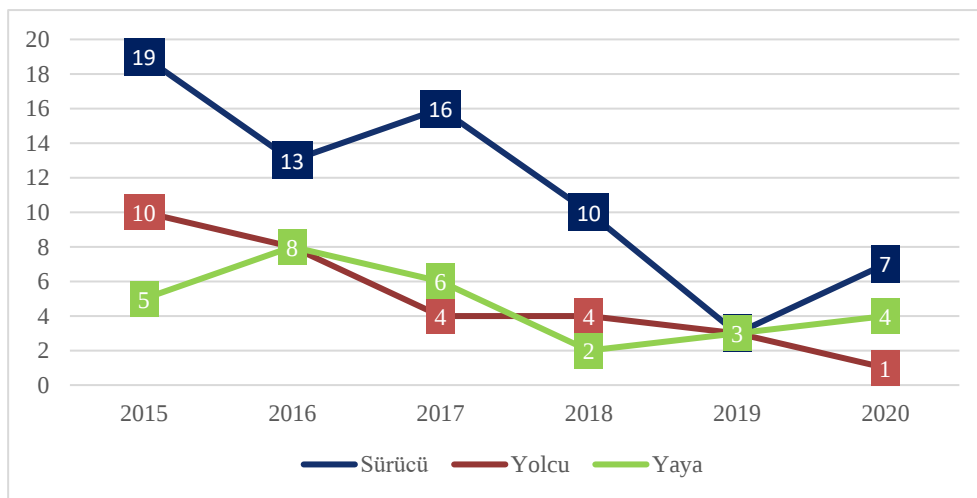
2.5.1. Birliktelik Kuralları elde edilirken kullanılan veriye ait tanımlayıcı istatistikler

Birliktelik Kurallarını elde edebilmek için uygulanan veri ön işleme sürecinden sonra Birliktelik Kurallarının elde edildiği veri seti oluşturulmuştur. Bu veri setine ait bilgilerin sunulduğu, verinin yapısını gösteren tanımlayıcı istatistikler Tablo 2.4'te verilmektedir.

Tablo 2.4. Birliktelik Kurallarının elde edildiği veri setinin tanımlayıcı istatistikleri.

Öznitelik	Sıklık	%	Öznitelik	Sıklık	%
Kaza ile ilgili etkenler			Sürücü ile ilgili etkenler		
Sürücü kaza sonucu durumu			Yaş grubu		
Ölü ya da yaralı	7050	49	Genç (18-26)	3453	24
Sağlam	7337	51	Ortayaş1 (26-40)	5035	35
Sürücü kusuru			Ortayaş2 (41-64)	4604	32
Öndeki taşıtı yakın takip	1150	8	Yaşlı (65+)	719	5
Yol vermeme	2446	17	Bilinmeyen	575	4
Hatalı dönüş	1583	11	Eğitim Durumu		
Hız	4748	33	İlkokul	5323	37
Hatalı şerit değiştirme	1583	11	Ortaokul	1439	10
Alkollü sürüş	1007	7	Lise	4172	29
Diğer	1870	13	Yükseköğretim	1583	11
Kaza oluş türü			Bilinmeyen	1870	13
Arkadan çarpma	2014	14	Yol ile ilgili etkenler		
Yandan çarpma	6474	45	Yol bölümü		
Devrilme/savrulma/yoldan çıkma	2014	14	Bölünmüş	4172	29
Karşılıklı çarpışma	1007	7	Çift yönlü yol	2590	18
Yayaya çarpma	1439	10	Tek yönlü yol	432	3
Engel/cisim ile çarpışma	1439	10	Kavşak	5323	37
Taşıt hasar derecesi			Dönel kavşak	1151	8
Ağır hasar	5899	41	Diğer yol kesimleri	719	5
Hafif hasar	6618	46	Yoldaki trafik ışığı		
Hasarsız	1870	13	Işıksız kavşak	5899	41
Çevresel etkenler			Işıklı kavşak	1151	8
Gün ışığı/aydınlatma			Kavşak yok	7337	51
Gündüz	9639	67	Taşıt ile ilgili etken		
Akşam/aydınlatma var	3309	23	Taşıt modeli		
Akşam/aydınlatma yok	1439	10	3 yaş ve altı	2877	20
Mevsim			4 ile 10 yaş arası	4604	32
Kış (aralık-şubat)	2877	20	11 yaş ve fazlası	6906	48
İlkbahar (mart-mayıs)	3309	23			
Yaz (haziran-ağustos)	3741	26			
Sonbahar (eylül-kasım)	4460	31			

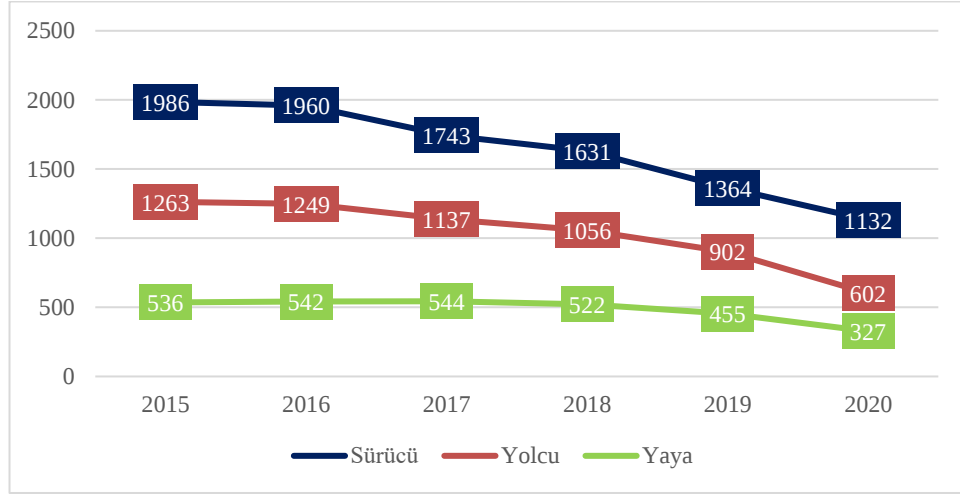
Şekil 2.46'da 2015-2020 yıllarında hayatını kaybeden kazazedelerin (sürücü, yolcu, yaya) yıllara göre değişimi sunulmaktadır.



Şekil 2.46. Trafik kazalarında hayatını kaybeden kazazede sayıları (2015-2020).

Trafik kazalarında hayatını kaybeden sürücü sayılarının 2016 yılında 2015'e göre %32 azaldığını, 2017'de ise 2016'ya göre %23 arttığı görülmektedir. 2020 yılında COVID-19 salgını etkisiyle azalan kaza sayısı neticesinde hayatını kaybeden kişi sayılarının da az olduğu yıl olmuştur.

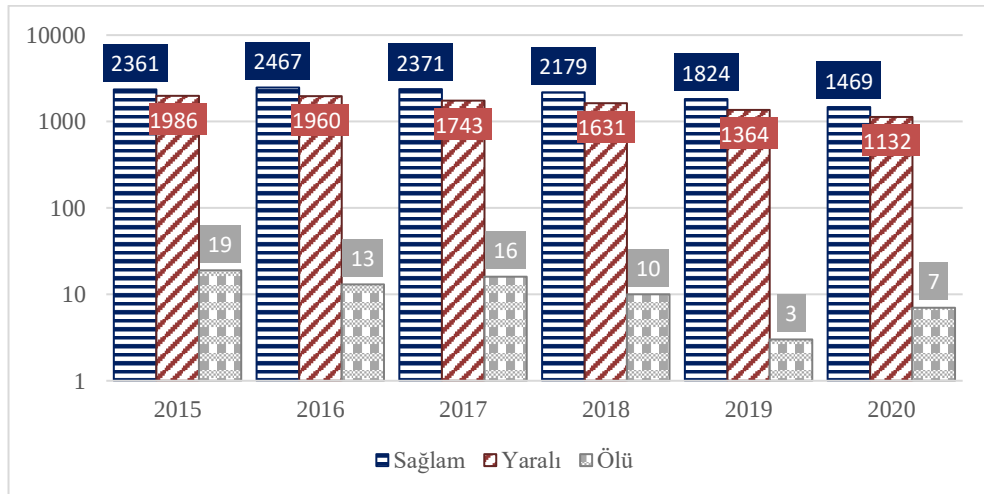
Şekil 2.47'de ise trafik kazalarında yaralanan kazazede sayılarının yıllık değişimi sunulmaktadır.



Şekil 2.47. Trafik kazalarında yaralanan kazazede sayıları (2015-2020).

Şekil 2.47'ye göre 2016 ile 2017 yılları arasındaki yaya yaralanma değerlerinin değişimi dışında her geçen yıl kazalarda yaralanma sayılarının azaldığı görülmektedir.

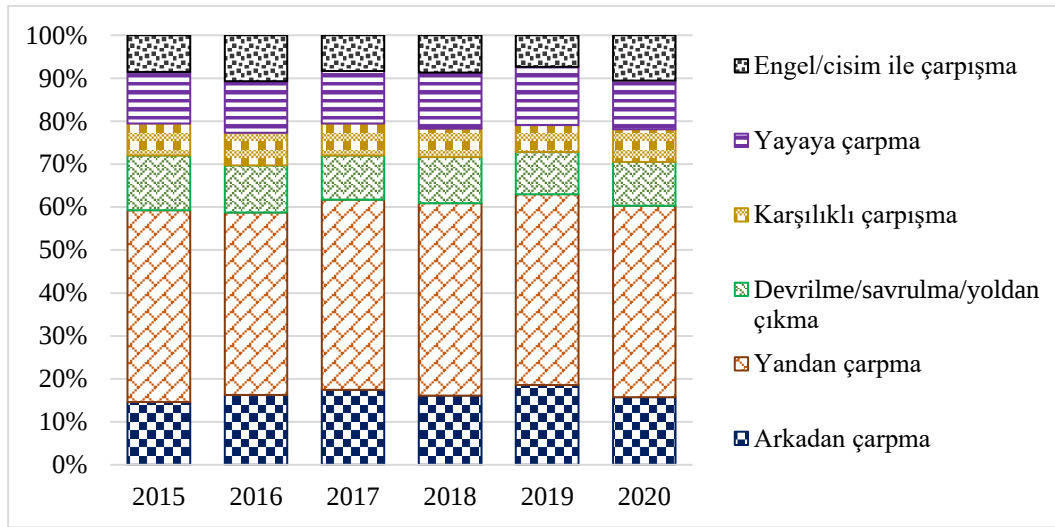
Şekil 2.48'de trafik kazalarına karışan sürücülerin kaza sonucu durumlarının dağılımları 2015-2020 yılları için verilmiştir.



Şekil 2.48. Trafik kazalarına karışan sürücülerin kaza sonucu durumları (2015-2020).

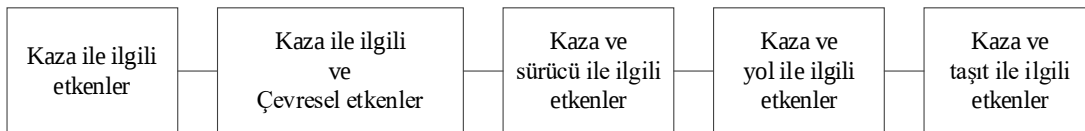
Şekil 2.48’de çalışma kapsamında incelenen yıllarda kaza sonucunda yaralanan ve sağlam sürücü sayıları hem COVID-19 salgını dışında hem de salgın etkisiyle azalmıştır. Hayatını kaybeden sürücü sayıları ise yıldan yıla bazen artmış bazen azalmıştır.

Şekil 2.49’da ise bir yıl içinde gerçekleşen kazaların kendi içindeki dağılımları verilmiştir. 2015-2020 yıllarında gerçekleşen kazaların ilgili altı yıl içindeki oranları sırasıyla %19, %20, %18, %17, %14 ve %12 olmakla beraber kaza oluş türlerinin oransal dağılımlarının çok benzer olduğu görülmektedir.



Şekil 2.49. Kaza oluş türlerinin yıllara göre dağılımı.

Sürücü, yaya veya yolcuların karıştığı ölümlü kazalar bu tez kapsamında incelenen 14,226 verinin sadece 126'sını oluşturduğundan, tüm ölümlü kazalar yaralanmalı kaza sınıfına dahil edilmiştir. Veri ön işleme süreci tamamlandıktan sonra, önemli örüntüler edinebilmek için veriler Şekil 2.50’de görüldüğü gibi beş ana gruba ayrılmıştır. Birinci grupta yalnızca Tablo 2.4’te yer alan kaza ile ilgili etkenler, diğer dört grupta ise bu etkenlerle beraber ayrı ayrı gruplanan çevre, sürücü, yol ve taşıt ile ilgili etkenler yer almaktadır.



Şekil 2.50. Önemli Birliktelik Kuralları elde edebilmek için oluşturulan beş ana grup.

Birliktelik Kurallarını elde ederken apriori algoritması kullanılmıştır. Bu algoritma, veri setini yinelemeli olarak tarayarak veri seti içerisinde önemli rol oynayan, referans

destek değeri dikkate alındığında frekans değeri yüksek olan öğeleri bulmaktadır (Agrawal vd., 1993).

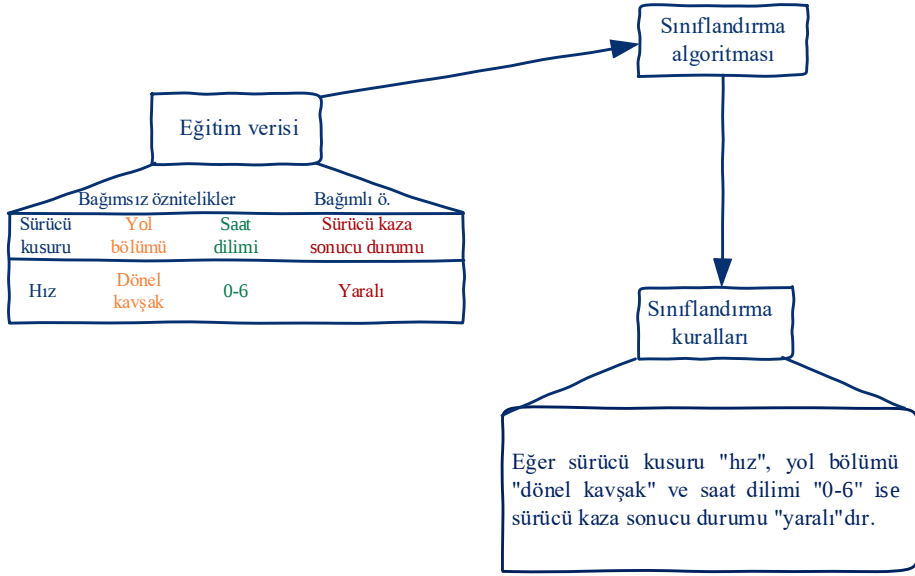
2.6. Karar Ağaçları İçin Veri Ön İşleme Süreci

Birliktelik Kurallarında “diğer” ya da az geçen başka sınıflar veri setinden çıkarılmamıştır. Çıkarılmamasının nedeni, bu sınıflarda yer alan verilerden kurallar elde edilse bile çalışma çıktısı olarak verilmeyerek, çalışma içeriği açısından önemli olmayan etkenlerin çalışma sonuçlarında yer almamasının sağlanmasıdır.

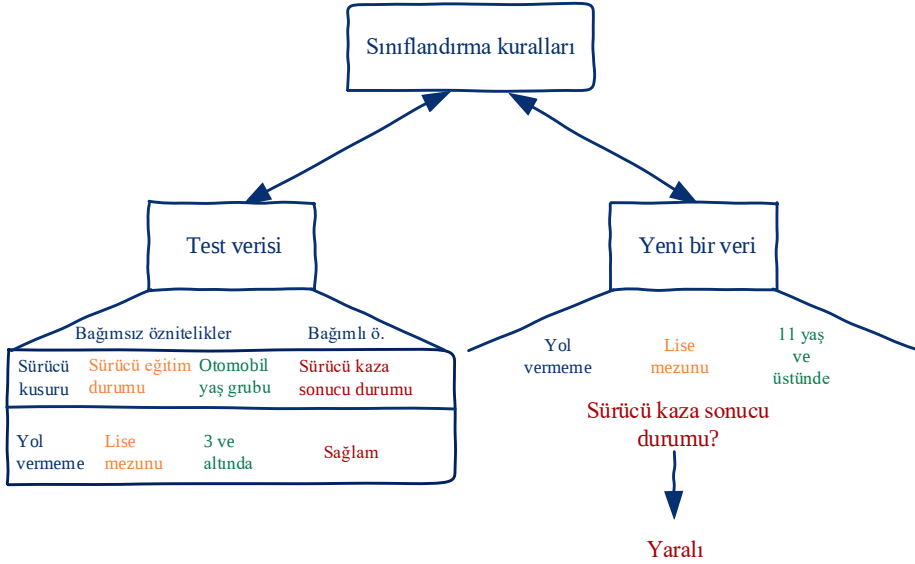
Karar Ağacında ise bu sınıflar ağacın belli aşamalarında yer alıp, sonuç olarak ağaçtan üretilecek olan bilginin bir parçası olmak durumunda olacağı için, analize başlamadan önce “diğer” gibi sınıflar veri setinden çıkarılmıştır. “Diğer” gibi sınıflarla elde edilecek olan bilginin kullanılması mümkün olmadığından bu işlem yapılmıştır.

2.6.1. Karar Ağacı elde edilirken kullanılan veriye ait tanımlayıcı istatistikler

Karar Ağacı analizinde sınıflandırma yapılacağı için Birliktelik Kuralları için ön işleme sürecinden geçen veri setinin tekrar düzenlenmesi gerekmiştir. Sınıflandırma, önemli veri sınıflarını tanımlayan modelleri çıkaran bir veri analizi biçimidir. Sınıflandırıcı olarak adlandırılan bu modeller, kategorik (ayrık/discrete, sırasız/unordered) sınıf etiketlerini tahmin etmektedir (Han vd., 2012). Bu çalışmada, bir trafik kazası sonucunda sürücünün sağlam, yaralı veya hayatını kaybetmiş olarak kategorize edilmesi için bir sınıflandırma modeli kurulmuştur. Öğrenme (sınıflandırma modelinin oluşturulduğu) ve sınıflandırma adımından (modelin verilen veriler için sınıf etiketlerini tahmin etmek için kullanıldığı) oluşan iki aşamalı bir süreç olan sınıflandırma, Şekil 2.51’de trafik kaza verileri kullanılarak gösterilmiştir.



(a)



(b)

Şekil 2.51. Veri sınıflandırma süreci: (a) Öğrenme: Eğitim verileri bir sınıflandırma algoritması tarafından analiz edilmektedir. (b) Sınıflandırma: Test verileri, sınıflandırma kurallarının doğruluğunu tahmin etmek için kullanılmaktadır, Han ve ark. (2012)'den uyarlanmıştır.

Tablo 2.5'te Karar Ağacı analizinde kullanılan veri setine ait tanımlayıcı istatistikler verilmiştir. Sınıflandırılan özneliğin sürücü kaza sonucu olduğu bu tabloda görülmektedir.

Tablo 2.5. Karar Ağacı için hazırlanan veri seti

Öznitelik	Sıklık	%	Öznitelik	Sıklık	%
Bağımlı değişken			Bağımsız değişkenler		
Sürücü kaza sonucu durumu			Sürücü alkol sonucu		
Yaralı	1849	0,59	Alkolsüz	2198	0,86
Sağlam	1297	0,41	Alkollü	344	0,14
Ölü	6	0,002	Sürücü belge sahipliği		
Bağımsız değişkenler			Ehliyet sahibi	2957	0,94
Sürücü kusuru			Ehliyeti yok	195	0,06
Hız	1529	0,49	Saat dilimi		
Yakın takip	341	0,11	00:00-06:00	467	0,15
Hatalı şerit değiştirme	539	0,17	06:00-12:00	621	0,20
Yol vermeme	491	0,16	12:00-18:00	1108	0,35
Hatalı dönüş	234	0,07	18:00-24:00	956	0,30
Alkollü sürüş	18	0,01	Yol bölümü		
Kaza oluş türü			Bölünmüş	1320	0,42
Yandan	1026	0,33	İki yönlü yol	528	0,17
Devrilme/savrulma/takla	963	0,31	Tek yönlü yol	84	0,03
Engele/cisme	543	0,17	3-4 kollu kavşak	910	0,29
Arkadan çarpma	402	0,13	Dönel kavşak	297	0,10
Karşılıklı çarpışma	163	0,05	Gün ışığı/aydınlatma		
Yolun yüzeyi			Gündüz	1919	0,61
Kuru	2428	0,77	Akşam/aydınlatma var	810	0,26
Kaygan	724	0,23	Akşam/aydınlatma yok	423	0,13
Kazaya karışan otomobil sayısı			Taşıt modeli		
Tek otomobil	1473	0,48	3 yaş ve altı	471	0,15
İki otomobil	1528	0,47	4 ile 10 yaş arası	957	0,30
Çok otomobil	151	0,05	11 yaş ve fazlası	1721	0,55
Sürücü yaş grubu			Haftanın günü		
Genç (18-25)	755	0,24	Pazartesi	395	0,12
Orta yaş 1 (26-40)	1252	0,40	Salı	389	0,12
Orta yaş 2 (41-64)	967	0,31	Çarşamba	413	0,13
Yaşlı (65+)	178	0,05	Perşembe	423	0,13
Sürücü öğrenim durumu			Cuma	426	0,00
İlkokul mezunu	1005	0,36	Cumartesi	503	0,16
Ortaokul mezunu	321	0,12	Pazar	603	0,19
Lise mezunu	1012	0,37	Düşey geçki geometrisi		
Yükseköğretim mezunu	422	0,15	Eğimli	2686	0,85
Sürücü cinsiyet			Eğimsiz	466	0,15
Erkek	2798	0,89	Yolun tipi		
Kadın	354	0,11	Bölünmüş yol	2042	0,65
Kaza mevsimi			İki yönlü yol	951	0,3
Sonbahar	768	0,24	Tek yönlü yol	132	0,04
Kış	679	0,21	Diğer	27	0,009
İlkbahar	721	0,23	Yerleşim yeri içi/dışı		
Yaz	984	0,31	Yerleşim yeri içi	2384	0,76
Trafik ışığı varlığı			Yerleşim yeri dışı	768	0,24
Var	345	0,11	Yatay geçki geometrisi		
Yok	2807	0,89	Düz yol	2695	0,85
			Dönemeç	457	0,14

Sürücü kusurları ham veri setinde 27 sınıfa ayrılmış durumdayken, Karar Ağacı analizi için en fazla gerçekleşen ve Birliktelik Kurallarında da öne çıkan, Tablo 2.5'te görülen sınıflar bırakılmıştır.

3. VERİ ANALİZLERİ

İlk aşamada, trafik kazalarının mevcut durumunu belirlemek, değişkenler arasındaki örüntüleri ortaya çıkarabilmek için öncelikle veri madenciliği yöntemlerinden olan Birliktelik Kuralları ile veri analizi yapılmıştır.

İkinci aşamada ise öngörü modeli kurmak için, yine veri madenciliği yöntemlerinden olan, makine öğrenmesi algoritmalarını (harzemli yolu) da kullanan Karar Ağaçları yönteminden yararlanılmıştır. Bu aşamada öngörü modeli kurarken tüm taşıtlar için model kurulduğunda sürücülere ait değişkenler yerine sürücü kaza sonucunu daha çok etkileyen (bu konuda diğer özniteliklere baskın gelen) taşıt cinsi, taşıt sayısı, yerleşim yeri içi/dışı öznitelikleri ile Karar Ağacı ortaya çıkmıştır. Ancak bu tez çalışmasında sürücü kusurlarından kaynaklanan kazalar ile ilgili ve bu kazalarda sürücü özelliklerine bağlı durumlar analiz edilmek istendiğinden, baskın olan değişkenler modelden çıkarılarak sürücülere ait özelliklerin Karar Ağacında yer alması sağlanmıştır.

3.1. Birliktelik Kuralları ile Veri Analizi

Birliktelik Kurallarını elde etmek için çalışmada kullanılan veri madenciliği programları; WEKA, Rapid Miner, Orange ve Rstudio'dur.

WEKA, 1993 yılında, Yeni Zelanda'da Waikato Üniversite'sinde geliştirilmiş, Waikato Environment for Knowledge Analysis kelimelerinin baş harflerinden oluşan aynı zamanda Yeni Zelanda'da yaşayan bir kuş ismi ile isimlendirilmiş, bir makine öğrenmesi programıdır. Açık kaynak kodlu, genel kamu lisansına sahip olan programın dili 1997'den beri JAVA'dır (Eibe Frank, Mark A. Hall, and Ian H. Witten, 2016). Bu tez çalışmasının başlangıcında, Weka programı kullanılarak CRISP-DM sürecinde sözü geçen verinin anlaşılması süreci üzerinde durulmuş, her öznitelik için grafikler elde edilerek verinin yapısı anlaşılmış, ayrıca veri ön işleme süreçlerinde de bu programdan faydalanılmıştır.

Rapid Miner, 2001 yılında Almanya'da Dortmund Üniversitesi'nde geliştirilmeye başlanan, dünyada en çok kullanılan açık kaynak kodlu veri madenciliği

programlarından biridir. Fizik, mühendislik, ilaç, kimya, dil ve sosyal bilimler başta olmak üzere çok çeşitli bilim alanlarında kullanılmaktadır (Sebastian Land & Simon Fischer, 2012). Rapid Miner programının lisansı süresince Karar Ağacı modelleri bu programda oluşturulmuştur.

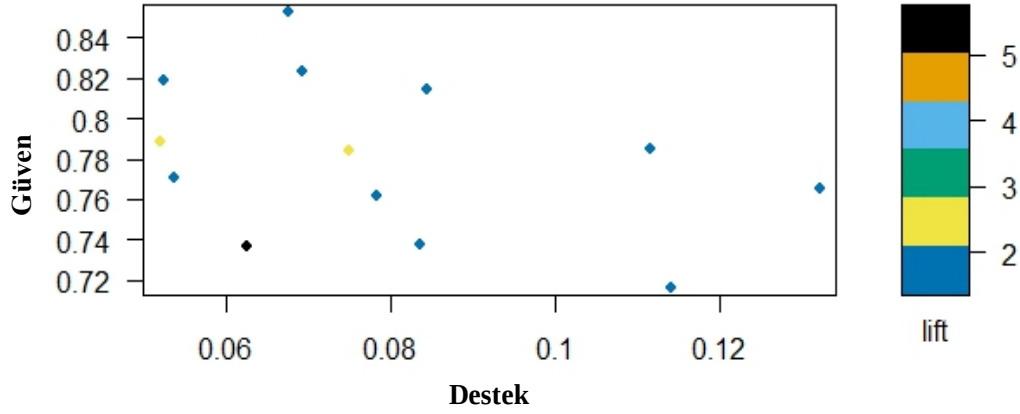
Orange, ilk sürümü 1996 yılında Slovenya’da Ljubljana Üniversitesi’nde ortaya çıkan, veri görselleştirme, makine öğrenmesi, veri madenciliği ve veri analizi için kullanılan bir yazılım paketi olup Python, C gibi programlama dillerinde yazılmaktadır (“Orange (Software)”, 2024). Bu çalışmada, bu program kullanılarak 2’den fazla öznitelik arasındaki ilişkilerin grafikleri kolaylıkla elde edilebilmiş, verinin bu detayda anlaşılması sağlanmıştır.

RStudio, R program dilinin (1993) eklentisi olarak 2011 yılında ortaya çıkan, istatistiksel hesaplama ve grafikler için kullanılan bir yazılımdır (“RStudio”, 2024). RStudio diğer programlardan farklı olarak, apriori ile elde edilen Birliktelik Kurallarına ait Destek, Güven ve Lift değerlerinin saçılım grafiğini üretebildiğinden, çalışmada Birliktelik Kuralları elde ederken bu programdan alınan sonuçlara yer verilmiştir.

Analizlerde sürücülerin kaza sonucunda karşı karşıya kaldıkları durumlar değerlendirilmiştir. Veri setini oluşturan kazaların sürücü kusurlarından kaynaklı yaralanmalı veya ölümlü kazalara ait raporlardan oluştuğu dikkate alındığında, Birliktelik Kurallarında sürücünün sağlam olduğu çıktısına ulaşılmış olsa bile, bu kazada mutlaka başka bir sürücü, yolcu ya da yaya yaralanmış veya hayatını kaybetmiştir.

Birliktelik Kurallarını elde etmeden önce Destek ve Güven için eşik değerlerinin belirlenmesi gerekmektedir. Literatürde bu değerlerin 0,001 ve 0,05 aralığında alındığı görülmüştür (Das vd., 2018; Geurts vd., 2005; Hong vd., 2020; Kong vd., 2020; Montella, 2011; Montella vd., 2012; Pande & Abdel-Aty, 2009; Xu vd., 2018). Bu çalışmada çeşitli denemeler yapılarak kurallar elde edilmiş, belirli bir değerden sonra kuralların yinelenmediği belirlendiğinden, Destek için 0,05, Güven için 0,70 değerleri eşik değerler olarak seçilmiştir. Lift değerinin 1’den büyük olması, öznitelikler arasında karşılıklı ilişkinin olduğu anlamına gelmekte olup bu değer 1’den uzaklaştıkça ilişkinin güçlü bir içeriğe sahip olma özelliği artmaktadır (Yu vd., 2019).

Sadece kazaya ait özniteliklerin yer aldığı veri setinde bu eşik değerler ile 13 tane Birliktelik Kuralı elde edilmiştir. Şekil 3.1’de bu kurallara ait Destek, Güven ve lift değerlerine ait saçılım grafiği verilmiş olup elde edilen kurallardan 9 tanesi lift değerlerine göre büyükten küçüğe doğru sıralanarak Tablo 3.1’de sunulmuştur.



Şekil 3.1. Kaza ile ilgili etkenler arasında bulunan 13 Birliktelik Kuralına ait saçılım grafiği.

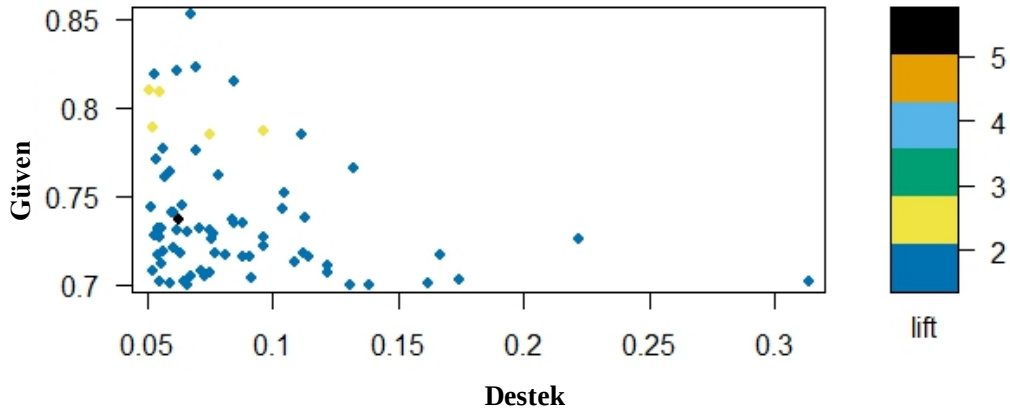
Tablo 3.1. Kaza ile ilgili etkenler arasında lift değerlerine göre sıralanan 9 önemli kural.

Kural	Öncül	Ardıl	Destek	Güven	Lift
1	Yakın takip	Arkadan çarpma	0,06	0,74	5,38
2	Yol vermeme, Hafif hasar	Yandan çarpma	0,07	0,78	1,76
3	Sağlam, Yol vermeme, Hafif hasar	Yandan çarpma	0,05	0,79	1,76
4	Hız, D/s/yoldan çıkma	Yaralı/Ölü	0,07	0,85	1,74
5	D/s/yoldan çıkma, Ağır hasar	Yaralı/Ölü	0,07	0,82	1,68
6	Hatalı dönüş, Hafif hasar	Sağlam	0,05	0,82	1,62
7	Yayaya çarpma	Sağlam	0,08	0,81	1,61
8	Hız, Ağır hasar	Yaralı/Ölü	0,11	0,72	1,46
9	Hatalı dönüş	Sağlam	0,08	0,74	1,46

Tablo 3.1’e göre, öndeki taşıtı yakın takip etme kusuru, tüm kazalar içerisinde yüksek bir paya sahip olduğundan, bu özniteliğe ait Destek değeri, eşik değeri olarak yukarıda belirlenmiş olan Destek değerinden büyük bir değerde edilmiştir. Bu durum bu kusura ait özniteliğin Birliktelik Kuralları arasında önemli bir yer edinmesi sonucunu doğurmuştur. Kurallar arasında en belirginini, taşıtları yakın takip etme kusuru yapan sürücülerin arkadan çarpma kazasına karışması arasındaki ilişkidir. Bu bulgu, sürücülerin öndeki taşıtla arasındaki mesafeyi doğru ayarlamadığı ve sonucunda ani fren vb. durumlarda arkadan çarpma şeklinde kaza yaptıklarını, bu türde kazaların tüm kazalar içerisinde dikkat çekici olduklarını ortaya çıkarmaktadır. İkinci kural, yol vermeme kusuru yapan ve taşıtları az hasar alan sürücülerin, yandan çarpma türündeki kazalara karışmaları olarak ortaya çıkmıştır. Taşıt hızını yol ve hava koşullarına

uydurmayan ve devrilme/savrulma/yoldan çıkma türünde kaza yapan sürücülerin kaza sonucunda yaralandığı ya da öldüğü 3. kural olarak ortaya çıkmıştır. Lift değeri 1'in üzerinde olan başka kurallar olsa da bunlar yukarıdaki tabloda yer alan kuralların içinde yer alan elemanlardan oluşan kümelerin kombinasyonları oldukları için tabloda verilerek tekrar edilmemiştir. Sürücü kusurunun yakın takip, kaza oluş türünün arkadan çarpma olduğu (lift değeri 5,38 olan kural) kural Şekil 3.2, Şekil 3.3, Şekil 3.4 ve Şekil 3.5'te de yer almıştır. Aşağıda Birliktelik Kuralları ile ilgili verilen diğer tablolar için de aynı durum geçerlidir.

Şekil 3.2'de kaza ve çevresel etkenler ile ilgili özniteliklerin bulunduğu veri seti ile elde edilen Birliktelik Kurallarındaki Destek, Güven ve Lift değerlerine ait saçılım grafiği verilmiştir.



Şekil 3.2. Kaza ve çevresel etkenler ile bulunan 72 Birliktelik Kuralına ait saçılım grafiği.

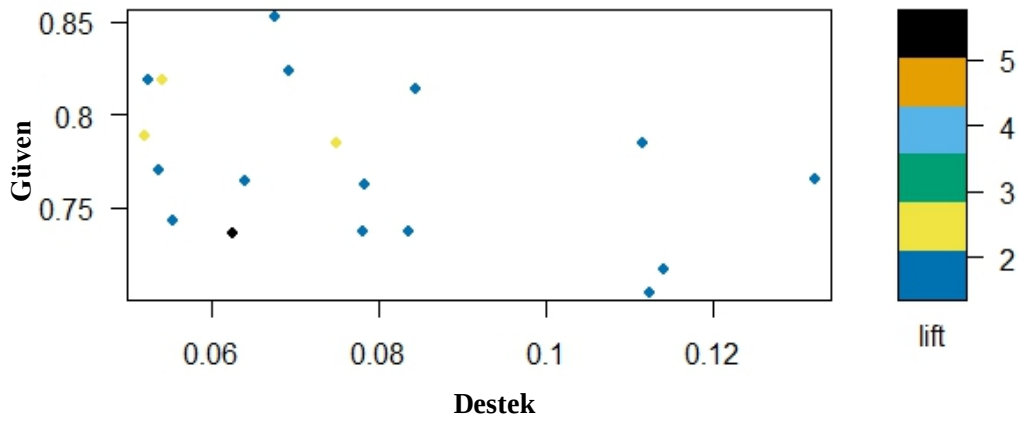
Tablo 3.2'de ise Şekil 3.2'de yer alan 72 kuraldan Lift değerine göre en önemli 10 Birliktelik Kuralı yer almaktadır.

Tablo 3.2. Kaza ve çevresel etkenler ile ilgili etkenler arasında lift değerlerine göre sıralanan 10 önemli kural.

Kural	Öncül	Ardıl	Destek	Güven	Lift
1	Gündüz, Yol vermeme, Hafif hasar	Yandan çarpma	0,05	0,81	1,81
2	Gündüz, Yol vermeme	Yandan çarpma	0,1	0,79	1,76
3	Gündüz, Yayaya çarpma	Sağlam	0,06	0,82	1,62
4	Gündüz, Devrilme/savrulma/yoldan çıkma	Yaralı	0,07	0,78	1,58
5	Gündüz, Hız, Ağır hasar	Yaralı	0,08	0,72	1,46
6	Yaz, Sağlam, Yandan çarpma	Gündüz	0,06	0,76	1,15
7	İlkbahar, Sağlam	Gündüz	0,09	0,73	1,12
8	Yaz, Hafif hasar	Gündüz	0,1	0,74	1,11
9	İlkbahar, Hız	Gündüz	0,05	0,73	1,11
10	Yaz, Hız	Gündüz	0,07	0,73	1,1

Tablo 3.2'den görülebileceği gibi en yüksek Destek değeri olan 1,81, gündüz, yol vermeme kusuru, taşıtın hafif hasarlı olması ile yandan çarpma türünde kaza gerçekleşmesi arasında büyük bir ilişki bulunmaktadır. Bir diğer kural ise gündüz ve yayaya çarpma türünde meydana gelen kazaların sürücünün yaralanmadığı kazalar olduğu gerçeğini ortaya koymaktadır. Öte yandan, gündüz ve devrilme/savrulma/yoldan çıkma (Dsy) türünde kazalar sıklıkla sürücünün yaralanmasına yol açmaktadır. Gündüz, yol vermeme kusuru, taşıtın hafif hasarlı olması ile yandan çarpma türünde kazaların gerçekleşmesi arasında da karşılıklı güçlü bir ilişki olduğu tablo verilerinden görülmektedir.

Şekil 3.3'te sürücü ve kaza ile ilgili etkenlerin bulunduğu veri seti ile Birliktelik Kuralları elde edilmiştir. Elde edilen 18 Birliktelik Kuralının Destek, Güven ve Lift değerlerinin saçılım grafiği verilmiştir.



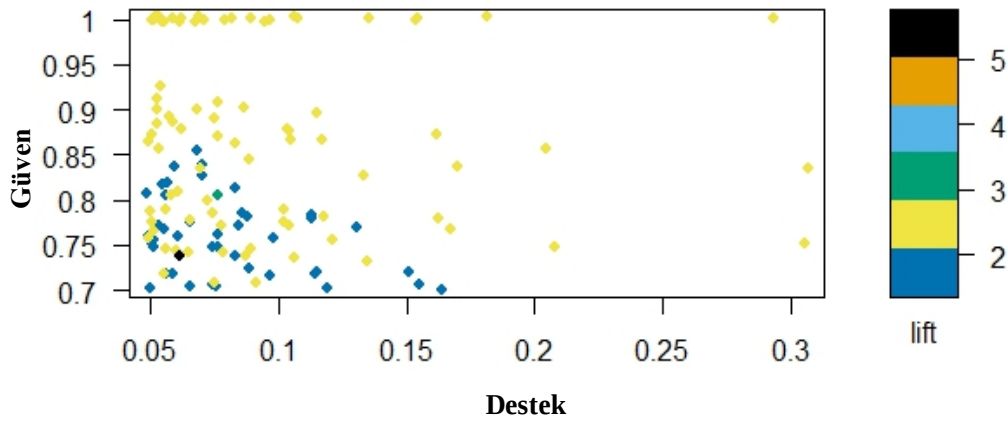
Şekil 3.3. Kaza ve sürücü ile ilgili etkenler ile bulunan 18 Birliktelik Kuralına ait saçılım grafiği.

Tablo 3.3'te bu 18 kuralın içinde en önemli olan 4 tanesi verilmiştir. İlk kural, ilköğretim mezunu olan sürücülerin yol vermeme kusuru işleyerek yandan çarpma türünde kazalara neden olmalarıdır. İkinci kural ise orta yaş 1 (26-40) grubunda yer alan sürücülerin yandan çarpma türünde yaptığı ve taşıtlarının hafif hasarlı olduğu kazalarda bu sürücülerin yaralanmaması olgusuna işaret etmektedir. 3. kural ise genç (18-25), taşıtları ağır hasar alan sürücülerin kazalarda yaralandığı sonucunu göstermektedir. 4.kural, 2. kural ile çok benzer olmakla birlikte bu kural, orta yaş 2 (41-64) grubunda yer alan sürücülerin de, orta yaş 1 (26-40) grubu gibi taşıtlarının hafif hasar aldığı durumlarda genellikle yandan çarpma türündeki kazalara neden olduğuna işaret etmektedir.

Tablo 3.3. Kaza ve sürücü ile ilgili etkenler arasında lift değerlerine göre sıralanan 4 önemli kural.

Kural	Öncül	Ardıl	Destek	Güven	Lift
1	İlkokul, Yol vermeme	Yandan çarpma	0,05	0,82	1,83
2	Ortayaş 1, Yandan çarpma, Hafif hasar	Sağlam	0,06	0,76	1,51
3	Genç, Ağır hasar	Yaralı	0,08	0,74	1,5
4	Ortayaş 2, Yandan çarpma, Hafif hasar	Sağlam	0,05	0,74	1,47

Yol ve kaza ile ilgili etkenlerin bulunduğu veri setinden Birliktelik Kuralları elde edilmiştir. 128 tane Birliktelik Kuralı elde edilmiş, bu kurallara ait Destek, Güven ve Lift değerlerine ait saçılım grafiği Şekil 3.4’te verilmiştir.



Şekil 3.4. Kaza ve yol ile ilgili etkenler ile bulunan 128 Birliktelik Kuralına ait saçılım grafiği.

Tablo 3.4’te yol ve kaza ile ilgili etkenler arasında bulunan en güçlü 11 ilişki yer almaktadır. Birbiri içinde tekrar eden kurallara yer vermemek için tablo 11 kuralla sınırlı tutulmuştur.

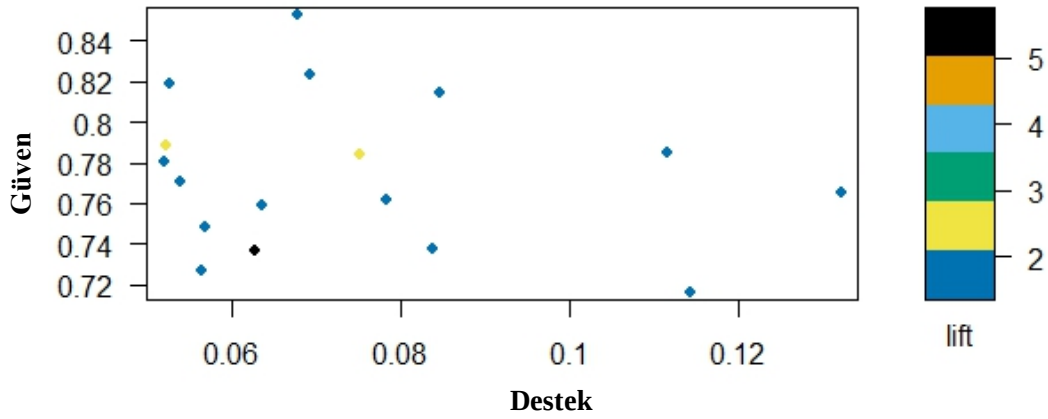
Tablo 3.4. Kaza ve yol ile ilgili etkenler arasında lift değerlerine göre sıralanan 11 önemli kural.

Kural	Öncül	Ardıl	Destek	Güven	Lift
1	Kavşak, Yol vermeme, Hafif hasar	İşıksız kavşak	0,05	0,92	2,27
2	Kavşak, Yol vermeme, Yandan çarpma	İşıksız kavşak	0,08	0,91	2,24
3	Kavşak, Hız, Yandan çarpma	İşıksız kavşak	0,07	0,9	2,21
4	Kavşak, Sağlam, Yandan çarpma, Hafif hasar	İşıksız kavşak	0,07	0,89	2,19
5	Kavşak, Yaralı, Hız	İşıksız kavşak	0,05	0,88	2,17
6	İşıksız kavşak, Yaralı	Kavşak	0,13	0,73	1,99
7	Bölünmüş yol, Cisme/engele çarpma	Kavşak yok	0,05	1	1,96
8	Bölünmüş yol, Hatalı şerit d,	Kavşak yok	0,05	1	1,96
9	Kavşak, İşıksız kavşak, Yol vermeme	Yandan çarpma	0,08	0,87	1,95
10	Kavşak, Ağır hasar	İşıksız kavşak	0,11	0,77	1,9
11	Dönel kavşak	Yandan çarpma	0,06	0,72	1,61

Tablo 3.4’e göre, 1. kuralda kavşakta yol vermeme kusuru sonucunda, taşıtların az hasarlandığı ve kazaların ışıksız kavşaklarda gerçekleştiği belirlenmiştir. 2.kuralda, 1.

kuraldan farklı olarak kaza türünün yandan çarpma olduğu belirlenmiştir. 3 ve 5. kurallarda kavşaklarda taşıt hızını yola uydurmama kusurunun mevcudiyeti belirlenmiş, 3. kuralda bu durumun yandan çarpma türünde kazaları, 5. kuralda ise bu tip kazalarda sürücünün yaralandığı ortaya çıkmıştır. 7. ve 8. kuralda ise, kavşak olmayan bölgelerdeki bölünmüş yollarda, engele/cisme çarpma ve hatalı şerit değiştirme kusurlarının yapıldığı belirlenmiştir.

Şekil 3.5'te taşıt ve kaza ile ilgili etkenlerin yer aldığı veri setinden elde edilen Birliktelik Kurallarının Destek, Güven ve Lift değerlerinin saçılım grafiği verilmiştir. Belirlenen en güçlü 4 kural Tablo 3.5'te sunulmuştur.



Şekil 3.5. Kaza ve taşıt ile ilgili etkenler ile bulunan 17 Birliktelik Kuralına ait saçılım grafiği.

Tablo 3.5'e göre, 11 yaş ya da daha yaşlı taşıt kullanan sürücülerin yol vermeme kusurları sonucunda yandan çarpma kazalarına neden olmaları durumu en fazla gerçekleşen durumlardan biri olmuştur. Aynı şekilde bu modellerde taşıt kullanan sürücülerin devrilme/savrulma/yoldan çıkma türünde kazalara karıştıkları ve kazalar sonucunda yaralandıkları belirlenmiştir. Bu kuralı destekleyen 3.kuralda da, yine aynı model taşıt kullananların taşıt hızını yol koşullarına uydurmama kusuru işledikleri, taşıtlarının ağır hasar aldığı ve sürücülerin yaralandığı görülmektedir. Son kuralda ise 4-10 yaş arası taşıt kullanan sürücülerin yandan çarpma türünde neden oldukları kazalarda taşıtlarının hafif hasar alması ile sürücülerin yaralanmaması arasında güçlü ilişki bulunmuştur.

Tablo 3.5. Kaza ve taşıt ile ilgili etkenler arasında lift değerlerine göre sıralanan 4 önemli kural.

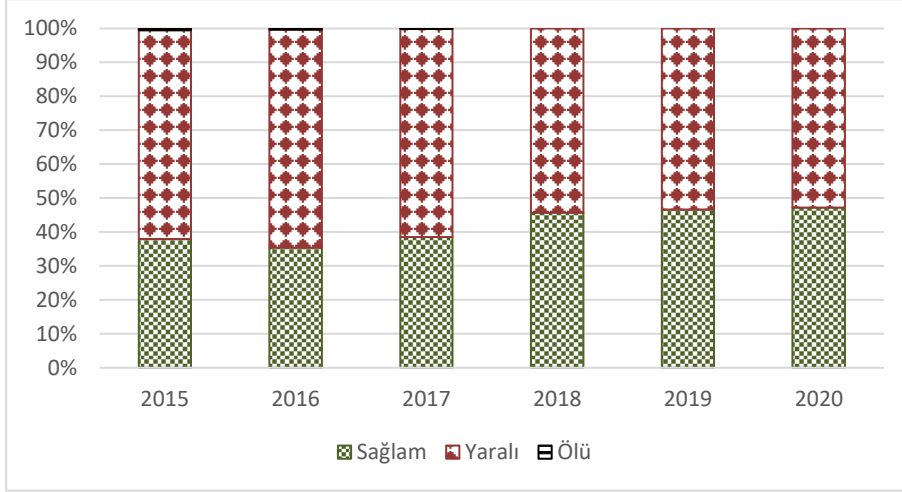
Kural	Öncül	Ardıl	Destek	Güven	Lift
1	11+, Yol vermeme	Yandan çarpma	0,06	0,76	1,7
2	11+, D/s/yoldan çıkma	Yaralı	0,05	0,78	1,59
3	11+, Hız, Ağır hasar	Yaralı	0,05	0,73	1,48
4	4-10, Yandan çarpma, Hafif hasar	Sağlam	0,06	0,75	1,48

Yukarıdaki tablodaki kurallardan da görülebileceği gibi, eski taşıtlar yeni taşıtlara göre göreceli daha ciddi kazalara karışabilmektedir. Kural 1, eski taşıt sürücülerinin daha dikkatsiz sürüş tutumlarına sahip olduklarını göstermektedir. Bu durum, eski taşıt sürücülerinin bir kazaya karıştıklarında taşıt hasarının maliyetinin nispeten düşük olacağı algısıyla ilgili olabilir. Öte yandan, yüksek hızda eski taşıtların karıştığı kazalar yaralanmalı ağır kazalarla sonuçlanırken, daha yeni model taşıtlar (4-10 yaş arasında), sürücülerin yaralanmadığı küçük kazalara karışmaktadırlar. 3 yaş ve altında kalan taşıtlar ile ilgili kurallar, veri setinde eşik değerlerin üstüne çıkacak kadar yer etmediğinden bulunamamıştır.

3.2. Karar Ağaçları ile Veri Analizi

Literatürde bazı çalışmalarda yaralanmanın şiddeti de bilindiğinden kaza şiddeti daha fazla sınıf içerebilmekte ancak bu çalışmada kullanılan veri setinde ise yalnızca ölümlü ve yaralanmalı kazalar şeklinde iki sınıf yer almaktadır. Yaralanmalı kaza sayısının baskınlığından dolayı (%99,1), sınıflandırma modeli kaza şiddeti üzerinden kurulmamıştır. Kaza şiddeti, kazazedelerin yaralanma şiddeti üzerinden de incelenmektedir. Bu teze temel oluşturan veri setinde kaza sonucu yolcu veya yolcuların durumları ile ilgili bilgilendirici kayıtların az olması, ayrıca yayaya çarpma kaza sayısının da veri setinde düşük oranda yer alması dikkate alındığında, trafik kazalarında en fazla zarar görenlerin de sürücü olduğu istatistiklerden anlaşıldığından, bu tezde sürücü kaza sonucu bilgilerine dayalı olarak yaralanma şiddetini tahmin eden bir sınıflandırma modeli kurulmuştur. Sürücü kusurlarının Türkiye’de yayımlanan kaza raporlarında %90 seviyesinde olması sürücü özelliklerinin, hatalarının kazaya etkisinin incelenmesi gerektiğini ortaya çıkarmaktadır. Sürücü özelliklerinin etkisini net bir şekilde görebilmek için kaza sonucuna etkisi çok yüksek olan kazanın yerleşim yeri içi/dışında olması, kazaya karışan taşıt sayısı gibi baskın öznitelikler ve taşıt büyüklüğü özneliği yalnızca otomobillerin karıştığı kazalar incelenerek ortadan kaldırılmıştır. Taşıt büyüklüğünün kazazedelerin yaralanma şiddetine olan etkilerinin

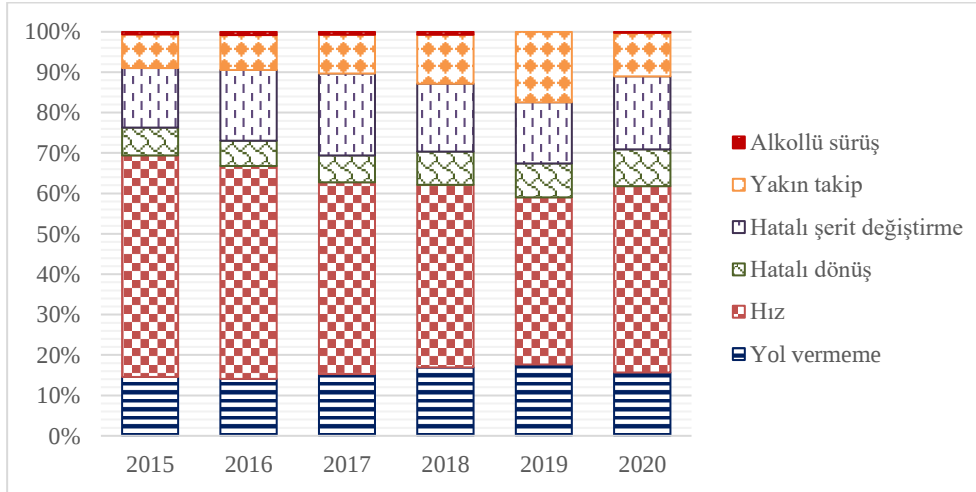
incelendiği pek çok araştırma literatürde bulunmaktadır. Bu çalışmalardan birinde Evans (1991), herhangi iki taşıtlı bir kazada, daha ağır olan aracın yolcularının daha hafif olan aracın yolcularından göreceli olarak daha düşük bir etkiyle kazayı atlatabildiklerini ifade etmiştir (Leonard Evans, 1991). Bu etkiyi ortadan kaldırmak için kaza iki ya da daha fazla taşıt içeriyorsa kazada yer alan diğer taşıtın otomobil olmadığı kazalar da veri setinden çıkarılmıştır. Bununla birlikte, otomobillerin savunmasız yol kullanıcısı olan yaya, motosiklet ve bisikletle çarpışması sonucu da sürücülerin kaza şiddetinin %94 oranında “sağlam” sınıfında yer aldığı bilindiğinden bu veriler de veri setinden çıkarılmıştır. İncelenecek kazaların yalnızca otomobillerin karıştığı kazalardan seçilmesinin nedeni, otomobillerin trafikteki taşıt kompozisyonundaki sahip oldukları yüksek orana bağlı olarak kaza veri setinde baskın bir etkiye sahip olma gerekçesine dayanmaktadır. Dolayısı ile yalnızca otomobillerin karıştığı kazalarda sürücülere ait etkin özniteliklerin analiz sonuçlarına bağlı olarak elde edilmesi, problemin ana yapısı ile derinlikli bir kavrayışın teminine olanak sağlayacağı gibi, trafik güvenliği stratejilerinin beklenen sonuçları üretmesine de önemli katkılar sağlamış olacaktır. Bu hedefler doğrultusunda, Sakarya’da 2015-2020 yılları arasında ölümlü-yaralanmalı türünde gerçekleşen otomobil kazalarına ait veri seti, bu çalışmadaki analiz çalışmaları ve model kurma süreçlerinde kullanılmıştır. Toplam 14,226 kaza içeren veri setinden yapılan veri ön işleme çalışmaları sonucunda elde edilen 3152 kazaya ait raporlama bilgilerine dayalı olarak analizler yapılmıştır. Bu kazalarda sürücülerin durumunun yıllara göre (kendi içindeki) dağılımları Şekil 3.6’da verilmiştir. Tüm yıllardaki veri sayıları farklı olup 2015 yılında kazaların %19’u, 2016’da %20’si, 2017’de %18’i, 2018’de %17’si, 2019’da %14’ü ve 2020’de %12’si gerçekleşmiştir.



Şekil 3.6. Karar Ağacında kullanılan veri setindeki sürücülerin kaza sonucu durumlarının yıllara göre oransal olarak dağılımı.

2015, 2016 ve 2017 yıllarında sürücülerin hayatını kaybettiği kazalar bulunmakta iken, 2018, 2019 ve 2020 yıllarında ise sürücülerin hayatını kaybettiği kaza bulunmamaktadır.

Şekil 3.7’de ise Karar Ağacı oluşturulurken kullanılan veri setindeki sürücülerin kusurlarının yıllara göre oransal dağılımı verilmektedir.



Şekil 3.7. Karar Ağacında kullanılan veri setindeki sürücü kusurlarının yıllara göre oransal olarak dağılımı.

Sürücü kusur dağılımları her yıl için birbirine oldukça yakın olmaktadır. Bazı yıllarda bazı kusurların dağılımı %1-5 arasında değişmektedir.

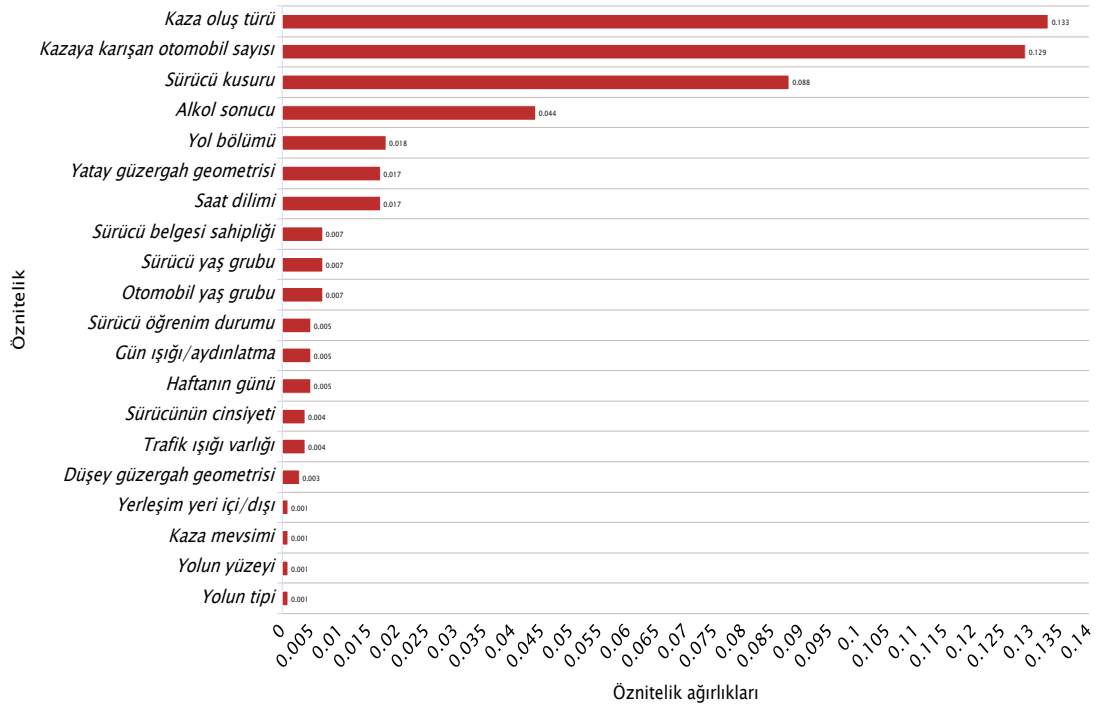
Literatürde en çok kullanılan Karar Ağacı yöntemlerinde, kural çıkarımında Gini indeksi (Sınıflandırma ve Regresyon Ağaçları / CART), entropi sayısı (ID3, C4.5 vb.) ve Ki-kare formülleri kullanılmaktadır. Bu çalışmada, bağımsız değişkenlerin

ağırlıkları bu yöntemler kullanılarak elde edilmiş, Şekil 3.8, Şekil 3.9 ve Şekil 3.10’da verilmiştir.

Entropi, bilginin sayısallaştırılmasıdır (Margaret H. Dunham, 2003). Denklem 3.1’de yer alan $p_1, p_2, p_3 \dots p_n$ değerleri olasılıkları, denklem 3.2’de yer alan $D(p)$, veri tabanının herhangi bir durumundaki halini göstermektedir.

$$p_1 + p_2 + p_3 + \dots + p_n = 1 \quad (3.1)$$

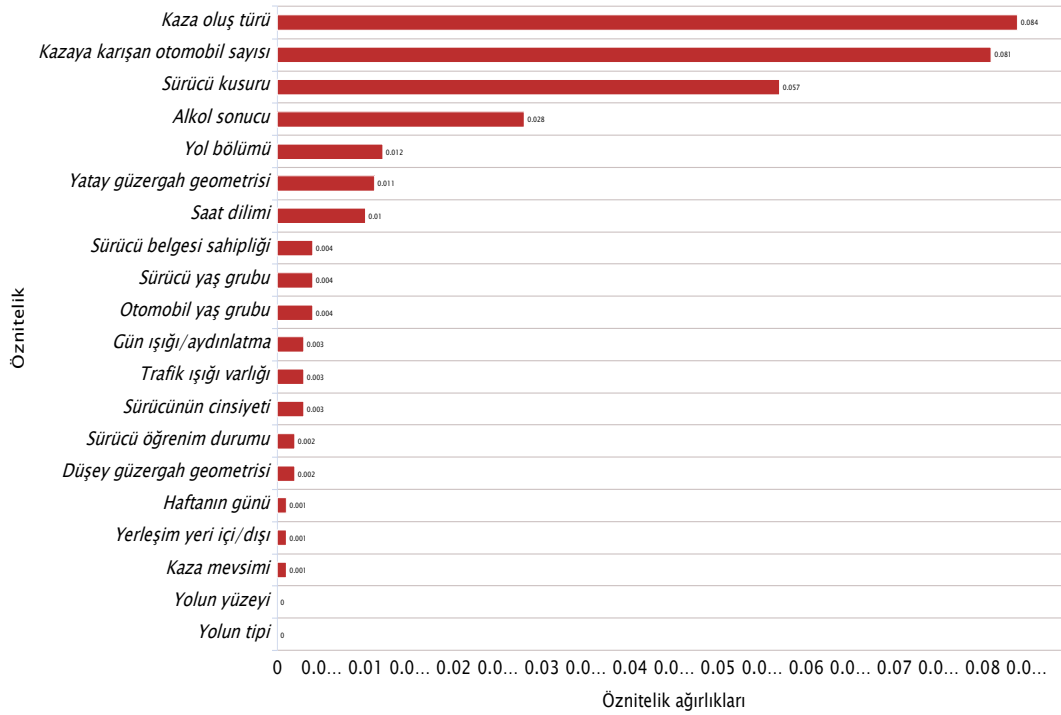
$$D(p_1, p_2, p_3, \dots, p_n) = \sum p_i \left(\log \left(\frac{1}{p_i} \right) \right) \quad (3.2)$$



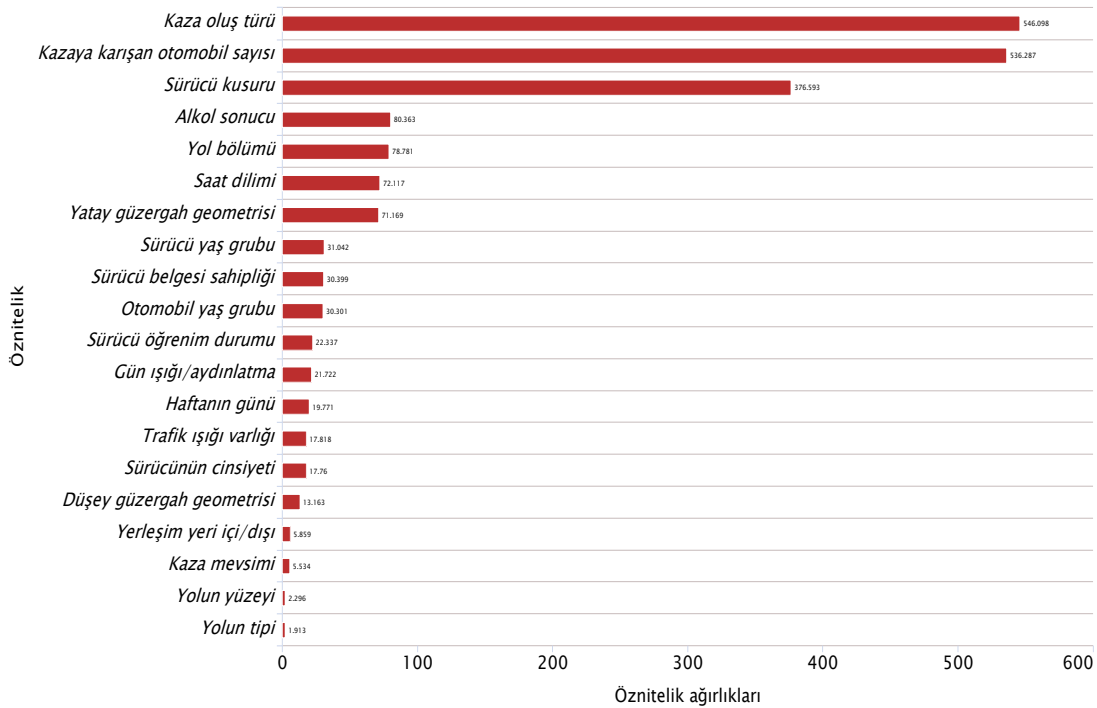
Şekil 3.8. Entropi sayısı formülü kullanılarak elde edilen öznitelik ağırlıkları sıralaması.

Gini indeksi, Karar Ağacındaki düğümlerin heterojenlik derecesini ölçmek için kullanılmaktadır. Ağaçlardaki dallanmanın, bölünmelerin en iyi şekilde olması için kullanılmaktadır (Han vd., 2012). Gini indeksi hesaplaması denklem 3.3’te verilmekte olup bu indeks büyüdükçe sınıf heterojenliği artmaktadır (Tangirala, 2020).

$$\text{Gini (D)} = 1 - \sum_{i=1}^n p_i^2 \quad (3.3)$$



Şekil 3.9. Gini indeksi formülleri kullanılarak elde edilen öznitelik ağırlıkları sıralaması.



Şekil 3.10. Ki-kare formülleri kullanılarak elde edilen öznitelik ağırlıkları sıralaması.

Özniteliklerin ağırlıklarının sıralamasının benzer olduğu Şekil 3.8, Şekil 3.9 ve Şekil 3.10 incelendiğinde görülmektedir. Bu değerler (indeksler) kullanılarak Karar

Ağaçları çizilmiştir. Ağacın en sade ve yorumlanabilir hali CHAID olduğu için bu çalışmada Karar Ağacı algoritmalarından olan CHAID tercih edilmiştir.

Sınıflandırma yönteminde veri seti, eğitim ve test olmak üzere iki gruba ayrılmaktadır. Bu ayrım, rastgele (random), katmanlı (stratified), doğrusal (satırların ilk %70'i, son %30'u şeklinde) yapılabilir. Çalışma kapsamında tüm ayrımlar denenmekle beraber, modellerin başarı sonuçları yakın olsa da trafik kazalarının yıldan yıla benzer şekilde (oluş türü, sürücü kusur dağılımı, sürücü kaza sonucu dağılımı) dağıldığı gerçeğinden dolayı doğrusal ayrımın uygulanması uygun bulunmuştur. Bu doğrultuda, veri setindeki veriler, 2015 yılından 2020 yılına doğru sıralanarak ilk %70'lik kısmının eğitim veri seti, sona kalan %30'luk verinin ise test veri seti olması sağlanmıştır. Eğitim seti sürücünün hayatını kaybettiği kazaları kapsamakla beraber, test veri setinde hayatını kaybeden sürücülerin yer almadığı Şekil 3.6'da görülmektedir. Bu durumda, ileride oluşturulacak modelin, sürücünün hayatını kaybettiği sınıfın başarımının "0" olacağını ifade etmek gerekmektedir. Sürücü kusurları dağılımının her yılda birbirine yakın olduğu ise Şekil 3.7'de görülmektedir.

CHAID yönteminde ağaçlar Ki-kare istatistiği değerlerine bağlı olarak dallara ayrılmaktadır. İstatistiksel anlamlılık testi olarak Ki-kare testi 1900 yılında Karl Pearson tarafından ortaya atılmıştır. Bu testte varyanslar, ortalamalar veya standart sapmalar değil, frekanslar önemlidir (Linoff ve Berry, 2011). Bu tez çalışmasında bağımlı değişken ile bağımsız değişken olarak seçilen değişkenler arasında anlamlı ilişki olup olmadığının belirlenmesi için Ki-kare bağımsızlık testi kullanılmıştır. Bu kapsamda bağımsızlık testinde sıfır ve alternatif hipotezler;

H_0 = "Sürücü kaza sonucu ile bağımsız değişken (sürücü kusuru, kaza oluş türü vb.) arasında birliktelik yoktur."

H_1 = "Sürücü kaza sonucu ile bağımsız değişken (sürücü kusuru, kaza oluş türü vb.) arasında birliktelik vardır."

şeklindedir.

Sıfır hipotezinin reddine veya kabulüne, %5 olarak belirlenen anlamlılık düzeyi değeri dikkate alınıp, Ki-kare istatistik ile Ki-kare kritik değerlerinin büyüklükleri arasında karşılaştırma yapılarak karar verilmiştir. Ki-kare istatistik değerleri kritik Ki-kare değerlerinden büyükse sıfır hipotezi reddedilmektedir. İlişkili olan iki değişkene istatistiksel olarak bağlı, ilişki olmayan değişkenlere ise istatistiksel olarak bağımsız

değişkenler denmektedir. Bağımlı öznelikle bağımsız öznelilik arasındaki ilişkinin büyüklüğünü sıralamak için aralarındaki Ki-kare istatistiği değerine bakılmaktadır. En yüksek Ki-kare değeri, sözü geçen iki değişken arasındaki ilişkinin en fazla olduğunu ortaya çıkarmaktadır (2011). Bağımlı değişkenle en fazla ilişkili olan (Ki-karenin en büyük olduğu) değişken, kurulacak ağaçta kök düğümde yer alır (Linoff & Berry, 2011).

Ki-kare testinde yer alan beklenen sıklık değerlerinin bağıntısı da aşağıda denklem 3.4'te yer almaktadır.

$$\text{Beklenen (B)} = \frac{\text{Satır Toplamı}}{\text{Toplam veri}} * \text{Sütun Toplamı} \quad (3.4)$$

Bu çalışmada bağımlı değişken olan sürücü kaza şiddeti ile, diğer öznelilikler arasında istatistiksel olarak anlamlı bir ilişki olup olmadığı, CHAID algoritmasında kullanılan Ki-kare bağımsızlık testi uygulanarak belirlenmiştir. Ki-kare istatistik değerinin hesaplanmasına örnek olması açısından Tablo 3.6'da veri setinde yer alan sürücü kusuru ve sürücü kaza sonucuna ait gerçekleşen (G) sıklık değerleri, Tablo 3.7'de beklenen (B) sıklık değerleri ile birlikte gerçekleşen değerler verilmiştir.

Tablo 3.6. Sürücü kusuru ve sürücü kaza sonucu durumu özneliğindeki gerçekleşen değerler

		Sürücünün kaza sonucu durumu			
		Sağlam	Yaralı	Ölü	Toplam
Sürücü kusuru	Hız	419	1106	4	1529
	Yakın takip	233	108	0	341
	Hatalı şerit değiştirme	192	345	2	539
	Yol vermeme	312	179	0	491
	Hatalı dönüş	140	94	0	234
	Alkollü sürüş	1	17	0	18
	Toplam	1297	1849	6	3152

Tablo 3.7'de üstte yer alan değerler Tablo 3.6'da da verilen gerçekleşen değerler, altta verilenler ise denklem 3.1 ile elde edilen beklenen değerlerdir. Denklem 3.5'te hız kusur sonucunda sağlam kalması beklenen sürücü sayısı 629,16 bulunmuştur.

$$\text{Beklenen değer} = \frac{1529}{3152} * 1297 = 629,16 \quad (3.5)$$

Tablo 3.7. Sürücü kusuru ve sürücü kaza sonucu durumu özniteliğindeki gerçekleşen ve beklenen değerler.

		Sürücünün kaza sonucu durumu			Toplam
		Sağlam	Yaralı	Ölü	
Sürücü kusuru	Hız	419 629,16	1106 896,93	4 2,91	1529
	Yakın takip	233 140,32	108 200,03	0 0,65	341
	Hatalı şerit değiştirme	192 221,79	345 316,18	2 1,03	539
	Yol vermeme	312 202,04	179 288,03	0 0,93	491
	Hatalı dönüş	140 96,29	94 137,27	0 0,45	234
	Alkollü sürüş	1 7,41	17 10,56	0 0,03	18
	Toplam	1297	1849	6	3152

Ki-kare değerinin hesabında kullanılan matematiksel yapı denklem 3.6’da verilmiştir. Bağlıdaki “G” harfi gerçekleşen gözlemi, “B” harfi gerçekleşmesi beklenen gözlemi ifade etmektedir. Denklem 3.7’te ise, Tablo 3.7’de yer alan, sürücü kusuru ve sürücü kaza şiddeti arasındaki ilişkinin belirlenmesi için yapılan Ki-kare hesabı yer almaktadır.

$$\chi^2 = \sum_{i=1}^k \frac{(G_i - B_i)^2}{B_i} \quad (2011) \quad (3.6)$$

$$(i = 1, 2, 3, \dots, k)$$

$$\chi^2 = \frac{(419-629,16)^2}{629,16} + \frac{(1106-896,93)^2}{896,93} + \dots + \frac{(0-0,03)^2}{0,03} = 376,59 \quad (3.7)$$

Çalışma içeriğinde kullanılan özniteliklere ait Ki-kare değerlerinin dağılımları Şekil 3.8’de verilmiştir. Kaza oluş türü değişkeni ile sürücü kaza şiddetinin Ki-kare istatistiğinin en büyük değeri alması, bu iki değişken arasındaki ilişkinin diğer bağımsız değişkenler ile sürücü kaza şiddeti değişkeni arasındaki ilişkilerden daha büyük olduğunu ortaya çıkmıştır. Yine elde edilen yüksek Ki-kare değeri, kazaya karışan taşıt sayısı ile bağımlı değişken olan sürücü kaza şiddeti arasında istatistiksel olarak güçlü ve anlamlı bir ilişki olduğunu ifade etmektedir. Bu bulgular, literatürdeki bir çok çalışma ile örtüşmekte ve uyumluluk göstermektedir (Chang & Wang, 2006; Ijaz vd., 2021; Montella, 2011; Prati vd., 2017; Wang & Kim, 2019). Bu bulgulara ilaveten, yerleşim yeri içi/dışı, kaza mevsimi, yolun yüzeyi ve tipi haricindeki bağımsız değişkenler ile bağımlı değişken arasında da istatistiksel olarak anlamlı bir ilişki olduğu belirlenmiştir.

Şekil 3.10’da verilen Ki-kare istatistik değerlerine ek olarak, Tablo 3.8’de serbestlik derecesi, anlamlılık düzeyi (level of significance) ve Ki-kare kritik değerleri verilmiştir. Serbestlik derecesi istatistikte bir istatistiğin kesin hesaplanmasında kullanılan değerlerin sayısının ne kadar değişme serbestisi olduğunu sayısal olarak vermektedir (“Serbestlik derecesi (istatistik)”, 2024). Bağımlı değişken (öznitelik) ile bağımsız öznitelik arasındaki serbestlik derecesi ($sd=df$ (degree of freedom)), iki öznitelikteki sınıfların bir eksiklerinin çarpılması ile elde edilmektedir (Yang vd., 2023). Anlamlılık düzeyi (α), bir sıfır hipotezini reddetme olasılığı olup 0,05 seçilerek hipotez testinin anlamlılık düzeyi olarak adlandırılmaktadır. Ki-kare kritik tablosundan Ki-kare kritik değerlerine bakılırken hem serbestlik derecesi hem anlamlılık düzeyi (α) değerlerinden yararlanılmaktadır.

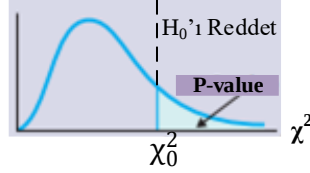
Tablo 3.8. Tüm özniteliklerin Ki-kare istatistik, kritik ve serbestlik dereceleri.

Öznitelik (Çıktı)	v_1-1	Öznitelikler (Girdiler)	v_j-1	sd	χ^2 ist	χ^2 kri
Sürücü kaza şiddeti (sağlam, yaralı, ölü)	2	Kaza oluş türü	4	8	546,10	15,51
		Kazaya karışan otomobil sayısı	2	4	536,29	9,49
		Sürücü kusuru	5	10	376,59	18,31
		Alkol sonucu	1	2	80,36	5,99
		Yol bölümü	4	8	78,78	15,51
		Saat dilimi	3	6	72,12	12,59
		Yatay geçki geometrisi	1	2	71,17	5,99
		Sürücü yaş grubu	3	6	31,04	12,59
		Sürücü belgesi sahipliği	1	2	30,40	5,99
		Otomobil yaş grubu	2	4	30,30	9,49
		Sürücü eğitim durumu	3	6	22,34	12,59
		Gün ışığı/aydınlatma	2	4	21,72	9,49
		Haftanın günü	6	12	19,77	21,03
		Trafik ışığı varlığı	1	2	17,82	5,99
		Sürücünün cinsiyeti	1	2	17,76	5,99
		Düşey geçki geometrisi	1	2	13,16	5,99
		Yerleşim yeri içi/dışı	1	2	5,86	5,99
		Kaza mevsimi	3	6	5,53	12,59
		Yolun yüzeyi	1	2	2,30	5,99
		Yolun tipi	3	6	1,91	12,59

$\alpha = 0,05$ (anlamlılık düzeyi)

Şekil 3.9’da Ki-kare istatistik ve Ki-kare kritik (χ_0^2) değerlerinin karşılaştırılması sonucunda sıfır hipotezinin reddine karar verileceği alan verilmiştir. Sürücü kusuru için Ki-kare istatistik değeri denklem 3.3’te hesaplandığı ve Tablo 3.8’de görüldüğü gibi 376,59’dur. Bu değişken ile sürücü kaza şiddeti arasındaki Ki-kare kritik değeri Ki-kare dağılımı tablosundan bakılarak 18,31 olarak saptanmıştır. Ki-kare istatistik değeri, Ki-kare kritik değerinden büyük olduğu için Şekil 3.11’deki grafikten

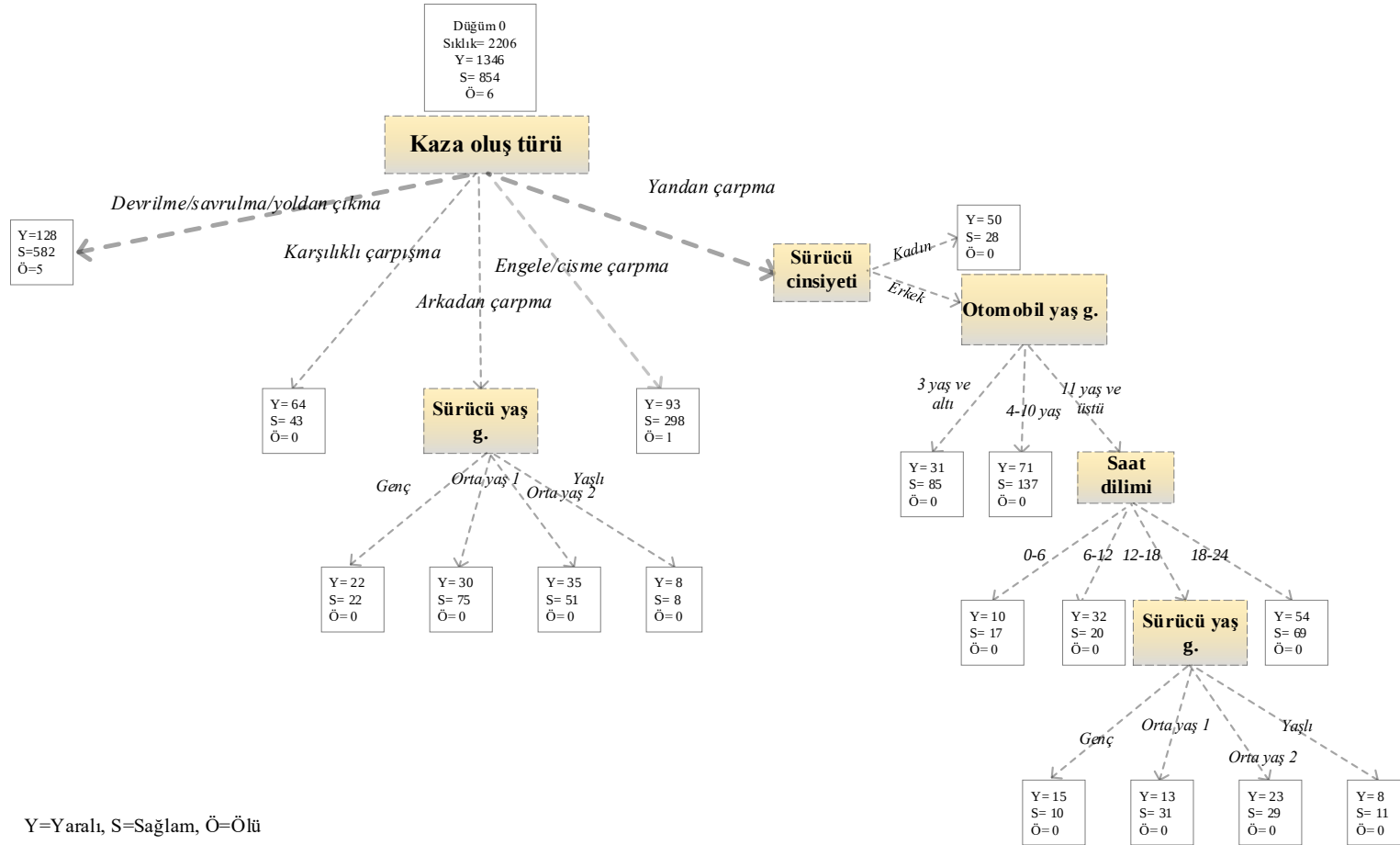
yararlanılarak sıfır hipotezinin reddine, böylelikle “Sürücü kaza sonucu ile bağımsız değişken (örn. sürücü kusuru) arasında birliktelik vardır.” sonucu elde edilmektedir. Diğer tüm değişkenler için bu hesaplama ve karşılaştırmalar yapılarak birliktelik (ilişki) varlığına karar verilmektedir.



Şekil 3.11. Ki-kare kritik ve Ki-kare istatistik değerlerinin görüldüğü Ki-kare dağılım grafiği, Neil A. Weiss (2011)’den uyarlanmıştır.

Tablo 3.8’de verilen bütün bağımsız (girdi) olan özniteliklerin her biri ile bağımlı (çıktı) olan öznitelikler aralarındaki ilişkinin istatistiksel anlamlılığını bulmak için her bir öznitelik çifti (bağımlı-bağımsız) için Ki-kare istatistik ve Ki-kare kritik değerleri arasında karşılaştırma yapılarak, sıfır hipotezinin reddine veya kabul edilmesine karar verilmiştir. Sıfır hipotezinin (“Sürücü kaza sonucu ile bağımsız değişken (örn. kaza oluş türü) arasında birliktelik yok”) reddi, istatistiksel olarak bu iki öznitelik arasındaki birlikteliğin anlamlı olduğu sonucunu doğurmaktadır.

Bağımlı değişkenle istatistiksel olarak anlamlı birlikteliği olan öznitelikler kullanılarak Şekil 3.12’de verilen CHAID Karar Ağacı oluşturulmuştur.



Şekil 3.12. Tüm özneliliklerin yer aldığı veri setiyle oluşturulan Karar Ağacı.

“Kaza oluş türü”, sürücü kaza sonucu ile en güçlü ilişkiye sahip olduğu için kök düğümde yer almıştır. Tüm veri setinin %83’ünü oluşturan devrilme/savrulma/yoldan çıkma özniteliği, ölümlü kaza sonuçlarının oluşmasında oldukça etkin bir rol taşımaktadır. %2 oransal değeri ile, veri setinde göreceli olarak düşük kabul edilebilecek karşılıklı çarpışma türünde olan kazaların, %60 oranında sürücülerin yaralanması ile sonuçlanması, önemli bir bulgu olarak elde edilmiştir. Veri setinin yalnızca otomobillerden oluşması ve sürücünün hayatını kaybettiği yalnızca 6 kazanın bulunmasından dolayı, bu tür kazalarda sürücünün hayatını kaybettiği durum görülmemiştir.

Yapılan analizlerde, arkadan çarpma şeklinde oluşan kazalara karışan sürücülerin yaşlarına bağlı olarak, sürücü kaza sonuçlarının değiştiği görülmüştür. Bu tip kazalarda genç (18-25) ve yaşlı (65+) sürücülerin yarı yarıya oranla yaralı ve sağlam olarak kazadan çıktığı, orta yaş 1 (26-40) grubunda yer alan sürücülerin %71 oranında, orta yaş 2 (41-64) grubunda olanların ise %59 oranında kazadan sağlam çıktığı belirlenmiştir.

Engele, cisme çarpma kazalarına karışan sürücülerden biri hayatını kaybetmiştir. Bu sayı, ağacı oluşturan eğitim veri setinde sürücünün öldüğü tüm kazaların %17’sine (1/6) karşılık gelmektedir. Bu tip kazalara karışan sürücülerin %76’sı kazadan sağlam olarak çıkmıştır.

Yandan çarpma türünde görülen kazalarda, sürücü ile ilgili kaza şiddetinin, ilk olarak sürücü cinsiyetine bağlı olduğu görülmüştür. Bu kazalarda hatalı olan sürücülerden ölen sürücü olmamıştır. Bu tip kazalara karışan kadın sürücülerin %64’ünün kendisini yaraladığı belirlenmiştir. Erkek sürücülerde ise ilk önce otomobil yaşına göre bir ayırım oluşmuş olup taşıtının modeli 3 yaş ve altı olanların %73’ü kazadan sağlam olarak, geri kalanı yaralı olarak çıkarken, taşıtının modeli 4-10 yaş arasında olan sürücülerin %66’sı sağlam olarak çıkmıştır. Taşıtı 11 yaş ve üzerinde olanlar arasında saat dilimine göre tekrar bir ayırım gerçekleşmiş olup 00:00-06:00 saatleri arasında sürücülerin %63’ünün sağlam olduğu, 06:00-12:00 saatleri arasında %62’sinin yaralandığı, 18:00-24:00 saatleri arasında yaklaşık olarak yarısının yaralanıp kalanının sağlam olarak kazadan çıktığı görülmüştür. 12:00-18:00 saatleri arasında ise sürücü kaza şiddeti, sürücü yaşına bağlı olarak değişmiştir. Genç (18-25) sürücülerin %60’ı yaralanmış, orta yaş 1 (26-40) grubundaki sürücülerin %70’inin, orta yaş 2 (41-64) grubundaki sürücülerin %56’sının, yaşlı sürücülerin ise %58’inin sağlam olduğu bulunmuştur. Bu

sonular, gen (18-25) srclerin kendilerinin kusurlu olduėu yandan arpma kazalarında daha Őiddetli bir arpıřmaya sebebiyet verdikleri ve sonucunda kendilerinin zarar grdklerini ortaya koymaktadır.

Bu sonular incelenirken, srclerin saėlam olduėu kazalarda bařka bir src ya da yolcu ya da yayadan en az birinin yaralanmıř ya da lmř olduėu bilgisi gz ardı edilmemelidir. Ayrıca bu veri setindeki tm srclerin kusurlu olduėu bilgisi de unutulmamalıdır.

3.2.1. Elde edilen Karar Aėacı modeli bařarım dzeyi

Sınıflandırma modellerinin tahmin bařarılarını lmek iin hata (karıřıklık/confusion) matrisi kullanılmaktadır. İki den fazla sınıf ieren sınıflandırma problemlerinde, tahmin edilen sınıf ile gerek sınıf karřılařtırılarak, her bir stunu gzlenen gerek durumu, her bir satırı ise tahmin edilen durumu temsil eden bir karıřıklık matrisi (confusion matrix) oluřturulur. Matrisin her bir stunu gzlemlenen gerek durumu, her bir satırı ise tahmin edilen durumu temsil eder. rneėin, gerekte yaralı olan biri satırda da yaralı hcresindeyse doėru tahmin edilmiř, satırlardan (tahmin) saėlam olan bir hcrede ise, yaralı olan birinin saėlam olarak yanlıř tahmin edildiėi anlařılmaktadır. Hcrelerdeki deėerler, verilerden kaının bu sınıfa atandıėını belirten deėerler olmaktadır. Karıřıklık matrisinde křegen zerindeki hcrelerdeki deėerler doėru sınıflandırılmıř veri sayısını, geri kalanlar ise yanlıř sınıflandırmaları temsil etmektedir (Han vd., 2012; Krstini vd., 2020).

Karıřıklık matrisi, sınıflandırmanın geerliliėini, doėruluėunu, ėrenme bařarısını lmede kullanılır. Bu deėerlendirme iin, hassaslık (precision) ve duyarlılık (recall/doėru pozitif oranı) gibi ltler kullanılmaktadır. Karıřıklık matrisi incelendiėinde, sınıfların bařarıları, sınıflar arasındaki baėlantılar hakkında deėerli gstergeler saėlamaktadır. Karıřıklık matrisi incelenerek sınıflandırma modelinin zayıf ynleri ortaya ıkarılabilir, daha ileri arařtırmalara ve model bařarısını iyileřtirmeyi saėlayacak bilgiler elde edinilebilir. Ayrıca, karıřıklık matrisinin analizi, eřitli veri zellikleri ve nesneleri arasındaki iliřkilere ıřık tutmakta ve verinin kendi yapısını ortaya ıkarabilmektedir.

Modeldeki bařarı lt olan "doėruluk (accuracy)" deėeri, tm sınıflardaki doėru tahmin edilen veri sayısının toplam veri miktarına blnmesiyle elde edilen orandır. "Kesinlik (precision)" lt ise bir sınıftaki doėru tahmin edilen veri sayısının, o sınıf

için tahmin edilen tüm veri miktarına bölünmesiyle elde edilen orandır. Bir diğer performans ölçütü olan ve duyarlılık olarak da bilinen "duyarlılık (recall)" ise bir sınıfta doğru tahmin edilen veri sayısının o sınıfta gözlemlenen toplam veri miktarına bölünmesiyle elde edilen orandır (Orr, 1972).

N*N boyutundaki bir hata matrisinde, sınıflandırma başarısı her sınıf için ayrı ayrı kesinlik ve duyarlılık değerleri ile ölçülebilir (Ufuk Çelik vd., 2017).

Tablo 3.9’da tüm değişkenlerin yer aldığı Karar Ağacı modelinin başarımlar değerleri verilmiştir.

Tablo 3.9. Tüm değişkenler için kurulan Karar Ağacı modelinin başarımları.

Tüm değişkenler		Gerçekleşen		
Tahmin edilen		Yaralı	Sağlam	Ölü
	Yaralı	227 (1)	111 (2)	0 (3)
	Sağlam	216 (4)	392 (5)	0 (6)
	Ölü	0 (7)	0 (8)	0 (9)
	Duyarlılık (%)	51,24	77,93	0,00
		Doğruluk: %65,43		

*Parantez içindeki değerler hücre isimleridir.

Her sınıf için doğru pozitif, doğru negatif, yanlış pozitif ve yanlış negatif değerleri aşağıda, Tablo 3.10’daki gibi hesaplanmaktadır:

Tablo 3.10. Her sınıfa ait gerçekleşen ve tahmin edilen verilere göre doğru ve yanlış sınıflandırma değerleri.

Yaralı (Sürücü)	Sağlam (Sürücü)	Ölü (Sürücü)
DP = 227 (1)	DP = 392 (5)	DP = 0 (9)
YP = 111+0 (2-3)	YP = 216+0 (4-6)	YP = 0+0 (2-3)
DN = 392+0+0+0 (5-6-8-9)	DN = 227+0+0+0 (1-3-7-9)	DN = 227+111+216+392 (1-2-4-5)
YN = 216+0 (4-7)	YN = 111+0 (2-8)	YN = 0+0 (3-6)

*Parantez içindeki değerler Tablo 3.7’deki hücrelerdir.

DP (Doğru pozitif): Doğru tahmin edilen pozitif veri sayısı

DN (Doğru negatif): Doğru tahmin edilen negatif veri sayısı

YP (Yanlış pozitif): Yanlış tahmin edilen pozitif veri sayısı

YN (Yanlış negatif): Yanlış tahmin edilen negatif veri sayısı

T (Toplam): Toplam veri sayısı

Sınıflandırmanın başarısını ölçmek için denklem 3.8, denklem 3.9 ve denklem 3.10’da verilen doğruluk, kesinlik ve duyarlılık hesaplamaları yapılmaktadır.

$$\text{Doğruluk} = \frac{\sum DP}{T} \quad (3.8)$$

$$\text{Kesinlik} = \frac{DP}{DP + YP} \quad (3.9)$$

$$\text{Duyarlılık} = \frac{DP}{DP + YN} \quad (3.10)$$

Denklem 3.11’de sınıflandırmanın doğruluğu, denklem 3.12 ve denklem 3.13’te tek bir sınıf için hesaplanabilen, burada sürücünün yaralı olması durumu için hesaplanan kesinlik ve duyarlılık değerleri verilmiştir.

Sınıflandırma doğruluğu:

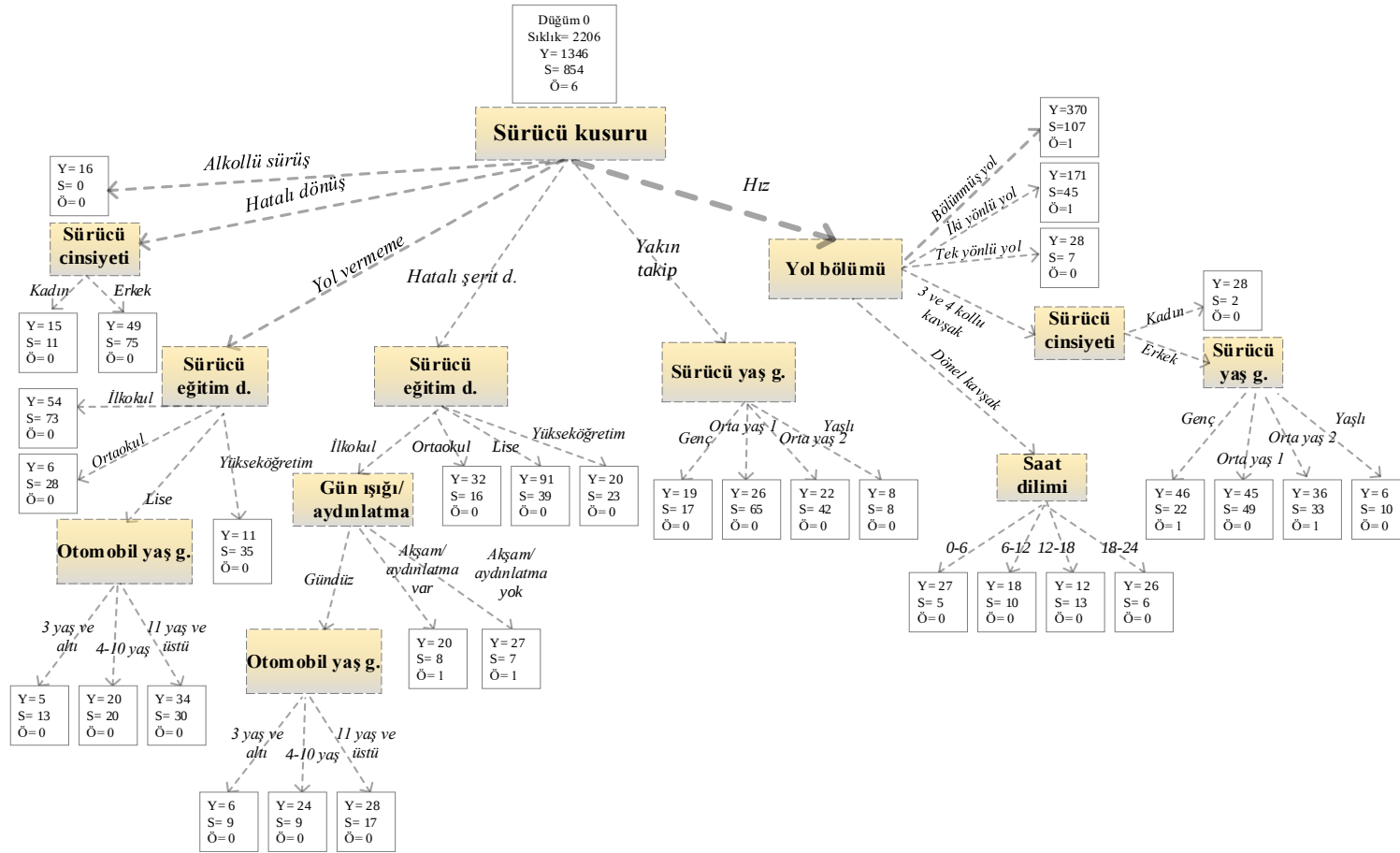
$$\text{Doğruluk} = \frac{\sum DP}{T} = \frac{227 + 392 + 0}{227 + 111 + 0 + 216 + 392 + 0 + 0 + 0 + 0} = \% 65,43 \quad (3.11)$$

Sürücünün yaralı olması sınıfı için kesinlik, duyarlılık değerleri:

$$\text{Kesinlik} = \frac{227}{227 + 111 + 0} = \%67,16 \quad (3.12)$$

$$\text{Duyarlılık} = \frac{227}{227+216+0} = \%51,24 \quad (3.13)$$

Daha sonra sürücü özelliklerinin sürücü kaza sonucuna etkisini görebilmek için tekrar veri ön işleme süreci uygulanarak bazı baskın (kaza oluş türü, kazaya karışan taşıt sayısı) ve sürücüden daha bağımsız (yatay ve düşey geçki geometrisi, haftanın günü, trafik ışığı varlığı) ve istatistiksel olarak bağımlı değişkenle ilişkisi anlamsız olan (yerleşim yeri içi/dışı, kaza mevsimi, yolun tipi ve yol yüzeyi) öznitelikler çıkarılarak kalan özniteliklerle yeni bir CHAID Karar Ağacı oluşturulmuştur. Oluşturulan bu Karar Ağacı, Şekil 3.13’te verilmiştir.



Y=Yaralı, S=Sağlam, Ö=Ölü

Şekil 3.13. Seçilen öz niteliklerin yer aldığı veri setiyle oluşturulan Karar Ağacı.

Bu Karar Ağacında ise, sürücü kusuru kök düğümde yer alarak, sürücü kaza şiddeti ile en fazla ilişkisi olan öznitelik olmuştur.

Sürücü kusuru olarak “hız kusuru” nun öne çıktığı, bu durumunda sürücülerin taşıt hızını yol ve havanın gerektirdiği koşullara uydurmamasından kaynaklandığı görülmektedir. Ağacın hız kusurunda bulunan yapraklarına bakıldığında bu kusuru yapan sürücülerin çoğunluğunun daha çok yaralandığı ortaya çıkmaktadır. Sonraki ayırımın (düğümün) yol bölümü özniteliği olduğu ve sürücü kaza şiddetinin de bu özniteliğe göre değiştiği görülmektedir. Bölünmüş yollarda hızını yol ve hava koşullarına uydurmayan sürücülerin %77 oranında yaralandığı belirlenirken, tüm veri setindeki sürücünün öldüğü kazalardan %17’sinin de bölünmüş yollarda gerçekleştiği belirlenmiştir. İki yönlü yollarda gerçekleşen kazalarda genel anlamda sürücülerin %79 oranında yaralandığı belirlenirken, tek yönlü yollarda bu oran çok yakın bir değerde %80 olarak hesaplanmıştır. Üç ve dört kollu kavşaklarda, sürücü cinsiyetine göre bir ayırım gerçekleşmiş, kadın sürücülerin %93 oranında kendilerini yaraladığı belirlenmiştir. Erkek sürücüler için bu tip kavşaklarda sürücünün yaşına göre tekrar bir ayırım ortaya çıkmış ve genç (18-25) sürücülerin bu tip kavşaklarda hız kusuru yaptıklarında tüm veri setinin %17’si oranında öldükleri, bu grubun içinde ise %67’sinin yaralandığı, %1’inin öldüğü, %32’sinin sağlam kaldığı belirlenmiştir. Orta yaş 1 (26-40) ve orta yaş 2 (41-64) grubundaki erkek sürücülerin sağlam ve yaralanma oranları, kendi grup içlerinde ikisi içinde yaklaşık olarak yarı yarıya olarak hesap edilmiştir. Orta yaş 2 (41-64) grubunda sürücünün öldüğü durum tüm veri seti içerisinde %17’lik bir oranla yer alırken, yaşlı erkek sürücülerin %62 oranında sağlam, %38 oranında yaralı olarak kazadan çıktığı belirlenmiştir. Analiz edilen veri setinde taşıt hızını yol ve hava koşullarına uydurmama nedeniyle (yalnızca otomobillerin karıştığı ve sürücülerin kusurlu olduğu) hiçbir yaşlının hayatını kaybetmediği görülmektedir.

Yol vermeme hatası yapan sürücülere ait kaza şiddetinin eğitim durumuna göre değiştiği bulunmuştur. İlkokul, ortaokul ve yüksek öğrenim mezunu olanların daha çok kazadan sağlam olarak çıktığı, lise mezunu sürücülerin ise kaza sonucunun otomobil yaşına bağlı olduğu görülmektedir. 3 yaş ve altı taşıt sahibi sürücülerin %72 oranla sağlam olarak kazadan çıktığı, 4 ve üzeri yaş taşıt sahiplerinin sağlam ve yaralı olma oranlarının yaklaşık olarak yarı yarıya olduğu belirlenmiştir.

Alkollü olarak taşıt kullanan sürücülerin tamamının kazadan yaralı olarak çıktığı, bir diğer ifade ile hiçbirinin sağlam olarak çıkmadığı sonucuna ulaşılmıştır. Hayatını kaybeden sürücü sayısı çok az olduğu için alkollü sürüşten dolayı hayatını kaybetme durumu analiz sonuçlarında belirlenememiş olsa da diğer kusurları işleyen ve hayatını kaybetmiş olan sürücülerin aynı zamanda alkollü olabileceği gerçeği unutulmamalıdır.

Hatalı dönüş yapan sürücülerin kaza sonucu durumları sürücü cinsiyetine göre değişiklik gösterdiği, kadın sürücülerin %58'inin yaralanırken, %42'sinin sağlam olduğu, erkek sürücülerin ise %60 oranı ile sağlam, %40'oranı ile de yaralandığı görülmüştür.

Hatalı şerit değiştirenlerin kaza sonucu durumu eğitim durumuna göre değişmiş olup orta okul mezunu olanların %67'si, lise mezunu olanların %70'i, yüksek öğrenim mezunu olanların %47'si kendisini yaralanmış olup bu eğitim seviyesinde olan sürücülere ait herhangi bir ölüm kaydına rastlanmamıştır. İlkokul mezunu olanların kaza sonucunun yapılan analizde gün ışığına göre farklılık gösterdiği, gündüz olan kazalarda da otomobil yaş grubuna bağlı olarak farklılıklar belirlenmiştir. 3 yaş ve altındaki taşıtları kullanan sürücülerin %60'ı kazadan sağlam olarak çıkarken, 4-10 yaş arasındaki taşıt sahipleri %73 oranında yaralanmıştır. 11 yaş ve üstü taşıt sürücüleri ise %62 oranında yaralanmıştır. Akşam olup aydınlatmanın da olduğu durumda gerçekleşen kazalarda sürücülerin %3'ü ölüp, %69'u yaralanırken, %28'i kazadan sağlam olarak çıkmıştır. Akşam ve aydınlatmanın olmadığı durumda gerçekleşen kazalarda da sürücülerin %3'ü ölmüş, %77'si yaralanmış ve %20'si sağlam olarak kazadan çıkmıştır.

Tablo 3.11'de seçilen değişkenlerin yer aldığı Karar Ağacı modelinin sürücü kaza sonucuna bağlı olarak sahip olduğu başarımlar değerlendirilmiştir.

Tablo 3.11. Seçilen değişkenler için kurulan Karar Ağacı modelinin başarımları.

Seçilen değişkenler		Gerçekleşen		
Tahmin edilen		Yaralı	Sağlam	Ölü
	Yaralı	244	122	0
	Sağlam	199	381	0
	Ölü	0	0	0
	Duyarlılık (%)	55,08	75,75	0,00
		Doğruluk: %66,07		
		Kesinlik (%)		

Eğitim verileriyle (tüm verinin %70'i) elde edilen Karar Ağacı üzerinde test verilerinin (tüm verinin %30'u) doğruluğu sorgulanarak belirlenip tahmin edilen ve gerçekleşen

(kaza verisi) sınıflar ile ilgili karışıklık (hata) matrisi ve başarımlarına ilişkin bilgiler Tablo 3.11’de verilmiştir. Tabloda sürücü kaza sonucunun yaralı olarak tahmin edilme olasılığının %55, sağlam olarak tahmin edilmesinin %75, ölü olarak tahmin edilmesinin %0 oranında doğru gerçekleşeceği belirtilmiştir. Ölü sınıfında hiç doğru tahmin olmamasının nedeni, veri setinde sürücünün öldüğü tüm kazaların eğitim veri setinde kalmasıdır.

4. BULGULAR VE TARTIŞMA

Bu çalışmada, Sakarya’da 2015-2020 yılları arasında trafik kazalarının gerçekleşmesine neden olan etkenler, veri setinde en sık geçen durumları belirlemek için apriori algoritması kullanılarak Birliktelik Kuralları ortaya çıkarılarak elde edilmiştir.

4.1. Birliktelik Kuralları Sonuçlarının Literatürle Değerlendirilmesi

- Taşıtın hızını yola uydurmama kusuru, taşıtın ağır hasar alması ve sürücünün yaralanması durumu kaza örüntülerinden biri olarak ortaya çıkmıştır. Bu bulguya uygun olarak Albuquerque ve diğ., engele çarpma ve yoldan çıkma türünde olan kazalarda, hızın yaralanma şiddetindeki en önemli etkenlerden biri olduğunu belirtmiştir (Albuquerque & Awadalla, 2020).
- Devrilme/savrulma/yoldan çıkma türünde gerçekleşen ve taşıtın ağır hasar aldığı kazaların %85’inde sürücünün kaza sonucu durumu yaralı/ölü şeklinde olmaktadır. Karabulut yaptığı çalışmada, motosiklet içeren, engele çarpma ve yoldan çıkma türündeki kazaların, sürücülerin kaza sonucunda yaralanması ya da ölümüyle ilişkili olduğunu belirlemiştir (Karabulut & Özen, 2023).
- Sürücünün öndeki taşıtı yakın takip etme kusuru işlemesi, arkadan çarpma türünde kazalar ile sonuçlanmaktadır. Yu ve diğerleri de bu sonuca ulaşmıştır (Yu vd., 2019). Bu durumu giderebilmek için sürücü asistanları geliştirilmekte, uyarlamalı hız kontrol sistemi ile, öndeki taşıt ile olan mesafenin uygun tutulması sağlanabilmektedir (T. Li vd., 2021).
- Yayaya çarpma türünde kazaların çoğunlukla sürücünün yaralanmaması ile sonuçlandığı bulunmuştur. Karabulut (2023) da çalışmasında benzer sonuca ulaşmıştır.
- Tüm kazaların %5’i gündüz, yol vermeme kusuru sonucu yandan çarpma türünde gerçekleşmiş olup taşıtların hafif hasar alması ile sonuçlanmıştır. Bunun temel nedeni, sürücülerin gündüz, ruhsal ve zihinsel olarak rahatlamış olması nedeniyle riskli manevralara yönelmeleri olabilmektedir. (Das vd., 2018).

- Analiz sonuçları, sürücünün ölüm ya da yaralanması ile sonuçlanan kazaların daha çok ilkbahar ve yaz mevsimlerinde gerçekleştiğini ortaya çıkarmıştır. Bu durumu destekleyecek şekilde, elde edilen kurallarda sürücülerin ilkbahar ve yaz mevsimlerinde gündüz saatlerinde hız yapma eğiliminde oldukları belirlenmiştir. Bu bulgu ile uyumlu olarak Çelik ve Oktay (2014), yaz aylarında ölümcül yaralanmaların 'yaralanma yok' ciddiyet düzeyine göre iki kat daha fazla meydana geldiğini belirterek, ölümcül olmayan yaralanma olasılığının %6,3 oranında arttığını ifade etmiştir (Çelik & Oktay, 2014).
- Elde edilen başka bir Birliktelik Kuralına göre, ilkokul mezunu sürücülerin ölümlü-yaralanmalı kazalarda ve yol vermeme kusuru işleyerek yandan çarpma türünde kaza oluşumlarında oldukça yüksek bir orana sahip oldukları görülmektedir. Çelik de sürücülerin ilkokul mezunu olmalarının kaza şiddetinin artmasında etkili olduğu sonucunu paylaşmıştır (Çelik & Oktay, 2014).
- Kazaların %8'inin, ışıksız kavşaklarda sürücülerin yol vermeme kusuru işleyerek yandan çarpma türünde gerçekleştiği belirlenmiştir. Yu ve diğerleri, kazaların %15'inin kavşaklarda yol vermeme kusuru sonucunda meydana geldiğini ifade etmektedir (Yu vd., 2019).
- 18-25 yaş arası genç sürücülerin karıştığı kazaların, yaralanmalara ve taşıtların ağır hasar almasına neden olduğu bulunmuştur. Bunun temel nedeni, bu sürücülerin nispeten hızlı taşıt kullanma eğiliminde olmalarıdır. Dolayısıyla bu bulgu, hız ile sürücülerin ağır yaralanmasıyla sonuçlanan şiddetli kazalar arasında güçlü bir ilişki olduğunu doğrulamaktadır. Yu ve diğerleri de 16-25 yaş aralığındaki sürücülerin diğer sürücülerden daha az trafik kurallarına uyduğunu, dolayısıyla daha çok sayıda ve daha şiddetli kazalara neden olduklarını belirtmiştir (Yu vd., 2019). Çelik ve Oktay (2014), başarılı uygulamaları bir rol model olarak alarak tüm yol kullanıcılarına, özellikle de yaşlı sürücülere yönelik eğitim kampanyalarının gerekliliğini vurgulamıştır (Çelik & Oktay, 2014). Bu çalışmada elde edilen analiz sonuçları ise, genç sürücülere yönelik kampanya ve/veya yaptırımlara daha çok ihtiyaç duyulduğu gerçeğini ortaya koymaktadır.
- Çalışmada analiz edilen ölümlü ve yaralanmalı kazaların ezici bir çoğunlukla ışıklı kavşaklardan ziyade ışıksız kavşaklarda gerçekleştiği görülmektedir. Çelik de çalışmada kavşaklarda ışık olmasının kazaların ölümlü olma olasılığını %41.8 oranında düşürdüğünü bulmuştur (Çelik & Oktay, 2014).

- Bu çalışmada taşıt yaşının kaza şiddeti ile ilişkili olduğu belirlenmiştir. 11 ve üstü yaşlardaki taşıt sahibi olan sürücülerin daha yüksek hızlarda taşıt kullanma eğiliminde oldukları ve daha şiddetli kazalara karıştıkları Birliktelik Kurallarında yer almıştır. Bédard ise çalışmasında taşıt yaşının, ölümlü kazalarda sürücü ve kaza ile istatistiksel olarak anlamlı bir ilişkisinin olmadığını ifade etmiştir (Bédard vd., 2002).
- Dönel kavşaklarda yaşanan kazaların %72'sinin yandan çarpma türünde olduğu bulunmuştur. Bu kazalarda sürücü, dönel kavşaktaki sürücüye yol vermeyerek çarpmaktadır. Bu durumun üç temel nedeni olabilir. Birincisi literatürde agresif sürücü olarak geçen sürücülerin trafik kurallarına uymaması durumu ile ilgili olabilmektedir (T. Özkan vd., 2010). İkincisi eğitim eksikliği ile ilişkilendirilebilir. Üçüncü neden, dikkatsizlik (ihmal) olarak ifade edilmektedir (Sümer & Özkan, 2002).

Trafik denetimlerinin ve cezaların artırılması yanında sürücü kurslarındaki teorik ve pratik eğitim süreçlerinin içerik olarak zenginleştirilmesi bu sorunların çözümünde olumlu etki meydana getirecektir (Factor, 2014; Paleti vd., 2010).

Birliktelik Kurallarından önce veri setinde yapılan veri ön işlemeye ek olarak, Karar Ağaçlarından önce de verinin ikinci bir ön işlemeden geçirilmesi sonucunda yalnızca otomobillerin karıştığı kazalara ait olarak elde edilen veri setinde, sürücü kaza şiddetini sınıflandırma tahmini amacıyla Karar Ağaçlarından CHAID kullanılmıştır. İlk olarak tüm değişkenlerin dahil edilmesi ile elde edilen bir Karar Ağacı söz konusu iken, ikinci olarak seçilen değişkenlerle ikinci bir Karar Ağacı elde edilmiştir.

4.2. Karar Ağaçları Sonuçlarının Literatürle Değerlendirilmesi

- İlk oluşturulan Karar Ağacında, kaza oluş türünün kök düğümde yer alarak sürücü kaza sonucunu en çok etkileyen öznitelik olduğu bulunmuştur. Montella ve ark., bisiklet kazalarında kaza şiddetini en iyi tahmin etmeyi sağlayan öznitelikleri bulmak için veri madenciliği yöntemlerinden olan CHAID ve Bayes Ağları'nı kullanmışlar ve benzer şekilde yol tipi, kaza oluş türü özniteliklerinin kaza şiddetini en çok etkileyen öznitelikler olarak bulmuşlardır (Montella vd., 2012).
- Bu çalışmada oluşturulan ilk Karar Ağacında kaza oluş türünden sonra sınıflandırmada etkisi olan özniteliklerden bir diğeri sürücü yaş grubudur. Wang ve Kim (2019), Maryland, ABD'de yaptıkları çalışmada Çok Terimli Lojit Model (MNL) ve Rassal Orman ile kaza şiddetini tahmin etmeye çalışmış olup kaza oluş

türü, sürücü yaşları ve hız limitini en büyük etkenler olarak ortaya çıkarmışlardır (Wang & Kim, 2019).

- Oluşturulan iki Karar Ağacında da, sürücü cinsiyeti ve saat dilimi de bu çalışmada kaza şiddetine etkisi olan öznitelikler olarak dikkat çekmektedir. De Oña ve ark., Granada, İspanya’da kırsal alanda gerçekleşen trafik kazalarının şiddetini tahmin etmek için Karar Ağaçları modelleri (CART, C4.5, ID3) kurmuşlar ve sürücü cinsiyeti ile zaman aralıklarının (0-6, 6-12, 12-18, 18-24) kaza şiddetine etkisi olduğunu bulmuşlardır (De Oña vd., 2013).
- Yine oluşturulan iki Karar Ağacında da, otomobil yaşı da bu çalışmada kaza şiddetinin sınıflandırılmasında etkili bir öznitelik olmuştur. Weast ve Monfort (2021), ölümlü kazalarda genç ve yetişkin sürücülerin kullandıkları taşıtların özelliklerini araştırmışlar, ölen genç sürücülerin yan hava yastıkları (hem baş hem de göğüs için) ve ESC (Electronic Stability Control=Elektronik Denge Kontrolü) gibi donatım özellikli taşıtlardan ziyade, daha az gelişmiş güvenlik özelliklerine sahip eski taşıtlar kullanma eğiliminde olduklarını, bu anlamda da bu tip taşıtlarla seyahat etme olasılıklarının yetişkinlerden fazla olduğunu belirtmişlerdir. Ayrıca çalışmada, otomobil yaşı arttıkça daha fazla kişinin hayatını yitirdiği, her iki sürücü grubu için de geçerli bir bulgu olmuştur (Weast & Monfort, 2021). Kadılar (2016) ise farklı olarak çalışmasında, sürücü cinsiyeti, yol yüzeyi, hava durumu ve taşıt yaşının kaza şiddeti üzerinde istatistiksel olarak anlamlı bir etkiye sahip olmadığını belirtmiştir (Kadılar, 2016).

5. SONUÇ VE ÖNERİLER

Trafik kazalarının oluşumuna birçok farklı etmen birlikte neden olabilmektedir. Kaza tutanaklarının raporlanmasından sonra Türkiye’de gerçekleşen trafik kaza kusurları insan (sürücü, yolcu, yaya) -yol-taşıt olarak sınıflandırıldığında, sürücü kusurlarının yaklaşık olarak %90 seviyelerinde olduğu görülmektedir. Türkiye’de, diğer ülkelerle kıyaslandığında yol kusurlarının etkisinin kaza raporlarında tam olarak gerçeği yansıtmadığı bilinse de sürücü kusurlarının tüm kusurlar içerisinde en büyük oranda olduğu bilinmektedir (NHTSA, 2018; Traffic Accidents in Germany, 2021; Yong-Seok Kim, 2002). Trafik kazalarının aynı anda birden fazla durumun örüntü halinde bir araya gelmesi ile gerçekleştiği dikkate alınarak yapılan bu çalışmada, kazaların meydana gelmesinde ortak etki gösteren ve birlikte yer alan etkenlerin bulunması, örüntülerin keşfedilmesi için veri madenciliği yöntemlerinden olan Birliktelik Kuralları kullanılmıştır.

Çalışma esnasında, Birliktelik Kurallarında kural keşfi için analize dahil edilen öznitelikler içinde grupların olması gerekliliği ortaya çıkmıştır. Bunun için öncelikle T. C. İçişleri Bakanlığı Emniyet Genel Müdürlüğü’nden alınan ham verinin ön işleme süreçlerinden geçmesi, analize hazırlanması gerekmektedir. Veri ön işleme süreçlerinden sonra veri setinde 11 tane öznitelik kalmıştır. Bu özniteliklerin tümünün bulunduğu veri setiyle analizler yapıldığında bazı değişkenlerin birbiriyle ilişkisi, veri setinde baskın olan özniteliklerden dolayı görülememektedir. Bunun için öznitelikler alt gruplara ayrılarak 5 farklı veri seti oluşturulmuştur. İlk veri setinde yalnızca kaza ile ilgili etkenler, ikinci veri setinde kaza ve çevresel etkenler, üçüncü veri setinde kaza ve sürücü ile ilgili etkenler bulunurken, dördüncü veri setinde kaza ve yol ile ilgili etkenler, beşinci veri setinde ise kaza ve taşıt ile ilgili etkenler yer almıştır.

Trafik işareti/işığına uymama durumu, sürücülerin kaza sonucunda ölmesi ya da yaralanmasıyla ilişkilidir (Karabulut & Özen, 2023). Ayrıca trafikte uyarı ve bilgilendirme levha/işaretlerinin sürücüler tarafından kolay ve doğru anlaşılabilmesi için standartların doğru bir şekilde uygulanması gerekmektedir (FHWA, 2009). Standartlarda iyileştirmeler yapılması ve her yerde aynı olması uygun olabilir.

KGM'nin yayınladığı Karayolu Trafik İşaretleme Standartlarında yol ver levhası yalnızca ters üçgen olarak verilmiştir (KGM, 2020). Sakarya'da da Türkiye'de birçok yerde olduğu gibi yol ver işaretleri bu standarttaki gibi olmakla beraber şehir içi yollarda bazı belediyeler yol ver işaretlerinin altına yol ver yazısı eklemektedir. Birleşik Krallık ve ABD gibi ülkelerde bu levhanın içinde yol ver anlamına gelen “give way” ya da “yield” yazmaktadır (Department for Transport, 2015; FHWA, 2009). Bu yazıların eksikliği sürücü farkındalığını azaltmaktadır. Türkiye'de de bu işaretlere yazı eklenerek standartların daha açık ve anlaşılır hale getirilmesi, sürücülerin yol verme davranışına uymaları açısından çok önemlidir.

Çalışmada öne çıkan durumlar aşağıda, maddeler halinde sıralanmıştır.

- Çalışmada, hız ve yakın takip kusuru önemli bir bulgu olarak ortaya çıkmaktadır. Trafik kanununda sürücülerin hızla ilgili uyması gereken kurallarda yakın takip kusurunun da beraber işlenebildiği, taşıt hızını sadece yoldaki hız sınırına göre değil, sürücünün sorumluluğunda olan birçok farklı etkene göre ayarlaması gerektiği görülmektedir.

2918 Sayılı Trafik Kanunu'nda 52. Madde, sürücülerin taşıt hızlarını;

1. Kavşaklara yaklaşırken, dönemeçlere girerken, tepe üstlerine yaklaşırken, dönemeçli yollarda ilerlerken, yaya geçitlerine, hemzemin geçitlere, tünellere, dar köprü ve menfezlere yaklaşırken, yapım ve onarım alanlarına girerken, hızlarını azaltmaları,
2. Kullandıkları aracın yük ve teknik özelliğine, görüş, yol, hava ve trafik durumunun gerektirdiği şartlara uydurmaları
3. Öndeki taşıtı izlerken yukarıdaki fıkrada belirlenen durumları göz önünde tutarak güvenli bir mesafe bırakmaları beklenmektedir (2918 sayılı Karayolları Trafik Kanunu, 1983).

Bu maddenin üç bendi de çok kapsamlı bilgiler içermekte, sürücülerin bu kanunun bilincinde olarak taşıt kullandıkları, bu kurallara uyacakları varsayılmaktadır. Bunun sonucu olarak hem hız hem de yakın takip kusurundan doğan arkadan çarpma kazaların önüne geçmek oldukça güçtür. Sürücülerin taşıt kullanım tarzlarından ve kendilerine sürüş konusunda fazla güvenlerinden doğan riskleri azaltmaya, gerekli durumlarda hızlarını düşürüp takip mesafesini koruyacakları inancından sıyrılmış önlemlere gereksinim

vardır. Bu önlemler, taşıtlardaki trafik güvenliğini arttırmaya yönelik donanımların zorunlu tutulması, yollarda fazladan önlemler almak (yol geometrisi ve işaretlemeler), yoldaki denetimler ve tüm yol kullanıcılarının (sürücü-yaya-yolcu) trafik eğitimlerinin arttırılması başta olmak üzere birçok kolu olan çalışmaların bütünüdür.

- Bu çalışmada elde edilen bir diğer önemli sonuç, kazadan sorumlu olan hatalı sürücülerin genellikle diğer taşıt sürücülerinin kendilerinden daha ciddi (ölüm veya yaralanma) derecede etkilenmelerine neden olmalarıdır. Bunun nedeni hatalı sürücülerin, yapmak üzere oldukları kazanın farkında olmaları nedeniyle, manevralarını, meydana gelecek kazadan en az etkilenecek şekilde ayarlamaları olabilir.
- Kavşaklarda gerçekleşen yandan çarpma türündeki kazalara en çok karışan yaş grubunun orta yaş 1 (26-40) yaş grubu olduğu bulunmuştur. Bu durum, esas olarak özellikle bu grup sürücülerin kavşaklarda risk alma davranışlarına bağlanabilir.
- Birliktelik Kurallarında orta yaş 1 (26-40) ve orta yaş 2 (41-64) gruplarının yandan çarpma türünde kazalara karışarak taşıtlarının hafif hasar alması ve sonucunda sağlam olarak kazadan çıktıkları durum tüm kazaların %11'idir. Bunun anlamı, bu yaş gruplarındaki sürücülerin yandan çarpma türünde kazalara karışarak başka bir kazazedenin yaralanması ya da hayatını kaybetmesine yol açtığıdır. Bu nedenle, özellikle orta yaş grubu (26-64) sürücülere yönelik bilinçlendirme kampanyaları da sürücülerin kural ihlallerinin ve bunlara bağlı olarak gerçekleşen trafik kazalarının azaltılmasında önemli rol oynayacaktır.
- Kaza oluş türü, düşey geçki geometrisi, haftanın günü, kazanın ortaya çıktığı mevsim, kazaya karışan otomobil sayısı, trafik ışığı varlığı, yatay geçki geometrisi, yerleşim yeri içi/dışı, yolun tipi ve yolun yüzeyi öznitelikleri çıkarıldıktan sonra oluşturulan ikinci Karar Ağacında, birinci Karar Ağacından farklı olarak; sürücü kusuru, yol bölümü, sürücü eğitim durumu ve gün ışığı/aydınlatma öznitelikleri Karar Ağacında yer almıştır. Bu tez içeriğindeki ana amaçlardan biri, trafik kazalarına karışan ve kusurlu olan sürücülere ait özelliklerin, kaza anındaki yolun durumunun (aydınlık/karanlık) ve hangi tip sürücülerin (yaş, cinsiyet, eğitim durumu) hangi yol bölümlerinde kazalara karıştıklarının ortaya çıkarılması olduğu için ikinci bir ağaç oluşturulmak istenmiştir. Oluşturulan bu ikinci ağaç ile

sürücüler ve sürücülerin farklı yol bölümü ve aydınlatma durumlarındaki karışıkları kazalar ile ilgili daha fazla açıklayıcı bilgi üretilebilmiştir.

Trafik kaza analizlerinde bilgilerin detaylı, güvenilir, tarafsız (kaza tutanakları tutulurken polis memurunun yorumunu içermemesi) olması elde edilecek sonuçlar açısından çok önemlidir. Bazı bilgilerin eksikliği çalışmada o öznelitliklerin kullanılamamasına neden olmaktadır. Bu durumlar aşağıda çalışmanın kısıtları olarak verilmiştir.

Çalışmanın kısıtları:

- Sürücülere ait yorgunluk, dikkatsizlik, agresif sürüş, cep telefonu kullanımı gibi durumlar, yeni taşıt teknolojileri sayesinde sürücü asistan sistemleriyle kontrol edilmeye çalışılmaktadır. Bu durumlara ait bilgiler veri setinde yer alamadığı için analiz süreçlerinde incelenememiştir. Kazaya neden olabilen diğer etkenler olarak bilinmelidir.
- Bu veri setinde kazaların %99'unda yol kusurunun bulunmadığı bilgisi vardır. Bu bilginin gerçeği yansıtmadığı, kazadan sonra yol hakkında daha detaylı veri girişi yapılabilmesi için kaza tutanaklarının bu doğrultuda geliştirilmesi, ileride yapılacak çalışmaların çıktılarına önemli katkılar sunacaktır.
- Maddi hasarlı kazaların da ölümlü-yaralanmalı kazalarla beraber incelenebilmesi, yaralanmalı kazalardaki yaralanma şiddeti derecesinin de tutanaklarda yer alması, bu tip çalışmaların kalitesini ve değerini arttıran bilgiler olacaktır.
- Sürücünün yaşının yanında sürücünün tecrübesi de trafik güvenliğinde önemli olmaktadır. Sürücülerin ehliyet aldığı yıl bilgisi de kaza tutanaklarında yer almalıdır.
- Ayrıca taşıtta çocuk koltuğu gibi ek koruyucu ekipmanların varlığı hem bu koltukta hem diğer koltuklarda emniyet kemeri kullanımı (sürücü-yolcu) bilgisi alınması çok önemlidir. Kaza tutanaklarında bu düzeyde bilgi alınması ileride yapılacak çalışmaların daha detaylı ve verimli olmasına katkı sağlayacaktır.
- Kaza şiddeti analizinde, trafik kazalarında en fazla yaralanan kişi olan sürücülerin kaza sonucuna göre inceleme yapılabilmektedir. Bunun nedeni veri setinde kazaların yalnızca binde ikisinin ölümlü olmasından dolayı kaza şiddeti

tahmini yapılamaması olmuştur. Yaya ve yolcularla ilgili incelemeler çalışma kapsamında yer almamaktadır.

Kaza sayısını öngören çalışmalarda trafik hacmi kaza analizlerinde önemli bir etken olarak dikkate alınmalıdır. Trafik hacminin yüksek olması trafik kaza sayısını arttırabilmektedir. Bu nedenle gündüz vakitlerinde daha fazla trafik kazası gerçekleşmesi olağan bir durum olmakla beraber bu çalışmada kullanılan yöntemlerde kazaların diğer özniteliklerle beraber sonuca etkisi araştırıldığından trafik hacminden bağımsız bir değerlendirme yapılması uygun görülmüştür.

Çalışmanın sonuçları, Sakarya'daki trafik kazalarının genel yapısı hakkında, birden fazla etkenin bir arada değerlendirildiği bilgiler vermektedir. Bu bilgilerin, yetkili kurumların trafik güvenliği konusunda alacağı önlemlerle ilgili katkı sunması beklenmektedir.

İleride yapılacak çalışmalarda bölgesel ve Türkiye genelinde de kaza analizleri yapılacak, şehir ve bölge bazlı benzerlik ve farklılıklar elde edilecektir. Ayrıca Jandarma Bölgesinde gerçekleşen trafik kazaları veri seti de temin edilerek analizler yapılacaktır.

Bu çalışmada kaza kara örüntüleri keşfedilmiş olup, konumdan bağımsız olarak çalışma sonlandırılmıştır. İleride yapılacak çalışmalardan biride kaza koordinatlarını da kullanarak yeni analizler yapmaktır. Bunun ilk adımı olarak bu tez kapsamında yazılan bilimsel araştırma projesinde Sakarya'daki bazı kaza kara noktalarında gözlemler yapılarak veri seti oluşturulmuştur. Bu noktalarda gerçekleşen kazalardaki kusur, geometrik unsur ve gözlemlerle elde edilen sürücü davranışları ile oluşturulan veri seti, bu çalışmada kullanılan veri seti ile birlikte analiz edilerek değerlendirmeler ve yorumlar yapılacaktır.

KAYNAKLAR

2918 sayılı Karayolları Trafik Kanunu (1983).

Adanu, E. K. (2021). Factors associated with driver injury severity of lane changing crashes involving younger and older drivers. *Accident Analysis and Prevention*, 8.

Aggressive driving enforcement. (2004). NHTSA.

Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. *Acm Sigmod Record*, 22(2), Article 2. <https://doi.org/10.1145/170036.170072>

Agrawal, R., Srikant, R., Road, H., & Jose, S. (1994). *Fast Algorithms for Mining Association Rules*.

Alavi, S. S., Mohammadi, M. R., Souri, H., Mohammadi Kalhori, S., Jannatifard, F., & Sepahbodi, G. (2017). Personality, Driving Behavior and Mental Disorders Factors as Predictors of Road Traffic Accidents Based on Logistic Regression. *Iranian Journal of Medical Sciences*, 42(1), Article 1.

Albuquerque, F. D. B. de, & Awadalla, D. M. (2020). Roadside Fixed-Object Collisions, Barrier Performance, and Fatal Injuries in Single-Vehicle, Run-Off-Road Crashes. *Safety*, 6(2), Article 2. <https://doi.org/10.3390/safety6020027>

AlKheder, S., AlRukaibi, F., & Aiash, A. (2020). Risk analysis of traffic accidents' severities: An application of three data mining models. *ISA Transactions*, 106, 213-220. <https://doi.org/10.1016/j.isatra.2020.06.018>

Anvari, M. B., Tavakoli Kashani, A., & Rabieyan, R. (2017). Identifying the Most Important Factors in the At-Fault Probability of Motorcyclists by Data Mining, Based on Classification Tree Models. *International Journal of Civil Engineering*, 15(4), 653-662. <https://doi.org/10.1007/s40999-017-0180-0>

Aoude, G. S., Desraj, V. R., Stephens, L. H., & How, J. P. (2012). Driver Behavior Classification at Intersections and Validation on Large Naturalistic Data Set. *IEEE Transactions on Intelligent Transportation Systems*, 13(2), 724-736. <https://doi.org/10.1109/TITS.2011.2179537>

Arthur Goodwin, Lewis Margolis, & Martha Waller. (2010). *ParentsTeensReport.pdf*.

Atalay, A. (2010). *Türkiye'deki Trafik Kazalarının Mekansal ve Zamansal Analizi*.

Augustynowicz, A. (2009). Preliminary classification of driving style with objective rank method. *International Journal of Automotive Technology*, 10(5), 607-610. <https://doi.org/10.1007/s12239-009-0071-8>

Bates, L., Evenhuis, A., & Lennon, A. (2020). Effectiveness of a pre-licensing driver education program on five psycho-social factors over twelve months. *Accident Analysis & Prevention*, 148, 105806. <https://doi.org/10.1016/j.aap.2020.105806>

- Batouli, G., Guo, M., Janson, B., & Marshall, W. (2020). Analysis of pedestrian-vehicle crash injury severity factors in Colorado 2006–2016. *Accident Analysis & Prevention*, 148, 105782. <https://doi.org/10.1016/j.aap.2020.105782>
- Bédard, M., Guyatt, G. H., Stones, M. J., & Hirdes, J. P. (2002). The independent contribution of driver, crash, and vehicle characteristics to driver fatalities. *Accident Analysis & Prevention*, 34(6), 717-727. [https://doi.org/10.1016/S0001-4575\(01\)00072-0](https://doi.org/10.1016/S0001-4575(01)00072-0)
- Cai, Q. (2020). Cause Analysis of Traffic Accidents on Urban Roads Based on an Improved Association Rule Mining Algorithm. *IEEE Access*, 8, 75607-75615. <https://doi.org/10.1109/ACCESS.2020.2988288>
- Chang, L.-Y., & Wang, H.-W. (2006). Analysis of traffic injury severity: An application of non-parametric classification tree techniques. *Accident Analysis & Prevention*, 38(5), 1019-1027. <https://doi.org/10.1016/j.aap.2006.04.009>
- Chiou, Y.-C., Fu, C., & Ke, C.-Y. (2020). Modelling two-vehicle crash severity by generalized estimating equations. *Accident Analysis & Prevention*, 148, 105841. <https://doi.org/10.1016/j.aap.2020.105841>
- Christopher Chabris, & Daniel Simons. (2018). *The Invisible Gorilla: And Other Ways Our Intuitions Deceive Us*. <http://www.theinvisiblegorilla.com/>
- Clifasefi, S. L., Takarangi, M. K. T., & Bergman, J. S. (2006). Blind drunk: The effects of alcohol on inattentive blindness. *Applied Cognitive Psychology*, 20(5), 697-704. <https://doi.org/10.1002/acp.1222>
- Constantinescu, Z., Marinoiu, C., & Vladiu, M. (2010). Driving Style Analysis Using Data Mining Techniques. *International Journal of Computers Communications & Control*, 5(5), 654. <https://doi.org/10.15837/ijccc.2010.5.2221>
- Çelik, A. K., & Oktay, E. (2014). A multinomial logit analysis of risk factors influencing road traffic injury severities in the Erzurum and Kars Provinces of Turkey. *Accident Analysis & Prevention*, 72, 66-77. <https://doi.org/10.1016/j.aap.2014.06.010>
- Çodur, M. Y. (2012). *Trafik Kaza Tahmin Modelleri: Erzurum İli Çevre Karayolları İçin Uygulamalar*.
- Das, S., Dutta, A., Jalayer, M., Bibeka, A., & Wu, L. (2018). Factors influencing the patterns of wrong-way driving crashes on freeway exit ramps and median crossovers: Exploration using ‘Eclat’ association rules to promote safety. *International Journal of Transportation Science and Technology*, 7(2), Article 2. <https://doi.org/10.1016/j.ijtst.2018.02.001>
- Das, S., Dutta, A., & Sun, X. (2020). Patterns of rainy weather crashes: Applying rules mining. *Journal of Transportation Safety & Security*, 12(9), Article 9. <https://doi.org/10.1080/19439962.2019.1572681>
- Das, S., Tamakloe, R., Zubaidi, H., Obaid, I., & Alnedawi, A. (2021). Fatal pedestrian crashes at intersections: Trend mining using association rules. *Accident Analysis & Prevention*, 160, 106306. <https://doi.org/10.1016/j.aap.2021.106306>

- Dawson, D., Sprajcer, M., & Thomas, M. (2021). How much sleep do you need? A comprehensive review of fatigue related impairment and the capacity to work or drive safely. *Accident Analysis & Prevention*, 151, 105955. <https://doi.org/10.1016/j.aap.2020.105955>
- De Oña, J., López, G., & Abellán, J. (2013). Extracting decision rules from police accident reports through decision trees. *Accident Analysis & Prevention*, 50, 1151-1160. <https://doi.org/10.1016/j.aap.2012.09.006>
- Department for Transport. (2015, Ekim 1). *The Highway Code—Traffic signs—Guidance—GOV.UK*. <https://www.gov.uk/guidance/the-highway-code/traffic-signs>
- Dünya Sağlık Örgütü. (2021). *Global Plan for the Decade of Action for Road Safety 2021-2030*. https://cdn.who.int/media/docs/default-source/documents/health-topics/road-traffic-injuries/global-plan-for-road-safety.pdf?sfvrsn=65cf34c8_35&download=true
- Dünya Sağlık Örgütü. (2022, Haziran 20). *Road traffic injuries*. <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>
- Dünya Sağlık Örgütü. *Mobile phone use: A growing problem of driver distraction*. (2023). <https://www.who.int/publications-detail-redirect/mobile-phone-use-a-growing-problem-of-driver-distraction>
- Eckerson, W. W., Hanlon, N., & Barquin, R. (2000). *Director of Education and Research* (4). 5(4), Article 4.
- Eibe Frank, Mark A. Hall, and Ian H. Witten. (2016). *The WEKA Workbench* (4. bs).
- Factor, R. (2014). The effect of traffic tickets on road traffic crashes. *Accident Analysis & Prevention*, 64, 86-91. <https://doi.org/10.1016/j.aap.2013.11.010>
- Feng, Y. R., Meuleners, L., Stevenson, M., Heyworth, J., Murray, K., Fraser, M., & Maher, S. (2021). Driving exposure, patterns and safety critical events for older drivers with and without mild cognitive impairment: Findings from a naturalistic driving study. *Accident Analysis & Prevention*, 151, 105965. <https://doi.org/10.1016/j.aap.2020.105965>
- FHWA. (2009). *MUTCD2009r1r2edition.pdf*.
- Fian Tabea, T. (2021). *From blackspots to blackpatterns: Pattern recognition with road traffic accident data. Illustrated with single-vehicle accidents with a single occupation and personal injury that occurred outside the built-up area on the Austrian road network between 2012 and 2019*.
- Funda TÜRE KİBAR. (2015). *Türkiye’de Kamyon Kazaları ile Trafik ve Karayolu Geometrik Özellikleri Arasındaki İlişkinin İstatistiksel ve Yapay Sinir Ağları Yöntemleri ile Modellenmesi*.
- Geurts, K., Thomas, I., & Wets, G. (2005). Understanding spatial concentrations of road accidents using frequent item sets. *Accident Analysis & Prevention*, 37(4), Article 4. <https://doi.org/10.1016/j.aap.2005.03.023>
- Google Earth. (2023).
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques Third Edition*. Elsevier. <https://doi.org/10.1016/B978-0-12-381479-1.00001-0>

- Hong, J., Tamakloe, R., & Park, D. (2020). Discovering Insightful Rules among Truck Crash Characteristics using Apriori Algorithm. *Journal of Advanced Transportation*, 2020, 1-16. <https://doi.org/10.1155/2020/4323816>
- Huang, P., & Winston, F. K. (2011). Young Drivers. İçinde B. E. Porter (Ed.), *Handbook of Traffic Psychology* (ss. 315-338). Academic Press. <https://doi.org/10.1016/B978-0-12-381984-0.10023-2>
- Ijaz, M., Lan, L., Zahid, M., & Jamal, A. (2021). A comparative study of machine learning classifiers for injury severity prediction of crashes involving three-wheeled motorized rickshaw. *Accident Analysis & Prevention*, 154, 106094. <https://doi.org/10.1016/j.aap.2021.106094>
- Injuries in the European Union 2009-2018. (2021). *Injuries in the European Union 2009-2018*. EuroSafe.
- Kadılar, G. O. (2016). Effect of driver, roadway, collision, and vehicle characteristics on crash severity: A conditional logistic regression approach. *International Journal of Injury Control and Safety Promotion*, 23(2), Article 2. <https://doi.org/10.1080/17457300.2014.942323>
- Karabulut, N. C., & Özen, M. (2023). Exploring Driver Injury Severity Using Latent Class Ordered Probit Model: A Case Study of Turkey. *KSCE Journal of Civil Engineering*, 27(3), Article 3. <https://doi.org/10.1007/s12205-023-0473-6>
- Karakuş, Y. İ. (2020). *Kentiçi Trafik Kazalarının Kümelenme Analizi ve Lojistik Regresyon Modeli ile İncelenmesi*.
- Kass, G. V. (1980). An Exploratory Technique for Investigating Large Quantities of Categorical Data. *Applied Statistics*, 29(2), 119. <https://doi.org/10.2307/2986296>
- Katanalp, B. Y., & Eren, E. (2020). The novel approaches to classify cyclist accident injury-severity: Hybrid fuzzy decision mechanisms. *Accident Analysis & Prevention*, 144, 105590. <https://doi.org/10.1016/j.aap.2020.105590>
- KGM. (2020). *KarayoluTrafikIsaretlemeStandartlari1.pdf*. KGM.
- KGM-Kaza kara noktaları. (2022, Mart 1). <http://yol.kgm.gov.tr/KazaKaraNoktaWeb/>
- KGM_Trafik Kazaları Özeti. (2022). KGM.
- Kong, X., Das, S., Jha, K., & Zhang, Y. (2020). Understanding speeding behavior from naturalistic driving data: Applying classification based association rule mining. *Accident Analysis & Prevention*, 144, 105620. <https://doi.org/10.1016/j.aap.2020.105620>
- Kotu, V., & Deshpande, B. (2014). *Predictive Analytics and Data Mining: Concepts and Practice with RapidMiner*. Morgan Kaufmann.
- Krstinić, D., Braović, M., Šerić, L., & Božić-Štulić, D. (2020). Multi-label Classifier Performance Evaluation with Confusion Matrix. *Computer Science & Information Technology*, 01-14. <https://doi.org/10.5121/csit.2020.100801>
- Leonard Evans. (1991). *Traffic safety and the driver*.

- Li, J., He, J., Liu, Z., Zhang, H., & Zhang, C. (2019). Traffic accident analysis based on C4.5 algorithm in WEKA. *MATEC Web of Conferences*, 272, 01035. <https://doi.org/10.1051/mateconf/201927201035>
- Li, T., Chen, D., Zhou, H., Laval, J., & Xie, Y. (2021). Car-following behavior characteristics of adaptive cruise control vehicles based on empirical experiments. *Transportation Research Part B: Methodological*, 147, 67-91. <https://doi.org/10.1016/j.trb.2021.03.003>
- Linoff, G. S., & Berry, M. J. A. (2011). *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management*. John Wiley & Sons.
- Lipovac, K., Đerić, M., Tešić, M., Andrić, Z., & Marić, B. (2017). Mobile phone use while driving-literary review. *Transportation Research Part F: Traffic Psychology and Behaviour*, 47, 132-142. <https://doi.org/10.1016/j.trf.2017.04.015>
- Ma, Z., Lu, X., Chien, S. I.-J., & Hu, D. (2018). Investigating factors influencing pedestrian injury severity at intersections. *Traffic Injury Prevention*, 19(2), Article 2. <https://doi.org/10.1080/15389588.2017.1354371>
- Maimon, O., & Rokach, L. (Ed.). (2005). *Data mining and knowledge discovery handbook*. Springer.
- Margaret H. Dunham. (2003). *Data Mining Introductory and Advanced Topics*.
- Montella, A. (2011). Identifying crash contributory factors at urban roundabouts and using association rules to explore their relationships to different crash types. *Accident Analysis & Prevention*, 43(4), Article 4. <https://doi.org/10.1016/j.aap.2011.02.023>
- Montella, A., Aria, M., D'Ambrosio, A., & Mauriello, F. (2012). Analysis of powered two-wheeler crashes in Italy by classification trees and rules discovery. *Accident Analysis & Prevention*, 49, 58-72. <https://doi.org/10.1016/j.aap.2011.04.025>
- Morgan, J.N., & Sonquist, J.A. (1963). Problems in the analysis of survey data, and a proposal. *Journal of the American statistical association*, 58(302), *Journal of the American statistical association*, 415-434.
- Neil A. Weiss. (2011). *Elementary Statistics (8th Edition)*. Addison Wesley.
- Orange (software). (2024). İçinde *Wikipedia*. [https://en.wikipedia.org/w/index.php?title=Orange_\(software\)&oldid=1193983524#External_links](https://en.wikipedia.org/w/index.php?title=Orange_(software)&oldid=1193983524#External_links)
- Orr, L. (1972). The Dependence of Transition Proportions in the Education System on Observed Social Factors and School Characteristics. *Journal of the Royal Statistical Society. Series A (General)*, 135(1), 74. <https://doi.org/10.2307/2345040>
- Özkan, Ö., & Öz, B. (2019). Sürücü Demografik Özelliklerinin Trafikte Kaçınma Davranışı ile İlişkisi. *Trafik ve Ulaşım Araştırmaları Dergisi*, 2(2), 108-122. <https://doi.org/10.38002/tuad.621367>

- Özkan, T., Lajunen, T., Parker, D., Sümer, N., & Summala, H. (2010). Symmetric Relationship Between Self and Others in Aggressive Driving Across Gender and Countries. *Traffic Injury Prevention*, 11(3), 228-239. <https://doi.org/10.1080/15389581003788864>
- Özkan, T., Lajunen, T., Parker, D., Sümer, N., & Summala, H. (2011). Aggressive driving among British, Dutch, Finnish and Turkish drivers. *International Journal of Crashworthiness*, 16(3), 233-238. <https://doi.org/10.1080/13588265.2010.536687>
- Paleti, R., Eluru, N., & Bhat, C. R. (2010). Examining the influence of aggressive driving behavior on driver injury severity in traffic crashes. *Accident Analysis & Prevention*, 42(6), Article 6. <https://doi.org/10.1016/j.aap.2010.05.005>
- Pande, A., & Abdel-Aty, M. (2009). Market basket analysis of crash data from large jurisdictions and its potential as a decision support tool. *Safety Science*, 47(1), Article 1. <https://doi.org/10.1016/j.ssci.2007.12.001>
- Prati, G., Pietrantonio, L., & Fraboni, F. (2017). Using data mining techniques to predict the severity of bicycle crashes. *Accident Analysis & Prevention*, 101, 44-54. <https://doi.org/10.1016/j.aap.2017.01.008>
- Redelmeier, D. A., & Tibshirani, R. J. (1997). Association between Cellular-Telephone Calls and Motor Vehicle Collisions. *New England Journal of Medicine*, 336(7), 453-458. <https://doi.org/10.1056/NEJM199702133360701>
- RStudio. (2024). İçinde Wikipedia. <https://en.wikipedia.org/w/index.php?title=RStudio&oldid=1193969134>
- Sebastian Land & Simon Fischer. (2012). *RapidMiner In Academic Use*.
- Senserrick, T., Boufous, S., Olivier, J., & Hatfield, J. (2021). At what stages of licensing do graduated driver licensing systems reduce crashes? Example from Queensland, Australia. *Accident Analysis & Prevention*, 152, 105989. <https://doi.org/10.1016/j.aap.2021.105989>
- Serbestlik derecesi (istatistik). (2024). İçinde Vikipedi. [https://tr.wikipedia.org/w/index.php?title=Serbestlik_derecesi_\(istatistik\)&oldid=31028953](https://tr.wikipedia.org/w/index.php?title=Serbestlik_derecesi_(istatistik)&oldid=31028953)
- Shahverdy, M., Fathy, M., Berangi, R., & Sabokrou, M. (2020). Driver behavior detection and classification using deep convolutional neural networks. *Expert Systems with Applications*, 149, 113240. <https://doi.org/10.1016/j.eswa.2020.113240>
- Singh, H., & Kathuria, A. (2021). Analyzing driver behavior under naturalistic driving conditions: A review. *Accident Analysis & Prevention*, 150, 105908. <https://doi.org/10.1016/j.aap.2020.105908>
- Sonquist, J. A., & Morgan, J.N. (1964). The detection of interaction effects: A report on a computer program for the selection of optimal combinations of explanatory variables. (No Title). <https://cir.nii.ac.jp/crid/1130000795534729728>
- Srikant, R., & Agrawal, R. (1995). *Mining Generalized Association Rules* (s. 26). *Stanford Machine Learning*. (2023). <http://www.holehouse.org/mlclass/index.html>

- Sümer, N., & Özkan, T. (2002). *Sürücü Davranışları, Becerileri, Bazı Kişilik Özellikleri ve Psikolojik Belirtilerin Trafik Kazalarındaki Rollerini*.
- Tangirala, S. (2020). Evaluating the Impact of GINI Index and Information Gain on Classification using Decision Tree Classifier Algorithm*. *International Journal of Advanced Computer Science and Applications*, 11(2). <https://doi.org/10.14569/IJACSA.2020.0110277>
- T.C. Milli Savunma Bakanlığı Harita Genel Müdürlüğü. (2023, Aralık 13). <https://www.harita.gov.tr/>
- T.C. İçişleri Bakanlığı Emniyet Genel Müdürlüğü-Cep Telefonu Kullanımı. (2023). <https://www.egm.gov.tr/cep-telefonu-kullanimi>
- The organisation and content of trauma memories in survivors of road traffic accidents | Elsevier Enhanced Reader. (2022, Kasım 22). <https://doi.org/10.1016/j.brat.2006.02.004>
- TÜİK-Trafik kaza istatistikleri. (2021). <https://cip.tuik.gov.tr/#>
- Türkiye İstatistik Kurumu. (2023). <https://data.tuik.gov.tr/Kategori/GetKategori?p=ulastirma-ve-haberlesme-112&dil=1>
- Ufuk Çelik, Eyüp Akçetin, & Murat Gök. (2017). *Rapidminer ile Uygulamalı Veri Madenciliği*.
- Ulaşan ve Erişen Türkiye. (2022). Ulaştırma ve Altyapı Bakanlığı.
- Vos, T., Lim, S. S., Abbafati, C., Abbas, K. M., Allebeck, P., Almasi-Hashiani, A., Alonso, J., Altirkawi, K. A., Alvis-Zakzuk, N. J., Amini, S., Amini-Rarani, M., Aminorroaya, A., Amiri, F., Amit, A. M. L., Amul, G. G. H., Anderlini, D., Andrei, T., Anjomshoa, M., Ansari, F., ... Murray, C. J. L. (2020). *Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019* (ss. 1204-1222). <https://linkinghub.elsevier.com/retrieve/pii/S0140673620309259>
- Wang, X., & Kim, S. H. (2019). Prediction and Factor Identification for Crash Severity: Comparison of Discrete Choice and Tree-Based Models. *Transportation Research Record: Journal of the Transportation Research Board*, 2673(9), 640-653. <https://doi.org/10.1177/0361198119844456>
- Watling, C. N., Mahmudul Hasan, M., & Larue, G. S. (2021). Sensitivity and specificity of the driver sleepiness detection methods using physiological signals: A systematic review. *Accident Analysis & Prevention*, 150, 105900. <https://doi.org/10.1016/j.aap.2020.105900>
- Weast, R. A., & Monfort, S. S. (2021). Characteristics of vehicles driven by teens and adults killed in crashes, 2013–2017. *Journal of Safety Research*, 77, 263-267. <https://doi.org/10.1016/j.jsr.2021.03.010>
- Wu, K.-F. (Ken), & Wang, L. (2021). Exploring the combined effects of driving situations on freeway rear-end crash risk using naturalistic driving study data. *Accident Analysis & Prevention*, 150, 105866. <https://doi.org/10.1016/j.aap.2020.105866>

- Wu, Y.-W., & Hsu, T.-P. (2021). Mid-term prediction of at-fault crash driver frequency using fusion deep learning with city-level traffic violation data. *Accident Analysis & Prevention*, 150, 105910. <https://doi.org/10.1016/j.aap.2020.105910>
- Xu, C., Bao, J., Wang, C., & Liu, P. (2018). Association rule analysis of factors contributing to extraordinarily severe traffic crashes in China. *Journal of Safety Research*, 67, 65-75. <https://doi.org/10.1016/j.jsr.2018.09.013>
- Yang, Y., Yi, F., Deng, C., & Sun, G. (2023). Performance Analysis of the CHAID Algorithm for Accuracy. *Mathematics*, 11(11), 2558. <https://doi.org/10.3390/math11112558>
- Yu, S., Jia, Y., & Sun, D. (2019). Identifying Factors that Influence the Patterns of Road Crashes Using Association Rules: A case Study from Wisconsin, United States. *Sustainability*, 11(7), Article 7. <https://doi.org/10.3390/su11071925>
- Zhang, G., Yau, K. K. W., & Chen, G. (2013). Risk factors associated with traffic violations and accident severity in China. *Accident Analysis & Prevention*, 59, 18-25. <https://doi.org/10.1016/j.aap.2013.05.004>
- Zhang, S., Abdel-Aty, M., Cai, Q., Li, P., & Ugan, J. (2020). Prediction of pedestrian-vehicle conflicts at signalized intersections based on long short-term memory neural network. *Accident Analysis & Prevention*, 148, 105799. <https://doi.org/10.1016/j.aap.2020.105799>
- Zhu, S. (2020). Investigation of vehicle-bicycle hit-and-run crashes. *Traffic Injury Prevention*, 21(7), Article 7. <https://doi.org/10.1080/15389588.2020.1805444>
- Zijian Zheng. (2018). *Applications of Data Mining Techniques In Transportation Safety Study*. North Dakota State University of Agriculture and Applied Science.

EKLER

EK A. Kaza Tutanağı

EK A

ÖLÜMLÜ/YARALANMALI TRAFİK KAZASI TESPİT TUTANAĞI

Sayfa 1/....

A. TUTANAĞI DÜZENLEYEN BİRİM ADI :		B. KONUM BİLGİSİ Koordinat X - E0 : Koordinat Y - N :	
TELEFON NO :	KAZA SIRA NO :	VERLEŞİM YERİ 1 İçi 2 Dışı :	
C. KAZANIN YERİ VE ZAMANI	YOLUN TİPİ 1 Bölünmüş yol 2 Tek yönlü yol 3 İki yönlü yol 4 Diğer (Kazanın Özeti Bölünmesine açıklama yapınız)	KAZA YERİNDEKİ AZAMI HIZ LİMİTİ : km/s	
TARİH :	YOLUN KAPLAMA CİNSİ 1 Asfalt 4 Parke 2 Sathi Kaplama 5 Stabilize 3 Beton 6 Toprak	ŞERİT SAYISI - GENİŞLİĞİ : m	
HAFTANIN GÜNÜ :	YOLUN SINIFI 1 Caddede 4 Devlet yolu 7 Orman yolu 10 Park alanı 13 Diğer (.....) 2 Sokak 5 İl yolu 8 Servis yolu 11 Tesis-mülk önü veya içi 3 Otoyol 6 Köy yolu 9 Bağlantı yolu 12 Su yolu taşıtı	YOL PLATFORM GENİŞLİĞİ : m	
SAAT/DAKİKA :	İLÇE :	YOL NO- KONTROL KESİM NO	
İL :	MAHL/KÖY :	Otoyol : O Devlet Yolu : D İl Yolu : Uzaklık : km mt	
KAZA YERİ ADRESİ (caddesi / sokağı) üzeri Öndü / yanı / arkası (caddesi / sokağı) ile (cad./sokağı) kavşağında (ilinden/üçesinden/mahallesinden/köy/öden) (ili / ilçesi/mah./köyü) yönüne km metrede		GÜN DURUMU 1 Gündüz 2 Gece 3 Alacakaranlık	
D. YOL GÜVENLİK EKİPMANLARI İLE ÇEVRE VE DİĞER ÖZELLİKLERİ 1 Var 2 Yok 3 Uygun Değil OTO KORKULUK : YAYA YOLU (Kaldırım) : cm EMNİYET ŞERİDİ / BANKET : cm YOL ŞERİT ÇİZGİSİ : TRAFİK İŞARET LEVHASI : Levha Adı: Kaza Nok. Uzaklık : 1) m 2) m 3) m Not: Uygun Değil seçeneğinin işaretlenmesi durumunda, Kazanın Özeti bölümünde tespiti ilişkin açıklama yapılmalı ve fotoğraflanmalıdır.		İŞIKLI / SESLİ İŞARET (Trafik Lambası) : 1 Var 2 Var (Bazı) 3 Yok AYDINLATMA : TRAFİK GÖREVLİSİ : 1 Var 2 Yok GÖRÜŞE ENGEL CİSİM : Var ise adı: KAZA SONRASI ARAÇ HARICİNDE HASAR GÖREN DİĞER ÜSURLAR : Var ise adı-1: Var ise adı-2: YOLDA ÇALIŞMA : Var ise; İşaretleme Var Yok Var ise; İşareti Personel Var Yok	
E. YOLUN GEOMETRİK ÖZELLİĞİ YATAY GÜZERGAH 1 Düz yol 2 Vinaj 3 Tehlikeli vinaj DÜŞEY GÜZERGAH 1 Eğimsiz 2 Eğimli 3 Tehlikeli eğim 4 Tepe üstü KAVŞAK 1 Üç yönlü (T) 5 Köprüli kavşak 2 Üç yönlü (Y) 6 Diğer kav. Çeşidi 3 Dört yönlü 7 Hemzemin geçit 4 Döneli kavşak 8 Kavşak yok		F. KAZAYA AİT ÖZELLİKLER OLUŞ ŞEKLİNE GÖRE KAZA TÜRÜ 8 Engel/cisim ile çarpışma 9 Yaygın çarpışma 10 Hayvana çarpışma 11 Devrileme, savrulma, tıklama 12 Yol kenarına çarpışma 13 Araçtan insan düşmesi 14 Araçtan cisim düşmesi 15 Park etmiş araç çarpışma 1- 2- ARAÇ SAYISINA GÖRE KAZA TÜRÜ 1 Tek araçlı 2 İki araçlı 3 Çok araçlı (Sayısı:)	
GEÇİT DURUMU 1 Kontrollü demiryolu 2 Kontrolsüz demiryolu 3 Okul geçidi 4 Yaya geçidi 5 Geçit yok		G. YOL SORUNU 4 Düşük banket 5 Yol sathında gevşek malzeme 6 Yolun münferit çıkışı 7 Diğer (*) 1 Tekerlek izinde oturma 2 Şerit çıkması 3 Kusurlu veya malzeme çıkması 4 Kazaya etken yol sorunu yok *Diğer seçeneğinin işaretlenmesi durumunda parantez içinde açıklama belirtilecek ve Kazanın Özeti bölümünde tespiti ilişkin gerekli açıklama yapılacaktır. Ayrıca yol sorununa ilişkin fotoğraflama yapılacaktır.	
DİĞER ÖZELLİK 1 Dar yol 2 Dar köprü 3 Köprü üstü 4 Köprü altı 5 Mesafe teli 6 Kazık 7 Tünel içi 8 Hışış		H. KAZA SONUCU KAZA ZEDE : ÖLÜ SAYISI : YARALI SAYISI : SÜRÜCÜ : VOLCU : YAYA : TOPLAM :	
YAYA GEÇİDİ ÖZELLİKLERİ 1 Üst geçit 2 Alt geçit 3 Sinyalizasyonlu geçit 4 Sinyalizasyonlu geçit 5 Diğer (.....) 6 Geçit yok		I. KAZAYA AİT FOTOĞRAF VE KAMERA KAYDI MEVCUT MU? A- EVET B- HAYIR 1 Fotoğraf 2 Kamera 3 Her ikisi de Mevcut Evete İse Açıklama Yazınız :	

Kaza Tutanağı A.1. Ölümlü-yaralanmalı trafik kazası tutanağı (Sayfa 1).

Sıra No Araç No		T.C. KİMLİK NO (Yabancı ise Ülke Adı ve Pasaport No)		DOĞUM YILI	KAZA ZEDE	KAZA SONUCU	YOLCU YAYA KURAL İHLALI	YAYA KAZA ÖNCESİ HAREKETİ	YAYA GYISI ÖZELLİĞİ	YAYA ALKOL KONTROLÜ	KONTROL NÖMRESİ (Pranid)	YOLCUNUN ARACI TA BULUN. YER	YOLCULAR İÇİN ENYET KANUNU (TUTU BAĞLANTI) SİSTEMİ VE KAZA DURUMU
		YAYA/YOLCU ADI SOYADI		DOĞUM YILI (R1)(R2)	1-Yolu 1-Yolu	1-Öm 1-Yolu							
1													
2													
3													
4													
5													
6													
7													
8													

YOLCU VE YAYALAR İÇİN KODLAR													
YAYA KAZA ÖNCESİ HAREKETİ				YAYA GYISI ÖZELLİĞİ		YAYA ALKOL KONTROLÜ		YOLCULARIN BULUNDUĞU YER				YOLCULAR İÇİN ENYET KANUNU (TUTU BAĞLANTI) SİSTEMİ VE KAZA DURUMU	
1 Karşıya Geçiyor 2 Duruyor 3 Yolda Yürüyor (Koşuyor) 4 Yaya Kalderiminde veya Yaya Dışında Duruyor/Yürüyor 5 Tespit Edilmedi				1 Açık 2 Kapa 3 Reflektör Malzemesi 4 Tespit Edilmedi		1 Trafik Zelen Korumu 2 Sağlık Kur. Kontrol Edildi 3 Kontrol Edilmedi		1 Ön Koltuk 4 Kaza İçi 2 Araç Koridor 5 Araç Dışı 3 Arka Koltuk 6 Tespit Edilmedi				1 Takılı 5 Kask Takılı 2 Takılı Değil 6 Kask Takılı Değil 3 Zorunlu Değil 7 Kask Yok 4 Belirli (Hem parsi kaza ve kask için)	

YAYALAR İÇİN TRAFİK KANUNU VE YÖNETMELİK MADDESİ İLE DİĞER KURAL İHLAL BİLGİSİ																					
1- 68/1-a-1		2- 68/1-a-2		3- 68/1-a-3		4- 68/1-b-1		5- 68/1-b-2		6- 68/1-c		7- 68/1-d		8- 68/1-e		9- 68/2		10- Yönet. Mad 138		11- Diğer (Kazanın Özet Bölümünde Belirtilen)	

YOLCULAR İÇİN TRAFİK KURAL İHLALI BİLGİSİ											
1 Taşıtın Sarkmak				2 Taşıt Dışında (Çamurluk v.s) Gayri Nizamî Seyahat Etmek				3 Taşıtlarda Kask veya Emniyet Kemeri Kullanılmamak			
4 Habersiz Taşıta Binmek ya da Taşıttan İnmek				5 Hareket Halindeki Taşıtlardan Yere Atılmak				6 Diğer (Kazanın Özet Bölümünde Belirtilen)			
7 Taşıt İçinde Gayri Nizamî Hareket Etmek				8 Ağız Yüklü Üzerinde Gayri Nizamî ve Tehlikeli Şekilde Seyahat Etmek							

Not: Kazazedelere ait doldurulması gereken bölümler tam ve doğru şekilde kodlanarak veya yazılarak doldurulacaktır. Özellikle T.C. Kimlik No (Yabancı ise Ülke Adı ve Pasaport No) bilgisi eksiksiz yazılacaktır. İlgili bölüme tespit edilemedi ya da bilinmiyor şeklinde metin yazılmayacak ve bu bölüm boş bırakılmayacaktır. Yaya/Yolculara ait kimlik bilgisinin tespiti için sevk edildikleri sağlık kuruluşu veya ilgili birimlerle irtibatla geçilerek tespit edilecek ve ilgili bölüme yazılacaktır.

M.KAZANIN ÖZETİ:											

Kaza Tutanağı A.3. Ölümlü-yaralanmalı trafik kazası tutanağı (Sayfa 3).

ÖZGEÇMİŞ

Ad-Soyad : Zeliha Çağla KUYUMCU

ÖĞRENİM DURUMU:

- **Lise** : 2009, Sakarya Anadolu Lisesi
- **Lisans** : 2013, Sakarya Üniversitesi, Mühendislik Fakültesi, İnşaat Mühendisliği
- **Yüksek Lisans** : 2017, İstanbul Teknik Üniversitesi, İnşaat Mühendisliği Anabilim Dalı, Ulaştırma Programı

YURT DIŞI DENEYİM:

- **Lisans** : 2013, Şubat-Haziran, Eötvös József Főiskola, Macaristan
- **Yüksek Lisans** : 2017, Şubat-Ağustos, Università degli Studi di Genova, İtalya

MESLEKİ DENEYİM:

- 2015-2016 yıllarında Yalova Üniversitesi Ulaştırma Mühendisliği Bölümünde araştırma görevlisi olarak çalıştı.
- 2016 yılının aralık ayından beri Sakarya Üniversitesi İnşaat Mühendisliği Bölümünde araştırma görevlisi olarak çalışmaktadır.

TEZDEN TÜRETİLEN ESERLER:

- Kuyumcu Zeliha Çağla, Aslan Hakan, Yurtay Nilüfer (2023). Identifying Interrelated Factors of Fatal and Injury Traffic Accidents Using Association Rules. Turkish Journal of Civil Engineering, 34, Doi: 10.18400/tjce.1322965 (Yayın No: 8517748)

DİĞER ESERLER:

- Çağlar Zeliha Çağla, Öğüt Kemal Selçuk (2015). Kent İçi Ulaşımında Taşıt Paylaşımının İncelenmesi Sakarya Üniversitesi Örneği. 11.Ulaştırma Kongresi(E/15/04), 559-568. (Tam Metin Bildiri/Sözlü Sunum) (Yayın No:2201234)
- Sarısoy Gürcan, Polat Ayşe, Çağlar Zeliha Çağla, Öğüt Kemal Selçuk (2015). Determination Minibuses Stop Delay in İstanbul. Sigma Journal of Engineering and Natural Sciences, 97-105., Doi: x (Yayın No: 3078113)-E-SCI.
- Çağlar Zeliha Çağla, Polat Ayşe, Sarısoy Gürcan, Öğüt Kemal Selçuk (2015). İstanbul da Park Et Bin Sisteminin İncelenmesi. Transist Uluslararası Ulaşım Teknolojileri Sempozyum ve Fuarı (Tam Metin Bildiri/Sözlü Sunum) (Yayın No:2286154)
- Sarısoy Gürcan, Polat Ayşe, Çağlar Zeliha Çağla, Öğüt Kemal Selçuk (2015). İstanbul da Minibüslerin Yolcu İndirme Bindirme Sürelerinin İncelenmesi. Transist Uluslararası Ulaşım Teknolojileri Sempozyum ve Fuarı (Tam Metin Bildiri/Sözlü Sunum) (Yayın No:2286721)
- Çağlar Zeliha Çağla, Polat Ayşe, Öğüt Kemal Selçuk (2016). İstanbul Yalova ve Sakarya Da Üniversite Öğrencilerinin Okul Yolculuklarında Bisiklet Kullanımının Araştırılması. Transist İstanbul Ulaşım Kongresi ve Fuarı (Tam Metin Bildiri/Sözlü Sunum) (Yayın No:3077302)
- Çağlar Zeliha Çağla, Aslan Hakan (2019). Türkiye’de Sürücü Hatalarına Bağlı Trafik Kaza Nedenlerinin Birliktelik Kuralları ile Analizi. International Symposium on Innovative Technologies in Engineering and Science 2019, Doi: 10.33793/acperpro (Tam Metin Bildiri/Sözlü Sunum) (Yayın No:5678680)
- Çağlar Zeliha Çağla, Aslan Hakan, Gözükızıl Gökçe (2019). Trafik Güvenliği Açısından Genç Sürücü Parametrelerinin Analizi. International Symposium on Innovative Technologies in Engineering and Science 2019, Doi: 10.33793/acperpro (Tam Metin Bildiri/Sözlü Sunum) (Yayın No:5678876)
- Çağlar Zeliha Çağla, Polat Alpan Mine, Aslan Hakan, Tanış Mustafa (2019). Investigation of The Effect of Demand and Inscribed Diameter on the Traffic Flow Patterns for Roundabouts: Case Study of Sakarya. ICEARC’19 International Civil Engineering and Architecture Conference (Tam Metin Bildiri/Sözlü Sunum) (Yayın No:5070965)
- Aslan Hakan, Yose Meutia Ayuni, Çağlar Zeliha Çağla (2020). Analyzing the Influence of Socio-Demographic Characteristics on Travel Mode Choice in Banda Aceh City Through Association Rules. International Symposium on Innovative Technologies in Engineering and Science 2020, 3(2), 703-714., Doi: 10.33793/acperpro.03.01.126 (Tam Metin Bildiri/Sözlü Sunum) (Yayın No:6900576)
- Kuyumcu Zeliha Çağla, Aslan Hakan, Yose Meutia Ayuni, Ahadi Suhrab (2020). Türkiye’de Trafik Kazaları ve Sürücülerin Kazalardaki Payı. International Symposium on Innovative Technologies in Engineering and Science 2020, 3(2), 694-702., Doi: 10.33793/acperpro.03.01.125 (Tam Metin Bildiri/Sözlü Sunum) (Yayın No:6900624)
- Çağlar Zeliha Çağla, Ahadi Suhrab, Aslan Hakan, Hashimi Sayed Naqibullah (2020). Investigation of the effect of Crumb Rubber Additive on the Fracture of characteristics of asphalt mixtures in control and asphalt rubber mixtures.

International Symposium on Innovative Technologies in Engineering and Science 2020, 3(2), 715-725., Doi: 10.33793/acperpro.03.01.127 (Tam Metin Bildiri/Sözlü Sunum) (Yayın No:6900642)

- Aslan Hakan, Aktaş Yusuf, Çağlar Zeliha Çağla (2020). Veri Madenciliği Uygulaması ile Sakarya Hane halkı Verileri Kullanılarak Yolculuk Analizlerinin Yapılması. International Symposium on Innovative Technologies in Engineering and Science 2020, Doi: 10.33793/acperpro.03.01.128 (Tam Metin Bildiri/Sözlü Sunum) (Yayın No:6900505)
- Kuyumcu Zeliha Çağla, Ahadi Suhrab, Aslan Hakan (2021). Analysis of Pedestrian-Vehicle Crashes Using Artificial Learning Methods: City of Sakarya Case Study. 5th International Symposium on Natural Hazards and Disaster Management, 4(2), 221-230., Doi: 10.33793/acperpro.04.02.54 (Tam Metin Bildiri/Sözlü Sunum) (Yayın No:7486559)
- Kuyumcu Zeliha Çağla, Aslan Hakan, Köse Muhammed Ali (2022). E-Skuter ve Bisikletin Kampüs Yolculuklarında Trafik Güvenliğine Etkisi: Sakarya Üniversitesi Örneği. International Symposium on Natural Hazards and Disaster Management 2022, Doi: 10.33793/acperpro (Tam Metin Bildiri/Sözlü Sunum) (Yayın No:7896706)
- Kuyumcu Zeliha Çağla, Turan İbrahim Can, Aslan Hakan (2022). Yolculuk Değerlerini Etkileyen Faktörlerin Veri Madenciliği Analizi ile İncelenmesi: Kahramanmaraş Örneği. International Symposium on Innovative Technologies in Engineering and Science 2022, Doi: 10.33793/acperpro (Tam Metin Bildiri/Sözlü Sunum) (Yayın No:7896694)
- Kuyumcu Zeliha Çağla, Kurtoğlu Halenur, Yıldız Özgenur, Kuruoğlu Serkan, Kepçe Âdem (2022). Veri Ön İşleme Sürecinin Türkiye Trafik Kaza Verilerine Uygulanması. International Symposium on Innovative Technologies in Engineering and Science 2022, Doi: 10.33793/acperpro (Tam Metin Bildiri/Sözlü Sunum) (Yayın No:7896678)
- Kuyumcu Zeliha Çağla, Aslan Hakan, Doğan Deniz Mert (2023). Motosiklet Kullanıcılarının Özellik ve Davranışlarının İncelenmesi Sakarya İli Motosikletli Kuryeler Örneği, 14. Ulaştırma Kongresi. (Tam Metin Bildiri/Sözlü Sunum)