



İSTANBUL MEDENİYET ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
İŞ SAĞLIĞI VE GÜVENLİĞİ ANABİLİM DALI

**Gemi İnşa Sektöründe İş Kazalarına Etki Eden
Faktörlerin Belirlenmesinde Makine Öğrenimi ve Küresel
Fuzzy-TOPSIS Yöntemlerinin Entegre Kullanımı:
Analitik ve Pratik Bir Yaklaşım**

Doktora Tezi

Fırat Aydın

Mayıs 2025



İSTANBUL MEDENİYET ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
İŞ SAĞLIĞI VE GÜVENLİĞİ ANABİLİM DALI

**Gemi İnşa Sektöründe İş Kazalarına Etki Eden
Faktörlerin Belirlenmesinde Makine Öğrenimi ve Küresel
Fuzzy-TOPSIS Yöntemlerinin Entegre Kullanımı:
Analitik ve Pratik Bir Yaklaşım**

Doktora Tezi

Fırat Aydın

Danışman

Prof. Dr. Emine Can

Eş Danışman

Doç. Dr. Neslihan Fidan Keçeci

Mayıs 2025

TEZ JÜRİSİ ONAYI

Fırat Aydın tarafından hazırlanan "Gemi İnşa Sektöründe İş Kazalarına Etki Eden Faktörlerin Belirlenmesinde Makine Öğrenimi ve Küresel Fuzzy-TOPSIS Yöntemlerinin Entegre Kullanımı: Analitik ve Pratik Bir Yaklaşım" başlıklı bu Doktora Tezi, İş Sağlığı ve Güvenliği Anabilim Dalı'nda hazırlanmış ve jürimiz tarafından kabul edilmiştir.

JÜRİ ÜYELERİ

İMZA

Tez Danışmanı:

Prof. Dr. Emine Can

İstanbul Medeniyet Üniversitesi

Üyeler:

Prof. Dr. Mithat Zeydan

İstanbul Medeniyet Üniversitesi

Prof. Dr. Hayriye Sundu Pamuk

İstanbul Medeniyet Üniversitesi

Prof. Dr. Prof. Dr. Berk Ayvaz

İstanbul Ticaret Üniversitesi

Doç. Dr. Adil Güler

Marmara Üniversitesi

Tez Savunma Tarihi: 29/Mayıs/2025

BEYANLAR

Yazım ve Kaynak Gösterme Kılavuzu Beyanı

Danışmanlığımda yazılan bu tezin APA yazım ve kaynak gösterme kılavuzunda belirtilen kurallara uygun olarak yapılandırıldığı ve bu kılavuzun APA kaynak gösterme standartlarının bu tezde tutarlı olarak uygulandığı tarafımdan incelenerek teyit edilmiştir.

İmza

Danışmanın Unvanı Adı Soyadı

Prof. Dr. Emine Can

Etik İlkeler Sadakat Beyanı

Hazırladığım bu tezin tamamen kendi çalışmam olduğunu, akademik ve etik kuralları gözeterek çalıştığımı ve her alıntıya kaynak gösterdiğimi beyan ederim.

İmza

Öğrencinin Adı Soyadı

Fırat Aydın

ÖZET

Gemi İnşa Sektöründe İş Kazalarına Etki Eden Faktörlerin Belirlenmesinde Makine Öğrenimi ve Küresel Fuzzy-TOPSIS Yöntemlerinin Entegre Kullanımı: Analitik ve Pratik Bir Yaklaşım

Aydın, Fırat

Doktora Tezi, İş Sağlığı ve Güvenliği Anabilim Dalı

Danışman: Prof. Dr. Emine Can

Mayıs 2025

Bu tez, gemi inşa sektöründe faaliyet gösteren tersanelerde meydana gelen iş kazalarının nedenlerini veri temelli ve analitik bir yaklaşımla incelemeyi amaçlamaktadır. Araştırmada, Türkiye’deki tersanelerde on yıl boyunca gerçekleşmiş 2020 iş kazasına ait, 17 değişken içeren bir veri kümesi kullanılmıştır. Hedef değişken kazanın nedeni olarak belirlenmiş ve veri ön işleme, etiketleme, dengeleme ve ölçekleme işlemleri uygulanmıştır. Modelleme sürecinde Random Forest (RF), Extra Trees (ET), XGBoost, SVM, KNN, Naive Bayes, Lojistik Regresyon, Karar Ağacı ve AdaBoost algoritmaları kullanılarak performans karşılaştırmaları yapılmıştır. En yüksek doğruluk ve başarı RF ve ET modelleriyle elde edilmiştir. Bu iki modelden elde edilen değişken önemlilik skorlarıyla en etkili dört faktör belirlenmiştir: Kazazedenin görevi, olay ayı, yaranın vücuttaki yeri ve kazanın türü. Bu faktörler, insan yargısını da içeren SF-Fuzzy TOPSIS yöntemiyle ayrıca değerlendirilmiş ve karar vericiler temelli alternatif bir sıralama oluşturulmuştur. Ardından, RF, ET ve SF-Fuzzy TOPSIS yöntemlerinin bilgi katkısı entropi tabanlı ağırlıklandırmayla hesaplanarak (%88,95 SF-Fuzzy TOPSIS, %7,31 RF, %3,73 ET) uzlaştırılmış nihai sıralama elde edilmiştir. Farklı ağırlık senaryolarıyla yapılan hassasiyet analizleri yöntemin sağlamlığını göstermiştir. Tez, makine öğrenmesi ile çok kriterli karar verme yöntemlerini birleştirerek sektöre özgü güçlü, esnek ve proaktif bir karar destek yapısı sunmaktadır.

Anahtar Kelimeler: İş kazaları, Makine öğrenmesi, Gemi İnşa Sektörü, SF-Fuzzy TOPSIS, Compromise, Entropi

ABSTRACT

Integrated Use of Machine Learning and Spherical Fuzzy-TOPSIS Methods in Determining the Factors Affecting Occupational Accidents in the Shipbuilding Industry: An Analytical and Practical Approach

Aydın, Firat

Ph. D. Thesis, Department of Occupational Health and Safety

Advisor: Prof. Dr. Emine Can

May 2025

This thesis aims to investigate the causes of occupational accidents in shipyards operating in the shipbuilding sector using a data-driven and analytical approach. The main objective is to estimate accident-causing factors based on historical data and prioritize them using multi-criteria decision-making (MCDM) techniques. A dataset comprising 2020 accidents and 17 variables, collected over ten years from Turkish shipyards, was utilized. The target variable was the cause of the accident. Various preprocessing steps such as labeling, balancing, and scaling were applied. Machine learning algorithms including Random Forest (RF), Extra Trees (ET), XGBoost, SVM, KNN, Naive Bayes, Logistic Regression, Decision Tree, and AdaBoost were implemented and compared using performance metrics. RF and ET achieved the highest accuracy. Based on variable importance scores from these models, the four most influential factors were identified: casualty's duty, month of the incident, injury location on the body, and type of accident. These factors were also evaluated using the SF-Fuzzy TOPSIS method, which incorporates human judgment, to obtain an alternative ranking. To reconcile results from different methods, entropy-based weighting was used, assigning weights as follows: SF-Fuzzy TOPSIS (88.95%), RF (7.31%), and ET (3.73%). A final ranking was produced through compromise scoring based on these weights. Sensitivity analysis with different weight scenarios confirmed the robustness of the approach. The findings show that machine learning models effectively predict accidents, and integrating these models with decision-making techniques like SF-Fuzzy TOPSIS offers a robust and flexible structure for sectoral decision-making. The thesis presents a proactive decision support framework for the shipbuilding industry by combining data science and occupational safety.

Keywords: Work Accidents, Machine Learning, Shipbuilding Industry, SF-Fuzzy TOPSIS, Compromise, Entropy

TEŞEKKÜR

Sevgili danışmanım sayın Prof. Dr. Emine Can hocama tez önerimden son ana kadar bana destek olduğu için teşekkürlerimi sunarım. Alana yeni bir şey katmak istememin heyecanını benimle tüm samimiyetiyle paylaştığı için ve yaşanan olumsuzluklarda hevesimi hiç kırmadan bana güç verdiği için kendisine minnettarım.

Bilgi ve donanımıyla bana hep destek olan, yönlendirmeleriyle tez sürecimi yürütmemi kolaylaştıran, her soruma titizlikle yanıt veren ve büyük katkı sunan eş danışmanım sayın Doç. Dr. Neslihan Fidan Keçeci hocama teşekkürlerimi sunarım.

Tez izleme komitemde yer alarak beni destekleyen sayın Prof. Dr. Mithat Zeydan ve sayın Prof. Dr. Berk Ayvaz hocama teşekkürlerimi sunarım.

Doktora maratonunun bu aşamasına ulaşabilmem konusunda bana çok yardımcı olan ve desteğini esirgemeyen çok değerleri dostlarım Sayın Hüseyin Enis Kara ve Sayın Murat Ramazan İltar'a teşekkürlerimi sunarım.

Akademik hayata başlamama ve çalışmalarımın ilk aşamasından bu aşamasına kadar desteklerini esirgemeyen sayın Prof. Dr. İdris Demir'e çok teşekkürler.

Fırat Aydın

İÇİNDEKİLER

TEZ JÜRİSİ ONAYI.....	III
BEYANLAR.....	IV
ÖZET	iv
ABSTRACT	v
TEŞEKKÜR	vi
İÇİNDEKİLER.....	vii
GRAFİK DİZİNİ	x
KISALTMALAR	xi
1. GİRİŞ.....	1
1.1. Tezin Amacı ve Kapsamı.....	1
1.2. Problemin Tanımı	5
1.3. Tezin Özgün Değeri ve Katkısı	6
1.4. Tezin Yapısı	6
BÖLÜM 2. KURAMSAL TEMELLER VE LİTERATÜR İNCELEMESİ	7
2.1 İş Kazası Tanımı ve Önemi	7
2.2 Gemi İnşa Sektöründe İş Kazaları	7
2.3. İş Kazalarına Etki Eden Faktörler	8
2.3.1. İnsan Faktörleri.....	10
2.3.2. Çevresel Faktörler.....	11
2.3.3 Teknik ve Operasyonel Faktörler	11
2.4 Örnek Olay İncelemeleri	12
2.5 Bulanık Mantık Yaklaşımı	12
2.6 Çok Kriterli Karar verme Metotları	13
2.6.1 TOPSIS Yöntemi.....	14
2.7. Fuzzy-TOPSIS Entegrasyonu	15
2.7.1 Küresel Fuzzy TOPSIS Metodu.....	15
2.8 Compromise (Uzlaştırma)	18
2.9 Entropi Yöntemi.....	19
2.10. Fuzzy–Makine Öğrenimi Entegrasyonu	20
BÖLÜM 3. ARAŞTIRMA METODOLOJİSİ	22
3.1 Veri Toplama Yöntemleri.....	23
3.1.1 Birincil Veri Kaynakları.....	23
3.1.2 İkincil Veri Kaynakları.....	24

3.2 Veri kümesi	24
3.2.1 Tanımsal İstatistikler	26
3.2.2 Veri Kümesi Özellikleri	30
3.2.3 Veri kümesi Hazırlama Süreci	30
3.2.4 Eksik Veri Yönetimi.....	31
3.2.5 Veri Normalizasyonu ve Dönüştürme	32
3.3. Makine Öğrenimi Modelleri.....	33
3.3.1. Sınıflandırma Problemlerinde Başarı Metrikleri	34
3.4. Küresel Fuzzy-TOPSIS Yönteminin Uygulanması (SF-TOPSIS)	34
3.4.1 Kriterlerin Belirlenmesi.....	35
3.4.2 Fuzzy Mantık ve TOPSIS Süreci	36
BÖLÜM 4. ANALİZ VE BULGULAR.....	40
4.1 Modelleme Süreci	40
4.2 Makine Öğrenimi Modellerinin Sonuçları.....	41
4.3 Fuzzy-TOPSIS Sonuçları.....	46
4.3.1 Kriter Ağırlıkları ve Sıralama	46
4.3.2 Entropi Tabanlı Ağırlıklandırma ve Uyumlaştırılmış Skorlar (Compromise Yaklaşımı)	47
4.4. Entegre Analiz ve Karşılaştırmalar.....	51
4.5 En Önemli Faktörlerin Belirlenmesi	52
BÖLÜM 5. TARTIŞMA	56
5.1. Bulguların Değerlendirilmesi.....	56
5.2. Metodolojik Değerlendirme ve Literatürle Karşılaştırma	57
5.3. Yöntemlerin Avantajları ve Dezavantajları	59
5.4. Pratik Uygulamalar ve Öneriler	59
BÖLÜM 6. SONUÇ VE ÖNERİLER	61
KAYNAKLAR	64
Ek 1.....	70
Ek-2	73
Ek 3.....	75

TABLO DİZİNİ

Tablo 1. Değişken Türleri ve Kategorik Dağılım Tablosu	25
Tablo 2. SF-Fuzzy TOPSIS'te Kullanılan En Etkili Dört Faktör.....	35
Tablo 3. Uzman Değerlendirme Matrisi	36
Tablo 4. Dilsel Terimler ve Bunlarla İlişkili Küresel Bulanık Sayılar.....	37
Tablo 5. Karar Vericilerin Önemi Derecelendirmek İçin Kullanılan Dilsel Terimler	37
Tablo 6. Uzmanların Küresel Bulanık Değerlendirme Matrisi	37
Tablo 7. Tüm Model Performansları.....	41
Tablo 8. Kaza Sınıfına Göre Extra Trees Modeli Performans Tablosu.....	42
Tablo 9. Optimize Extra Trees Modeli Karmaşıklık Matrisi.....	43
Tablo 10. Kaza Sınıfına Göre Random Performans Tablosu.....	44
Tablo 11. Optimize Random Forest Modeli Karmaşıklık Matrisi.....	45
Tablo 12. SF-TOPSIS Sıralaması	46
Tablo 13. Uyum analiz metrikleri.....	49
Tablo 14. Nihai Uzlaştırma Skorlarına Göre Sıralama	52
Tablo 15. Kriter Skorları.....	53
Tablo 16. En Önemli Dört Özellik.....	54

GRAFİK DİZİNİ

Grafik 1. Süreç Akış Diyagramı	22
Grafik 2. Yaranın Vücuttaki Yeri Sınıf Dağılımı.....	26
Grafik 3. Kazazedenin Görevi Sınıf Dağılımı.....	27
Grafik 4. Kaza Sayısına Göre Kaza Türleri.....	27
Grafik 5. Kazalı Sayısı Dağılımı	28
Grafik 6. Aylara Göre Kaza Dağılımı	28
Grafik 7.Görev Türüne Göre Kazalı Sayısı Dağılımı	29
Grafik 8. Kaza Türü Dağılımı.....	29
Grafik 9. Hassasiyet Analiz grafiği.....	49
Grafik 10. Yöntemlerin Uyum Analiz Grafiği.....	50

KISALTMALAR

ARAS Additive Ratio Assessment

ÇKKV Çok Kriterli Karar Verme

EDAS Ortalama Çözüm Uzaklığına Göre Değerlendirme

ESAW Avrupa İş Güvenliği

FML Fuzzy Machine Learning

HFACS Human Factor Analysis and Classification System

İSG İş Sağlığı Ve Güvenliği

KKD Kişisel Koruyucu Donanımlar

KNN K-En Yakın Komşu

NACE Ekonomik Faaliyet Sınıflaması

NumPy Sayısal Hesaplamalar Ve Çok Boyutlu Dizi İşlemler

ML Machine Learning

MOORA Multi-Objective Optimization by Ratio Analysis

Pandas Veri Çerçevesi Kütüphanesi

SAW Simple Additive Weighting

SCM Swiss Cheese Model

SF Spherical Fuzzy

SGK Sosyal Güvenlik Kurumu

SFN Spherical Fuzzy Numbers

SMOTE Synthetic Minority Over-sampling Technique

SVM Destek Vektör Makineleri

SWAM Küresel Ağırlıklı Aritmetik Ortalama

SWG Küresel Ağırlıklı Geometrik Ortalama

TOPSIS Technique for Order Preference by Similarity to Ideal Solution

TÜİK Türkiye İstatistik Kurumu

VIKOR VlseKriterijumska Optimizacija I Kompromisno Resenje

XAI Açıklanabilir Yapay Zekâ

1. GİRİŞ

1.1. Tezin Amacı ve Kapsamı

Gemi inşa sektörünün üretim üsleri olan tersanelerin ağır, yoğun ve mesleki kalifikasyon gerektiren imalat yapılan yerler olduğu bilinmektedir. Bu alanlarda yapılan işlerin elektrik, mekanik, boya, yüksekte çalışma gibi çok çeşitli olması buralarda fiziksel, kimyasal, biyolojik, ergonomik ve dolayısıyla Psikososyal risk etmenleri oldukça fazladır. Bunun yanı sıra Türkiye’de “30.1- Gemi ve Tekne Yapımı” iş kolunda gerçekleştirilen projelerin çok büyük bir kısmında kalifikasyon, birden fazla iş türü olması ve başka bir çok nedenle işlerin çoğunluğu alt taşeron firmalara yaptırılmakta bu durum iletişim, kontrol ve koordinasyon eksikliği yaratmakta olup iş kazalarının ve meslek hastalıklarının neden olan etkenlerin bertarafını olumsuz etkilemektedir (Ayvaz ve Ergör, 2023). Tersanelerde yaşanan iş kazaları, insan yaralanmaları ve ölümlerin yanında ciddi ekonomik kayıplara sebep olabilmektedir. Bu kayıplar birçok ülke ve tersane işletmeleri için çok ciddi sonuçlar doğurmaktadır. Çünkü tersaneler birçok üretim ve sanayi alt kolunu desteklediği gibi savunma sanayi ile istihdam içinde önem arz etmektedir. İş kazaları görevlerle bağlantılı olarak ortaya çıkmakta ve çalışanların yaralanmasına, hastalanmasına veya ölümüne neden olabilmektedir. Sanayileşmiş ülkelerde ve gelişmekte olan ülkelerde çalışanların yaklaşık yarısının işyerlerinde ölümcül yaralanma riskine maruz kaldığı raporlanmıştır (World Health Organization, 1995). Gemi inşa, bakım ve onarım alanlarında yaşanacak iş kazaları çalışanların psikolojik ve davranışsal durumlarında yıkıcı değişikliklere neden olmakta, iş sırasında travmaya neden olmuş deneyimler güvensizlik hissi gibi diğer ilgili sağlık sorunlarına neden olabilmektedir. Bu durum sonuç olarak iş performansını bozacak, işin ilerlemesini geciktirecek, proje yürütmede zaman kaybı ve üretkenlik kaybına neden olacaktır (Aziz ve ark., 2018).

İş kazaları ile ilgili istatistiki veriler meslekten ve mesleğin icra edildiği alanlardan kaynaklı yaralanma ile olayların nasıl meydana geldiğine dair ayrıntılı bir açıklama sağlayabilir. Gemi inşa, bakım ve onarım yapılan tersanelerde yaşanan iş kazalarının analizinin yapılarak öngörüle bulunulması, kaza olmasını etkileyen faktörlerin gelecekte yaşanabilecek kazaların tahmini ve önlem alınmasının sağlanması bu nedenle önemlidir. İş kazaları ile ilgili istatistiki veriler meslekten ve mesleğin icra edildiği alanlardan kaynaklı yaralanma ile olayların nasıl meydana geldiğine dair ayrıntılı bir açıklama sağlayabilir. Gemi inşa, bakım ve onarım yapılan tersanelerde yaşanan iş kazalarının

analizinin yapılarak öngöründe bulunulması, kaza olmasını etkileyen faktörlerin gelecekte yaşanabilecek kazaların tahmini ve önlem alınmasının sağlanması bu nedenle önemlidir. Özellikle iş kazası ile ilişkili istatistik ve verilerden bilgi elde edilmesi, iş kazalarının varlığını azaltmak için bu veriler ile yoğun araştırmalar hayati önem taşımaktadır. Bu önem derecesine binaen meslek ve mesleğin uygulandığı alanlardan elde edilecek verilerin analizinde genel tanımlayıcı bir istatistiksel analiz, karar ağaçları, genelleştirilmiş doğrusal model ve bulanık sinirsel yöntemler, mesleki yaralanma oranlarını ölçmek için Bayes ağ tekniklerini, mesleki yaralanma riski değerlendirmeleri için Poisson modelleri gibi yöntemler kullanılmıştır (Khairuddin ve ark., 2022). Ancak, olaylar ve zaman geçtikçe, istatistik veriler genellikle önemsiz sayılır ve herhangi bir takip ve analiz işlemleri derin ve yeterli olarak yapılmaz. Bu nedenle kuruluşlar ve sektörlerin kaza önleme ve iş sağlığı güvenlik göstergelerinin çoğu geriye dönük olarak kalmaktadır. Sosyal Güvenlik Kurumu (SGK) istatistiklerinde Türkiye’de gemi inşa sektörü Ekonomik Faaliyet Sınıflamasında (NACE) “30-Diğer Ulaşım Araçları İmalatı” ana iş kolunda yer almaktadır. Gemi inşa sektörü istatistiki verileri “30.1- Gemi ve Tekne Yapımı” alt iş kolunun altında yer alan “30.1.1- Gemilerin ve yüzen yapıların inşası” ile “30.1.2- Eğlence ve spor amaçlı teknelerin yapımı” alt iş kollarına ayrılmaktadır. T.C. Ulaştırma ve Altyapı Bakanlığının Kasım 2023 verilerine göre Türkiye’de Gemi inşa, bakım ve onarım faaliyeti yürüten tersane sayısı 85’tir. Bu sayının azlığına rağmen SGK istatistiklerinde “30.1- Gemi ve Tekne Yapımı” iş kolu yer alan sadece 4-1/a kapsamında ki sigortalılar arasında yaşanan iş kazaları 2020 yılında 6.785, meslek hastalıkları sayısı 19, 2021 yılında 8.598 iş kazası 45 meslek hastalığı, 2022 yılında ise 8.771 iş kazası 36 meslek hastalığı vakası gerçekleşmiş ve artış sergilemiştir.

Türkiye’deki tersane tarihine bakıldığında istatistiki anlamda büyük veriler elde edilebilmektedir. Bu veriler ile uygun strateji ve bilgi birikimiyle birlikte, sektörler kendilerini tahmin etme yeteneği ile donatmalı ve bu tür olayların gelecekteki oluşumlarını öngörebilmelidir. Bu, ilgili ekosisteme ve endüstrilere yeni katma değer sağlarken, mesleki ve melek uygulama alanlarındaki iş kazası riski öngörerek uygunsuz alanlara ve eğitim yetersizliğine sahip çalışanlara odaklanmayı sağlayacaktır. Bu nedenle İSG risklerinin değerlendirilmesinde büyük verilerin ve çoklu değişkenlerin işlenebildiği makine öğrenimi teknolojisinden yararlanılması gerekli görülmektedir. Makine öğrenimi (ML) ve yapay zekâ teknolojisi paralel işleme özelliği ve yüksek boyutlu verileri verimli bir şekilde işleyebilmesi nedeniyle gelişmekte olan bir teknolojidir. Büyük veriler,

gelecekteki eğilimleri veya olayları belirlemek için özellikle keşif, tanımlayıcı, tahmin ve kural koymada belirleyici olmasından ötürü çok önemlidir.

Bu çalışma, mesleki yaralanma sonuçlarını ve bunları etkileyen faktörleri daha tutarlı, sağlam ve zamanında öngörebilmek amacıyla, makine öğrenimi temelli bir tahmin modeli geliştirir. Kodlar, Google Colab ortamında Python 3.10 kullanılarak (Python Software Foundation) yürütülmüştür. Model, Türkiye’deki gemi inşa, bakım ve onarım faaliyetlerine ilişkin kaza verilerini işler; böylece kaza sebepleri ile kritik odak noktalarını tanımlar. Elde edilen tahminler, SF-TOPSIS yöntemiyle hesaplanan uzman öncelikleriyle karşılaştırılır. Böylelikle kaza sebebi öngörülerinin doğruluğu ile kazalara etki eden faktör sıralamasını belirleme yetkinliği bütüncül biçimde değerlendirilir. Bu bütünlük yaklaşım, yönetsel stratejilerin iyileştirilmesine ve iş güvenliği performansının sürdürülebilir biçimde artırılmasına katkı sunar.

Türkiye’de faaliyet gösteren iş yerlerinin faaliyet konularına göre SGK tarafından belirlenen bir kodlama sistemi ile tanımlanmıştır. Bu kodlama sistemi, işyerlerini belirli sektörler ve alt sektörler sınıflandırmak için kullanılır. Bu çalışma kapsamında “30-Diğer Ulaşım Araçları İmalatı” ana iş kolunda yer almakta olan “30.1- Gemi ve Tekne Yapımı” alt iş kolunun altında yer alan “30.1.1- Gemilerin ve yüzen yapıların inşası” ile “30.1.2- Eğlence ve spor amaçlı teknelerin yapımı” iş kollarında yer alan tersaneler yer almaktadır. “30.1- Gemi ve Tekne Yapımı” iş kolunda gerçekleşen 2020 kazanın verileri kullanılacaktır. Bu 2020 kaza verisi için 17 ayrı ana veri faktörü ve bunlara ait 200 alt veri kümesi kullanılacaktır. Mevcut büyük miktarda veri sayesinde makine öğrenimi, mühendislik, tıp bilimi, finans gibi çeşitli alanlarda kullanılan gelecekteki olayları tahmin etmede geleneksel istatistiksel karşılığın yerini alır ve çok faydalı sonuçlar verir (Sarkar ve ark., 2019). Ancak literatür incelendiğinde makine öğrenimi tekniklerinin iş kazası analizlerinde sınırlı olarak kullanıldığı görülmektedir (Bevilacqua ve ark., 2008). Şimdiye kadar mesleki analiz üzerine yapılan çalışmalar, tahmin gücü ve açıklama kapasitesi açısından makine öğrenimi tekniklerinin kullanıldığını göstermektedir. Gerçekleşmiş kazalardan elde edilen tarihsel verilere dayanan makine öğrenimi yöntemleri, tahmin fonksiyonları ve olay sonuçları üzerinde etkisi olan tahmin edicilerin önemi açısından geleneksel istatistiklere göre avantaj sağlar. Makine öğreniminin potansiyel faydaları, yalnızca büyük miktarda veriyi işleme kapasitesinden değil, aynı zamanda: büyük boyutlu problemlerle başa çıkma yeteneklerinden, karmaşıklıktan bağımsız olarak veri üretim yapısını yeniden üretmedeki esnekliklerinden ve kuralların çıkarılması yoluyla

öngörücü ve yorumlayıcı potansiyelleri ve makine öğrenimi tekniklerinin kapasitesi nedeniyle, iş kazası analizleri de dahil olmak üzere birçok alanda başarıyla kullanılmaktadır (Khairuddin ve ark., 2022).

Python kullanarak yapılacak bir tahmin modellemesi, makine öğrenimi veya istatistiksel yöntemler kullanarak verileri analiz ederek belirli bir olayın veya değerin tahmin edilmesini amaçlar. Bu modellemeler, çeşitli Python kütüphaneleri ve araçları kullanılarak gerçekleştirilebilir. Tahmin modellemesi, verilerin analiz edilmesi ve bir hedef değişkenin gelecekteki değerlerinin tahmin edilmesi için bir matematiksel veya istatistiksel modelin oluşturulması sürecidir. Model, mevcut verilerden öğrenir ve bu öğrenmeyi kullanarak yeni veriler üzerinde tahminler yapar. Python kullanılarak yapılan tahmin modellemesi genellikle makine öğrenimi tekniklerini içerir. Python, tahmin modellemesi için birçok popüler kütüphaneye sahiptir. Bunlar arasında en yaygın olanları NumPy (Sayısal hesaplamalar ve çok boyutlu dizi işlemleri için kullanılır), Pandas (etkili bir veri çerçevesi kütüphanesi) (Bullejos ve ark., 2022), Scikit-learn (Makine öğrenimi algoritmalarını içeren geniş kapsamlı bir kütüphane) TensorFlow ve Keras'dır. (Derin öğrenme ve yapay sinir ağı modelleri için kullanılan popüler kütüphane) Statsmodels'dır. (Regresyon analizi, zaman serisi analizi ve hipotez testleri gibi istatistiksel yöntemler sunar) (Li ve ark., 2019).

Makine öğrenimi ile yapılacak tahmin modellemesi matematiksel bir yöntem olarak kullanılırken, iş sağlığı ve güvenliğinde insan yargılarının değerlendirmeye katılması, risklerin ve kazalara etki faktörlerin belirlenmesinde daha gerçekçi bir yaklaşım sunacaktır. Tahmin algoritmaları ile elde edilen sonuçların bulanık mantık ile desteklenmesi belirsizliklerin ele alınmasını da sağlayacaktır. Bu amaçla kullanılan Küresel Bulanık TOPSIS yöntemi, bulanık küme teorisi prensiplerini kullanmaktadır. Bu yöntemde, karar verme sürecindeki kriterlerin belirsizliği veya belirsizlikleri, bulanık küme teorisi ile temsil edilir. Kriterlerin belirsizliği, her bir kriterin bulanık küme olarak ifade edilmesiyle modellenir. Böylece, her bir kriterin alt ve üst sınırları belirlenerek belirsizlik ifade edilir. Karar vericilere daha geniş bir seçim alanı sağlamak amacıyla yeni bir bulanık küme olan küresel bulanık sayılar kavramı geliştirilmiştir (Kutlu Gündoğdu ve Kahraman, 2020). Küresel bulanık küme ve çok kriterli karar verme (ÇKKV) tekniklerini birleştirmek, iş kazalarının nedenlerini belirlenmesine yardım edecek yaklaşımlara ışık tutacaktır. Bu ÇKKV teknikleri TOPSIS, AHP, ELECTRE

PROMETHEE yöntemleridir (Sikalo ve ark., 2023). Bu çalışmada TOPSIS yöntemi tercih edilmiştir.

Sonuç olarak, Python ile geliştirilen tahmin modeli çıktıları uzman görüşlerine dayalı Küresel Bulanık Alternatifler sıralaması, İdeal Çözüme Benzerlik Yoluyla Tercih Sıralaması metodu ile desteklenecektir. Bu sayede veriler analiz edilirken, gelecekteki kaza nedenleri tahmin edilecek ve kazalara ait değişkenlerin hangi ölçüde etkili olduğu saptanacaktır. Hem makine öğrenmesi hem de küresel Fuzzy yöntemi kullanılarak yapılan karşılaştırma entegre sonuçlar sağlayarak yöntemin tahmin performansını pekiştirecektir.

1.2. Problemin Tanımı

Gemi inşa sektöründe kalifikasyon gerektiren çok fazla alana yönelik uygulama yapılması ve yaşanan iş kazaları ile meslek hastalıklarında yıllara göre artış görülmektedir. Bu artış iş kazalarının önlenmesine yönelik çalışmanın yapılması gerekliliğini göstermektedir. Türkiye’de tersanelere ilişkin iş kazası verileri oldukça sınırlıdır, bazı tersanelerde 2010 yılından itibaren iş kazası verileri kayıt altına alınmakta, kaza verilerinin kayıt altına alınmasına ilişkin bir standart bulunmamaktadır. Her tersanenin veri toplama kriterleri vardır. Kaza verileri, verilerdeki kalıpları kategorize etmek ve sınıflandırmak için birçok farklı bağlamda kullanılabileceği belirtilmiştir (Izci ve ark., 2023). Bunun yanı sıra gemi inşası iş kolunda makine öğrenimi yöntemleri ile yapılmış çalışmaların çok kısıtlı olması, yaşanan iş kazalarının ve kazaları değerlendirilirken etken faktörlerin sıralanması için makine öğrenimi yöntemleri ile küresel bulanık TOPSIS yönteminin bir arada kullanıldığı (Lin, Shen, Zhou, & Xu, 2021) çok az çalışma olması, “30.1-Gemi inşası” iş kolunda 2016-2022 yıllarında yaşanan İş kazalarının sayısı, iş kazasının gerçekleştiği zaman, iş kazasının nedeni, kazazedenin yaşı, kaza türü, kazazedenin görevi, yaranın vücuttaki yeri, iş kazasına uğrayanın cinsiyeti gibi farklı veri setlerinin kullanılması ile birlikte bu verilerin etki derecelerinin neler olduğunun bilinmesi tedbir stratejileri belirlenmesini kolaylaştıracağı gibi, veri toplama ve veri çeşitliliğinin belirlenmesini de önem arz eden unsurlardan biridir.

Bu çalışma, iş kazalarının yaş, cinsiyet, kaza nedeni gibi çeşitli faktörlere göre dağılımını inceleyerek, risk gruplarını ve kazaya neden olan temel etkenleri belirlemeyi amaçlamaktadır. Makine öğrenimi modelleriyle kazaların nedenlerini tahmin etmek ve SF-Fuzzy-TOPSIS yöntemiyle bu kazalara etki eden faktörleri önceliklendirmek, farklı yaklaşımların karşılaştırmalı analizine olanak tanımaktadır. Bu iki yöntemin

entegrasyonu sayesinde hem tahmin başarısı hem de karar destek süreci güçlenmekte, kazaları önlemeye yönelik stratejiler geliştirilebilmektedir. Elde edilen bulgular, gemi inşa sektöründe riskli durumları ve odak noktalarını belirleyerek literatüre özgün katkılar sunmayı hedeflemektedir.

1.3. Tezin Özgün Değeri ve Katkısı

Analiz sonuçları ve önerilen önleyici stratejiler, “30.1.1- Gemilerin ve yüzen yapıların inşası” iş kolunda İSG politikalarının iyileştirilmesi için temel bilgilere odaklanılmasını sağlayabilir. Bunun yanı sıra tahmin modellemesinde ortaya çıkacak sonuçlar çerçevesinde risk değerlendirmesi ve yönetimi, işyeri düzenlemesi ve ergonomi, kişisel koruyucu donanımlar (KKD) gibi konulara dikkat edilerek kaza ve kazalardan etkilenimin azalmasına ve farklı disiplinlerden gelen bu iki yöntemin birleştirilerek kullanılması çalışmanın ve yöntemin etkinliği arttıracaktır. Bunun yanında İş kazalarının sayısı, iş kazasının gerçekleştiği zaman, iş kazasının nedeni, kazazedinin yaşı, kaza türü, kazazedinin görevi, yaranın vücuttaki yeri, iş kazasına uğrayanın cinsiyeti gibi farklı veri setlerinin kullanılması tezin özgün yanlarından biridir.

1.4. Tezin Yapısı

Tez, ana metin kısmında altı bölümden ve bunlara ek olarak Ekler başlıklı yedinci bir kısımdan oluşur. Birinci bölüm giriş niteliğinde olup tezin amacı, kapsamı, problem tanımı, özgün değeri ve genel yapısını açıklar. İkinci bölüm kuramsal temelleri ve literatür incelemesini sunar; iş kazası kavramından gemi inşa sektöründeki kaza dinamiklerine, makine öğrenimi ve bulanık mantık yaklaşımlarından TOPSIS ve entropi yöntemlerine kadar çalışmanın teorik dayanaklarını sistematik biçimde anlatır. Üçüncü bölüm araştırma metodolojisini ele alır; veri toplama süreçleri, veri setinin özellikleri, ön-işleme adımları, makine öğrenimi modellerinin kurulumu ve küresel Fuzzy-TOPSIS prosedürü ayrıntılarıyla açıklanır. Dördüncü bölümde analizler ve bulgular yer alır: makine öğrenimi modellerinin performans sonuçları, SF-TOPSIS çıktıları, entropi-temelli ağırlıklandırma ile elde edilen uzlaştırılmış skorlar ve bütünlüklü değerlendirmeler bu kısımda sunulur. Beşinci bölüm tartışmaya ayrılmıştır; bulgular literatürle karşılaştırılır, kullanılan yöntemlerin güçlü ve zayıf yanları irdelenir, uygulamaya yönelik öneriler geliştirilir. Altıncı bölümde ise çalışmayı özetleyip ana sonuçları ve ileri araştırma önerilerine değinilmiştir. Son olarak Yedinci kısım (Ekler), ayrıntılı verilere, ilave tablolara ve destekleyici materyallere için ek dokümanları içerir.

BÖLÜM 2. KURAMSAL TEMELLER VE LİTERATÜR İNCELEMESİ

2.1 İş Kazası Tanımı ve Önemi

İş yerinde veya iş sırasında ortaya çıkan; ölümcül veya ölümcül olmayan olay olarak tanımlanmaktadır. İş kazası, işyerinde veya işin yürütümü esnasında kaza sonucu oluşan ölüm, herhangi bir kişisel yaralanma veya hastalık anlamına gelir. İş kazası terimi genellikle ani, dışsal ve istem dışı bir olay olarak anlaşılır (Hämäläinen ve ark., 2006). Bunun yanı sıra İşyeri kazaları, ekonomik faaliyette meydana gelen bir veya daha fazla işçinin yaralanmasına veya hayatını kaybetmesine neden olan öngörülemeyen ve kasıtsız olaylardır. Kaza, bir yaralanma, maddi hasar veya çevresel hasara neden olabilecek bir ön olaydır şeklinde de tanımlanır (Remeshevskaya ve Gurets, 2023). İş kazaları, çalışanın hayatını ve sağlığını doğrudan etkilediği gibi, işletmeler açısından da üretim kayıpları, tazminatlar ve hukuki sorunlar doğurarak ekonomik zararlara neden olmaktadır. Ayrıca işletmelerin itibarını olumsuz etkileyerek, çalışanların moral ve motivasyonunu düşüren olaylar olarak ta ele alınmalıdır (Yorulmaz ve Öztürk, 2024).

2.2 Gemi İnşa Sektöründe İş Kazaları

Tersaneleri, limanları ve ilgili denizcilik operasyonlarını kapsayan denizcilik endüstrisi, küresel ticarete ve ekonomiye önemli ölçüde katkıda bulunan hayati bir sektördür. Tersaneler olarak da bilinen gemi inşa ve onarım tersaneleri, gemilerin inşa, bakım ve onarımının gerçekleştiği bu endüstrinin ayrılmaz bileşenleridir. Tersanelerdeki işin karmaşık ve tehlikeli doğası göz önüne alındığında, kaza/olay raporlarının analizi, personelin güvenliğini ve refahını, gemilerin ve altyapının bütünlüğünü sağlamak için son derece önemlidir (Jeong, 2021). Bununla beraber iş kazalarının büyük çoğunluğu tehlikeli davranışlar ve tutumlardan kaynaklanmakta olup, bu durum insan faktörlerinin İSG açısından belirleyici bir unsur olduğunu göstermektedir (Guldenmund, 2000).

Tersaneler genellikle kaynak, kesme, ağır kaldırma ve yüksekte çalışma gibi yüksek riskli faaliyetlerle karakterize edilir ve bunların hepsi çalışanları ve varlıkları potansiyel tehlikelere yatkın hale getirir. Deniz ortamı bu riskleri daha da kötüleştirir ve değişken hava koşulları ile su yakınında veya su üzerinde çalışmanın doğasında var olan tehlikeler gibi ek zorluklar ortaya çıkarır. Sonuç olarak, tersaneler gıda ve tarım gibi diğer birçok endüstri arasında daha yüksek sıklıkta işyeri kazalarına ve olaylarına karşı hassas olarak kabul edilebilir (Celebi ve ark., 2010).

Gemi ve tekne inşa endüstrisi, çok çeşitli deniz araçlarının imalatıyla ilgilenir. Gemi inşa süreci, birleştirme, kurma, temizleme, boyama, donatma ve test etmeyi içerir. Gemi inşa işçileri çok çeşitli fiziksel işler yaparlar ve zamanlarının çoğunu ayakta durarak, yürüyerek, eğilerek ve vücutlarının bazı kısımlarını bükerek geçirirler. Bazen ağır ekipman veya malzeme taşımak zorunda kalırlar. Gemi inşa işleri, çelik ve diğer metalleri kesmeyi, şekillendirmeyi ve birleştirmeyi içerir (You ve ark., 2019). Ayrıca, gemi inşa işi, dar alanlar ve önemli yükseklikler gibi çeşitli son derece tehlikeli durumlarda yapılır. Bu nedenle, diz çökme veya baş üstü çalışma gibi birçok rahatsız edici çalışma pozisyonu vardır (Barlas, 2012). Ayrıca, çok fazla iş açık havada yapılır ve aşırı hava koşullarının etkileri tehlikeli koşullara neden olabilir veya bunları kötüleştirebilir. Ek olarak, işçiler için önemli riskler oluşturabilecek bazı kimyasallar, boyalar ve çözücüler kullanılmaktadır (Kim ve ark., 2017).

Tersane işi geleneksel olarak tehlikelidir ve inşaat ile genel endüstrinin iki katından fazla yaralanma-kaza oranına sahiptir. Gemi inşa faaliyetleri en tehlikeli işyeri operasyonlarından bazılarını içerir. Tehlikeler arasında toksik maddelere maruz kalma, tehlikeli atmosferler, elektrik çarpması, düşmeler, yangınlar ve patlamalar bulunur. Ayrıca, gemi inşa işi ergonomik olarak tehlikelidir ve genellikle ağır fiziksel çalışma, rahatsız edici duruşlar, sık sık eğilme ve dönme çalışması gerektirir. Çoğu işlem metalle çalışmayı içerir ve bu genellikle kabul edilebilir güvenli sınırların üzerinde gürültü seviyeleri üretir (Jeong, 2021). Bunun yanında yüksek irtifalardan düşme, elektrik çarpmasına maruz kalma, yangın ve/veya patlama, nesnelere çarpma veya çarpma, boğulma vakaları tersane alanında sık görülen kazalardır. (Barlas ve Izci, 2018)

Tersanelerdeki kaza/olay raporları, istenmeyen olayların ayrıntılarını özetleyen, nelerin yaşandığına, dâhil olan bireylere ve varlıklara, anında ortaya çıkan sonuçlara dair kapsamlı bir hesap sağlayan önemli belgeler olarak hizmet eder. Bu raporlar yalnızca bürokratik formaliteler değildir; şirketler için geriye dönük analiz, risk değerlendirmesi ve önleyici strateji politikalarının uygulanması için paha biçilmez araçlardır.

2.3. İş Kazalarına Etki Eden Faktörler

Kaza odaklı araştırmaların son gelişmelerine rağmen, kazalara ilişkin anlayışımızın henüz tam olarak tamamlanmadığı ve karmaşık sosyal ve teknik sistemlerde kazaların meydana gelmeye devam ettiği genel olarak kabul edilmektedir. Bu nedenle, bunlardan ders çıkarmak çok önemlidir. Bu amaçla, kazaların nedenini araştırmak için çeşitli modeller

geliştirilmiştir. Geçtiğimiz yüzyılda, insan hatalarına odaklanan farklı modeller geliştirilmişken, bugün sosyo-teknik sistemlerde meydana gelen kazaların birbirleriyle etkileşim halinde olan çok çeşitli insan ve sistemsel faktörlerden kaynaklandığı yaygın olarak kabul edilmektedir (Salmon ve ark., 2017)

Kaza ve olay istatistiklerini analiz etmek, güvenli ve etkili bir işyerini beslemek için çok önemlidir. Bu, altta yatan nedenleri saptamak, önleyici eylemler başlatmak, sağlam bir güvenlik ahlakını beslemek ve düzenlemelere uyumu sağlamak gibi birden fazla yönü kapsar. Proaktif bir duruş benimseyerek ve bilgili seçimler için verileri kullanarak, kuruluşlar geçmiş kazalardan dersler çıkarabilir, uygulamalarını sürekli olarak iyileştirebilir ve paydaşlara karşı sorumluluklarını sürdürebilir, böylece daha güvenli ve daha dayanıklı bir sektörü destekleyebilir. Bu raporların yapılandırılmış analizlerini yürüterek, tersaneler eğilimleri tanıyabilir, altta yatan nedenleri belirleyebilir ve kazaları ve olayları etkileyen bağlamsal unsurları kavrayabilir. Bu prosedür, hassas müdahaleler oluşturmak, güvenlik prosedürlerini iyileştirmek ve tersane ortamındaki genel güvenlik kültürünü yükseltmek için hayati öneme sahiptir (Fu ve ark., 2020).

Çok sayıda mesleki güvenlik düzenlemesi, işyeri kazalarının ve olaylarının kapsamlı bir şekilde araştırılmasını ve raporlanmasını zorunlu kılar. Bu gerekliliklere uyulmaması, tersaneler için önemli yasal sonuçlar ve mali cezalarla sonuçlanabilir. Düzenleyici uyumluluğa ek olarak, kaza/olay raporlarının analizi tersanelerin mali istikrarı üzerinde doğrudan etkilere sahiptir. Kazalar ve olaylar, tıbbi masraflar ve ekipman hasarı gibi önemli doğrudan maliyetlerin yanı sıra kayıp çalışma saatleri ve artan sigorta primleri gibi dolaylı maliyetlere de neden olabilir. Proaktif analiz ve müdahale, bu maliyetleri azaltarak tersane operasyonlarının ekonomik sürdürülebilirliğini artırabilir. Kaza/olay raporlarının analizi, tersanelerin itibarını korumada ve geliştirmede de kritik bir rol oynar. Çalışanlar, müşteriler ve düzenleyici kurumlar dâhil olmak üzere paydaşlar, güvenlik standartları ve uygulamaları konusunda giderek daha fazla bilinçlenmektedir. Sıkı analiz ve güvenlik uygulamalarının iyileştirilmesi yoluyla güvenliğe kararlı bir bağlılık gösteren tersanelerin rekabetçi denizcilik sektöründe güven ve itibar kazanma olasılığı daha yüksektir. Tersanelerdeki kaza/olay raporlarının sistematik analizi, güvenlik, düzenleyici uyumluluk, finansal istikrar ve itibar yönetimi üzerinde etkileri olan çok yönlü bir çabadır. Tersane operasyonlarında risk yönetiminin vazgeçilmez bir bileşenidir ve sektörün sürekli iyileştirme ve iş gücü ile varlıklarının korunmasına olan bağlılığını vurgular. Diğer sektörlerde işyeri güvenliği ve kaza analizi üzerine çok sayıda araştırma ve çalışmaya

rağmen, tersanelere odaklanmada belirgin bir eksiklik vardır. Herhangi bir çalışma varsa bile, bunlar genellikle belirli bir yer veya tersane ile sınırlıdır ve öncelikli olarak insan etkilerine odaklanır (DeCamp ve Herskovitz, 2015).

2.3.1. İnsan Faktörleri

2004 ve 2014 yılları arasında, Türkiye tersanelerinde kayıtlı ölümcül iş kazaları hakkında kapsamlı bir araştırma yürütüldü ve Türkiye Çalışma ve Sosyal Güvenlik Bakanlığı'nın vaka raporları analiz edildi. Bu araştırma, İstanbul tersanelerindeki işçiler ve ustabaşılar arasında bir işyeri anketi içeriyordu ve ölümlerin başlıca nedenlerini yüksekten düşme, elektrik çarpması, yangın veya patlama, nesnelere çarpma veya nesnelere çarpma ve boğulma olarak kategorize etti. Dikkat çekici bir şekilde, bu ölümcül kazaların %43,7'si gemilerin içinde meydana gelirken, %56,3'ü tersanelerde meydana geldi. Anket, kişisel veriler, işyeri ortamı, kişisel faktörler ve işe gidip gelme faktörleri olmak üzere dört kategoride 56 sorudan oluşuyordu (Barlas ve Izci, 2018).

Gemi inşası ve onarımı, dünyanın en tehlikeli endüstrileri arasında yer almakta olup, oldukça karmaşık görevlerin eşzamanlı olarak yürütülmesini gerektirmektedir. İşçilerin dikkatsizliği, yetersiz iş sağlığı eğitimi, kazaların maliyetinin farkında olunmaması, insan kaynaklı hatalı işlemler ve yetersiz çalışma ortamı, mesleki kazalar için temel risk faktörleri olarak kalmaktadır (Tsoukalas ve Fragiadakis, 2016).

İşçiler arasında düşük eğitim düzeyi önemli bir risk faktörüdür. Tersane işçilerinin ortalama eğitim düzeyi yaklaşık 8,1 yıldır ve bu da çoğunun yalnızca ortaokulu bitirdiğini göstermektedir. Bu eğitim eksikliği, etkili iletişimi ve güvenlik protokollerinin anlaşılmasını engelleyebilir. Yorgunluk ve stres gibi faktörler yaygındır. Birçok işçi, özellikle uzun çalışma saatlerinden sonra yorgun ve uykulu hissettiğini bildirmektedir ve bu da kaza olasılığını artırmaktadır (Barlas ve Izci, 2018).

Kazaların %98,8'inde doğrudan nedenler veya diğer bir deyişle insan faktörleri etkili olmuştur. Bu faktörler arasında durumsal farkındalık kaybı (%61,1), fiziksel stres (%20,7) ve zihinsel stres (%14,2) yer almaktadır. (Wiegmann ve Shappell, 2001)'in HFACS (Human Factor Analysis and Classification System) çerçevesine göre, doğrudan nedenler genellikle operasyonların keskin ucundaki "Güvensiz Eylemler" ile ilgilidir (Shameli, 2024).

2.3.2. Çevresel Faktörler

İşyeri Koşulları: Kötü hava koşulları, gemi üretiminde aşırı kapasite ve yoğun çalışma programları gibi faktörler kazalara önemli katkıda bulunur. Bu koşullar, yönetimin güvenlik protokollerini ihmal etmesine yol açabilir. İşe gidip gelmekle ilişkili koşullar ve stres, işçilerin tersaneye vardıklarında hazır olma ve uyanık olma durumlarını da etkileyebilir (Barlas ve Izci, 2018).

Yetersiz iş sahası çevresi, iş kazaları için temel risk faktörleri arasında kalmaktadır. Özellikle kötü ergonomik koşullar, düzensizlik ve yetersiz aydınlatma gibi unsurlar kazaların oluşmasına zemin hazırlamaktadır. İş sahasındaki kaygan veya engellerle dolu zeminler, dar ve ani dönüşlü ya da uygun olmayan malzemeden yapılmış merdivenler ile çalışma alanları ve koridorların düşmeye karşı yeterince korunaklı olmaması gibi durumlar önemli riskler yaratmaktadır. Ayrıca acil çıkış yollarının kapalı ya da uygunsuz şekilde işaretlenmiş olması ve yetersiz veya uygunsuz havalandırma gibi çevresel koşullar, ölümle sonuçlanabilecek seviyede yüksek risk taşımaktadır. Bunlara ek olarak, çalışma istasyonlarında görüşün yetersiz olması veya kullanılan cihazların göstergelerinin net olarak okunamaması da çevresel risk faktörlerini artırmaktadır (Tsoukalas ve Fragiadakis, 2016).

İşyeri faktörleri kazaların %45,6'sında rol oynamıştır. Bu kategoride en etkili faktörler yetkin personel eksikliği (%24,3) ve kötü çalışma ortamı (%12,4) olarak belirlenmiştir. James Reason'ın Swiss Cheese Model (SCM) Mark III'üne göre, işyeri faktörleri, çalışanların içinde bulunduğu koşulları ve ortamı şekillendiren unsurlardır (Shabani ve ark., 2024).

2.3.3 Teknik ve Operasyonel Faktörler

Gemi inşa sektöründeki iş kazalarına etki eden teknik faktörler arasında fiziksel çalışma ortamının yetersizliği ve uygunsuzluğu, yüksek iş yükü ve çalışma temposu, güvenlik uygulamalarının etkinliği, güvenlik donanımlarının durumu ile güvenlik eğitimlerinin yeterliliği ve düzenliliği bulunmaktadır. Operasyonel faktörler ise yönetimin güvenlik değerlerine verdiği önem, doğrudan amirlerin güvenlik yaklaşımı ve çalışanlarla iletişim şekilleri, çalışanların güvenlik kurallarına uyumu ve güvenlik faaliyetlerine katılımı, iş güvencesizliği algısı, organizasyonel sistemlerin etkinliği ile işyeri kültürü ve çalışanlar arası ilişkisel çatışmalar olarak sıralanmaktadır. Bu teknik ve operasyonel faktörlerin

etkili yönetilmesiyle sektördeki iş kazalarının azaltılabileceği vurgulanmaktadır (Kim ve ark., 2017).

Sistemik faktörler kazaların %53,3'ünde etkili olmuştur. Bu kategoride en etkili faktörler yetersiz iş yöntemleri ve prosedürleri (%30,2) ile yetersiz eğitim ve yetkinlik (%26,6) olarak tespit edilmiştir. Sistemik faktörler, organizasyonel veya sistem düzeyindeki sorunları temsil eder ve genellikle olumsuz olaylara zemin hazırlayan gizli koşullardır (Shameli, 2024).

2.4 Örnek Olay İncelemeleri

Örnek olay incelemeleri, belirli bir durum veya olayı ayrıntılı ve kapsamlı şekilde analiz etmek için kullanılan nitel araştırma yöntemidir. Bu yöntem sayesinde belirli olayların gerçek hayattaki dinamikleri daha iyi anlaşılır ve teorik bilgilerin pratik uygulamalarla ilişkilendirilmesi sağlanır. Ayrıca örnek olay incelemeleri, farklı koşullarda benzer durumlarla karşılaştırma yapma olanağı tanır ve karmaşık süreçlerin daha iyi kavranmasını sağlar (Matías ve ark., 2008).

İşyeri kazası verilerinin Avrupa İş Güvenliği (ESAW) metodolojisine göre tanımlandığı ve değişkenleri, kaza anında kazazede tarafından gerçekleştirilen ana iş türünü, kazadan hemen önce kazazede tarafından gerçekleştirilen aktiviteyi, kazaya yol açan normalikten sapan son olay, kazazedenin yaralanma şeklini, kazazedenin temas ettiği ve yaralandığı alet, nesne veya enstrüman, kazanın kazazede üzerindeki fiziksel sonuçlarını "Yaralanma Türü", yaralanan vücut bölümünü ve varsa kazazedenin kaza nedeniyle iş göremez durumda olduğu takvim günü sayısının yanı sıra, iş kazasının tarihi, mağdurun cinsiyeti, doğum tarihi ve mesleği hakkında bilgilerin kullanıldığı bir veri kümesi kullanılan çalışmalar yapılmıştır (Pietilä ve ark., 2018).

2.5 Bulanık Mantık Yaklaşımı

Bulanık mantık, geleneksel küme teorisinin genişletilmiş bir halidir ve esnek düşünmeye izin vererek öznellik, belirsizlik ve muğlaklığı dikkate alır. Dilsel değişkenlerin belirsizliğini ve kesin olmayan bilgiyi temsil etmeyi mümkün kılan matematiksel bir yöntemdir. Özellikle geleneksel modellerle ifade edilmesi zor olan parametreleri değerlendirmede oldukça kullanışlıdır. Bulanık mantık, dilsel ifadelerin muğlaklığını sayısallaştırır ve belirsiz koşullarda daha doğru karar vermeye olanak sağlar (Oubellouch ve Aziz, 2024)

Bulanık mantık, doğrusal olmayan giriş-çıkış ilişkilerini modellemesi ile bilinen ve endüstriyel uygulamalar ile mesleki sağlık değerlendirmelerinde yaygın olarak kullanılan bir yöntemdir. Bulanık mantık, belirsizlikler içeren insan duruşları gibi sistemlerde meydana gelen belirsizlikleri ve algısal farklılıkları gidermek için kullanılır. Bu yaklaşımda giriş değişkenleri üyelik fonksiyonları ile bulanıklaştırılır ve çıktılar değerlendirilmesinde 'eğer-ise' kuralları kullanılır. Sonrasında bulanık çıktılar, net sayılara dönüştürülerek değerlendirilir. Böylece değerlendirmelerde daha hassas sonuçlar elde edilir ve gözlemsel hataların azaltılması sağlanır (Zhao ve ark., 2025)

Bilgiye dayalı karar sistemlerinin (Bilgi Tabanlı İşleme) bir parçası olan bu teknik, kazaların meydana gelme sıklığının hem de kaza sonuçlarının değerlendirilmesinde risk analizi alanında yaygın olarak uygulanmaktadır. Özellikle, yöntem risk değerlendirmesinde çok farklı tipte parametrelerin (olayların sıklığı, stres faktörü, maruz kalma süresi, vb. gibi) hesaba katılmasına izin vererek, bunların kazayı başlatmadaki etkileşimlerini vurgulamaktadır. Bulanık Mantık yönteminin etkili bir şekilde uygulanmasının sınırlayıcı faktörü, referans parametrelerinin doğru bir şekilde ayarlanabilmesi için yeterli ve homojen sayıda kaza kaydının olmasıdır (Murè ve Demichela, 2009).

2.6 Çok Kriterli Karar verme Metotları

ÇKKV metotları, farklı kriterleri ve parametreleri göz önünde bulundurarak en uygun seçimi yapmak için kullanılır. Bu yöntemler genellikle belirsizlikler ve öznel değerlendirmelerle karşılaşılan durumlarda kullanılır ve değerlendirmeleri daha tutarlı hale getirir. Karar verme süreçlerinde, bulanık mantık ve risk değerlendirme yöntemlerin entegre kullanımı, belirsiz ve karmaşık durumlarda risk değerlendirme süreçlerini iyileştirmek için daha etkili sonuçlar üretir. Bu yöntemler, karar vericilere daha net ve eyleme dönüştürülebilir risk değerlendirmeleri sunarak, risk temelli tasarım süreçlerinde daha net kararlar alınmasına imkân sağlar (Oubellouch ve Aziz, 2024).

ÇKKV, modern karar bilimi ve yöneylem araştırmasının önemli bir dalıdır. ÇKKV, aralarında çatışan çok sayıda kriteri dikkate alarak, mevcut alternatifler içinden en uygun olanını seçmeyi amaçlar. ÇKKV yöntemlerinin en önemli özelliklerinden biri, optimal çözümü aramak yerine karar vericilerin tercihlerini yansıtmasıdır. ÇKKV yöntemleri toplum, ekonomi, mühendislik ve yönetim alanlarındaki birçok problemin çözümünde geniş bir kullanım alanına sahip olup, temel varsayımlar, teori, hesaplama süreçleri ve

karmaşıklık açısından birbirlerinden farklılık gösterirler. Bu yöntemler arasında yakınlık skoruna göre sıralama yapan yöntemler İdeal Çözüm Benzerliğe Göre Tercih Sıralama Tekniği TOPSIS (Technique for Order Preference by Similarity to Ideal Solution), Çelişen kriterler altında uzlaşık çözümle optimizasyon yöntemi VIKOR (VlseKriterijumska Optimizacija I Kompromisno Resenje), Oran Analizi ile Çok Amaçlı Optimizasyon – Tip I MOORA I (Multi-Objective Optimization by Ratio Analysis – Type I) yöntemleri, Orijinal skora göre sıralama yapan yöntemler Referans Noktalı Oran Analizi ile Çok Amaçlı Optimizasyon – Tip II MOORA II (Multi-Objective Optimization by Ratio Analysis with Reference Point Approach – Type II) ve Karmaşık Orantılı Değerlendirme Yöntemi COPRAS (Complex Proportional Assessment) yöntemleri, pozitif skora göre sıralama yapan yöntemler Basit Ağırlıklı Toplama Yöntemi SAW (Simple Additive Weighting) ve Toplam Oran Değerlendirmesi ARAS (Additive Ratio Assessment) yöntemleri gibi yöntemler yer almaktadır. Yakınlık skoruna göre sıralama yapan yöntemler TOPSIS, ideal çözüme yakınlığı temel alarak alternatifleri sıralamakta kullanılır (Li ve ark., 2024). Bu çalışmada TOPSIS yöntemi kullanılmıştır.

2.6.1 TOPSIS Yöntemi

TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution- İdeal Çözüme Benzerliğe Göre Tercih Sırasının Tekniği), ÇKKV tekniği olup birden fazla kritere dayalı olarak bir dizi alternatif arasından en iyi alternatifi seçmek için kullanılır. Alternatifleri ideal çözüme olan geometrik uzaklıklarına göre sıralamakta yaygın olarak kullanılır. Yöntem, ideal çözüm kavramına dayanır; ideal çözüm, tüm kriterlerde en iyi performansı gösteren çözümdür. Yöntem, her alternatifin ideal çözüme olan mesafesini hesaplayarak, ideal çözüme en yakın alternatifi seçer. Bu yöntem, Hwang ve Yoon (1981) tarafından ortaya atılmıştır. "TOPSIS metodunun temel prensibi, en iyi alternatifin pozitif ideal çözüme (her kriter için mümkün olan en iyi sonuç) en kısa Öklid mesafesine ve negatif ideal çözüme (en kötü sonuç) ise en uzun Öklid mesafesine sahip olmasıdır. TOPSIS, her kriterin olumlu ya da olumsuz etkisi olduğunu varsayar ve tüm kriterlerin ortak bir ölçekte normalleştirilebileceğini kabul eder. Bu yöntem, mühendislik, yönetim, ekonomi ve çevre bilimleri gibi çeşitli alanlarda karmaşık karar verme durumlarında başarıyla uygulanmaktadır." (Ahmed ve ark., 2024)

2.7. Fuzzy-TOPSIS Entegrasyonu

"Fuzzy TOPSIS, bulanık mantık teorisini klasik TOPSIS yöntemiyle birleştiren çok kriterli bir karar verme tekniğidir. Klasik TOPSIS yöntemi alternatifleri ideal ve negatif ideal çözüme olan uzaklıklarına göre sıralarken, bulanık TOPSIS yöntemi ise karar vericilerin kriterlere ilişkin belirsizliğini ve muğlaklığını dikkate alır ve dilsel terimlerle ifade edilen kararları sayısal forma dönüştürür. Bu yöntem, farklı kriterler altında çok sayıda alternatif arasından ideal çözüme en yakın olan alternatifi belirlemek amacıyla kullanılır. Özellikle karmaşık ve belirsiz karar verme ortamlarında, yatırım kararları gibi finansal problemlerde karar vericilerin belirsizlik ve sübjektif değerlendirmelerini daha etkin bir şekilde yansıtmak için tercih edilir. Fuzzy TOPSIS, klasik TOPSIS yönteminin belirsiz ve karmaşık ortamlarda karar verme etkinliğini artırmak için bulanık mantık teorisi ile entegre edilmesi sonucu geliştirilmiştir. Bu yöntem, kriter ve alternatiflere ilişkin belirsizlikleri daha etkin şekilde işleyebilmek adına dilsel ifadeleri sayısal değerler haline dönüştüren bulanık kümeler teorisini benimsemiştir. Bu sayede karar vericilerin sübjektif ve belirsiz değerlendirmeleri sistematik olarak analiz edilip objektif hale getirilebilir." (Sachdeva ve ark., 2023)

2.7.1 Küresel Fuzzy TOPSIS Metodu

Küresel bulanık TOPSIS (SF-TOPSIS) yöntemi, sübjektif yargılardan kaynaklanan belirsizliği ve objektif verilerdeki bulanıklığı eş zamanlı olarak ele alabilmesi nedeniyle farklı belirsizlik ve bulanıklık türlerini içeren ÇKKV problemlerinde geniş bir kullanım alanına sahiptir (Sharaf, 2023)

SF-TOPSIS yöntemi, klasik TOPSIS yönteminin küresel bulanık kümelerle (Spherical Fuzzy Sets) genişletilmiş hâlidir. Küresel bulanık kümeler, üyelik, üyelik dışı ve tereddüt derecelerini üç boyutlu bir küresel yüzey üzerinde bağımsız olarak tanımlayabilen bir bulanık küme yaklaşımıdır. Küresel bulanık kümeler, üyelik (μ), üyelik dışı (ν) ve tereddüt (π) derecelerini üç boyutlu bir küresel yüzey üzerinde bağımsız olarak tanımlayabilen bir bulanık küme yaklaşımıdır ve Formül 1’de yer alan koşulu sağlar.

$$0 \leq \mu_{\{A\}(u)}^2 + \nu_{\{A\}(u)}^2 + \pi_{\{A\}(u)}^2 \leq 1 \quad (1)$$

Bu yöntem sayesinde karar vericiler, değerlendirmelerini yaparken üyelik, üyelik dışı ve tereddüt düzeylerini daha bağımsız ve esnek biçimde belirleyebilirler. Küresel Fuzzy TOPSIS yönteminin temel amacı, farklı kriterlere göre değerlendirilen alternatiflerin

ideal çözüme olan mesafelerini küresel bulanık ortamda hesaplayarak, en uygun ve dengeli alternatifi belirlemektir. Bu yöntem, karar vericilerin farklı kriterlere dair belirsizliklerini ve tereddütlerini üç boyutlu küresel ortamda ifade ederek daha hassas ve gerçekçi değerlendirmeler yapmalarını sağlar (Kutlu Gündoğdu ve Kahraman, 2019). Küresel Fuzzy TOPSIS yöntemi aşağıdaki dokuz aşamadan oluşmaktadır.

SF TOPSIS yönteminde karar vericiler tarafından yapılan değerlendirmelerin temsil edilmesi amacıyla küresel bulanık karar matrisi (**D**) oluşturulur. Bu matris, her bir alternatifin (Kaza Sebebi) her bir kritere (kazazedenin görevi, Olay Ayıvb.) göre değerlendirmesini ifade eder. Karar vericiler, alternatifleri ve kriterleri değerlendirmek için belirlenen dilsel terimlerle değerlendirme matrisini Formül 2 gösterilen şekilde oluşturulur.

$$D = \begin{bmatrix} (\mu_{11}, \nu_{11}, \pi_{11}) & (\mu_{12}, \nu_{12}, \pi_{12}) & \cdots & (\mu_{1n}, \nu_{1n}, \pi_{1n}) \\ (\mu_{21}, \nu_{21}, \pi_{21}) & (\mu_{22}, \nu_{22}, \pi_{22}) & \cdots & (\mu_{2n}, \nu_{2n}, \pi_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ (\mu_{m1}, \nu_{m1}, \pi_{m1}) & (\mu_{m2}, \nu_{m2}, \pi_{m2}) & \cdots & (\mu_{mn}, \nu_{mn}, \pi_{mn}) \end{bmatrix} \quad (2)$$

Spherical Fuzzy Set teorisinde iki küresel bulanık sayının çarpımı, bu sayıların bileşenleri olan üyelik (μ), karşıtlık (ν) ve kararsızlık (π) derecelerine dayalı olarak gerçekleştirilir. Formül 3'te yer alan bu işlem, alternatiflerin ve kriterlerin küresel bulanık temsilleriyle yapılan matematiksel işlemlerde kritik rol oynar.

$$\tilde{A}_s \otimes \tilde{B}_s = \left\{ \mu_{A_s} \mu_{B_s}, \sqrt{v_{A_s}^2 + v_{B_s}^2 - v_{A_s}^2 v_{B_s}^2}, \sqrt{(1 - v_{B_s}^2) \pi_{A_s}^2 + (1 - v_{A_s}^2) \pi_{B_s}^2 - \pi_{A_s}^2 \pi_{B_s}^2} \right\} \quad (3)$$

İki küresel bulanık sayının çarpımının yeni bir küresel bulanık sayı üretmesini sağlar. Bu işlem özellikle normalize edilmiş karar matrislerinin ağırlıklandırılması, ideal çözüm noktalarının hesaplanması ve ÇKKV adımlarında kullanılmaktadır. Formülün bu yapısı hem belirsizlik hem de karşıtlık düzeylerini dikkate alarak, daha gerçekçi ve esnek modellemelere olanak tanır

SWGM (Spherical Weighted Geometric Mean) operatörü ise küresel ağırlıklı geometrik ortalama anlamına gelir. Bu operatörde, her bir değerlendirme kriterinin karar vericiler tarafından yapılan puanlamaları, yine üyelik, karşıtlık ve kararsızlık bileşenleri özelinde ayrı ayrı hesaplanır. Bu değerlendirmeler, Formül 4'te yer alan Küresel Ağırlıklı

Aritmetik Ortalama (SWAM) veya Formül 5'te yer alan Küresel Ağırlıklı Geometrik Ortalama (SWGM) gibi küresel bulanık toplama operatörleri ile birleştirilir.

$$SWAM_{\omega} = \left\{ \sqrt{1 - \prod_{i=1}^n (1 - \mu_i^2)^{\omega_i}}, \prod_{i=1}^n v_i^{\omega_i}, \sqrt{1 - \prod_{i=1}^n (1 - \mu_i^2)^{\omega_i} - \prod_{i=1}^n (1 - \mu_i^2 - \pi_i^2)^{\omega_i}} \right\} \quad (4)$$

$$SWGM_{\omega} = \left\{ \prod_{i=1}^n \mu_i^{\omega_i}, \sqrt{1 - \prod_{i=1}^n (1 - v_i^2)^{\omega_i}}, \sqrt{\prod_{i=1}^n (1 - v_i^2)^{\omega_i} - \prod_{i=1}^n (1 - v_i^2 - \pi_i^2)^{\omega_i}} \right\} \quad (5)$$

SF-TOPSIS yönteminde, alternatiflerin birbirine kıyasla ne kadar tercih edilir olduğunu değerlendirmek amacıyla her alternatif için bir skor fonksiyonu hesaplanır. Bu fonksiyon, alternatifin üyelik (μ), karşıtlık (v) ve kararsızlık (π) derecelerini dikkate alarak alternatifin genel tercih düzeyini yansıtır. Formül 6, bir alternatifin net pozitifliğini, yani olumlu yönlerinin olumsuz yönlerine göre ne kadar ağır bastığını gösterir. Sonuç pozitifse alternatifin daha çok tercih edilmesi gerektiği; negatifse daha az tercih edilebilir olduğu yorumlanır. Bu hesaplama, alternatiflerin sıralanmasında kritik bir rol oynar ve karar vericiye doğrudan rehberlik eder.

$$Score(\tilde{A}_s) = (\mu_{A_s} - \pi_{A_s})^2 - (v_{A_s} - \pi_{A_s})^2 \quad (6)$$

Küresel bulanık sayılarla yapılan değerlendirmelerde, bir alternatifin küresel bulanık temsilinin geçerliliğini ve güvenilirliğini ölçmek amacıyla Formül 7'de yer alan doğruluk fonksiyonu kullanılır. Bu fonksiyon, küresel bulanık sayının üç temel bileşeni olan üyelik (μ), karşıtlık (v) ve kararsızlık (π) derecelerini kareleriyle birlikte toplar.

$$Accuracy(\tilde{A}_s) = \mu_{A_s}^2 + v_{A_s}^2 + \pi_{A_s}^2 \quad (7)$$

Bu formül, bir küresel bulanık sayının geometrik olarak birim küre üzerindeki konumunu ifade eder. Küresel bulanık setlerin tanımı gereği, bu üç bileşenin karelerinin toplamı 1'e eşit ya da daha küçük olmalıdır ($0 \leq \mu_{\{A\}(u)}^2 + v_{\{A\}(u)}^2 + \pi_{\{A\}(u)}^2 \leq 1$). Bu toplam ne kadar 1'e yaklaşırsa, söz konusu alternatifin değerlendirmesi o kadar net ve güvenilir olarak kabul edilir.

SF-TOPSIS yönteminde, her bir alternatifin ideal çözüme ne kadar yakın olduğunu ölçmek amacıyla bir uzaklık fonksiyonu tanımlanır. Bu uzaklık fonksiyonu, alternatifin (X_i) küresel bulanık karar değerleri ile pozitif ideal çözüm (X^*) arasındaki farkı hesaplayarak, alternatifin toplam idealden ne kadar uzak olduğunu sayısal olarak Formül 8 ile hesaplayarak ifade eder.

$$D(\tilde{X}_i, \tilde{X}^*) = \sqrt{\frac{1}{2n} \sum_{j=1}^n [(\mu_{ij} - \mu_j^*)^2 + (v_{ij} - v_j^*)^2 + (\pi_{ij} - \pi_j^*)^2]} \quad (8)$$

Formül 8, klasik Öklidyen mesafe yaklaşımına benzer şekilde çalışır; ancak üçlü yapıya sahip küresel bulanık sayılar için özelleştirilmiştir. Her kriter için üç bileşen (üyelik, karşıtlık, kararsızlık) ayrı ayrı ele alınarak farkların kareleri hesaplanır ve tüm kriterler üzerinden toplanır. Elde edilen sonuç, alternatifin ideal çözüme olan toplam uzaklığını gösterir. Uzaklık değeri ne kadar küçükse, ilgili alternatif pozitif ideale o kadar yakındır ve bu durum alternatifin tercih edilirliliğinin yüksek olduğunu gösterir. Bu hesaplama süreci, negatif ideal çözüme olan uzaklıkla birlikte değerlendirilerek göreceli yakınlık katsayısı (closeness coefficient) hesaplanır. Böylece TOPSIS yönteminin temel sıralama mekanizması tamamlanır

Ardından ağırlıklandırılmış küresel bulanık karar matrisi oluşturularak, her bir alternatifin Küresel Bulanık Pozitif İdeal Çözüme (SF-PIS) ve Küresel Bulanık Negatif İdeal Çözüme (SF-NIS) mesafeleri Formül 9 ile hesaplanır. Son olarak, alternatiflerin ideal çözümlere olan yakınlıklarına göre sıralama yapılarak en uygun alternatif seçilir." (Kutlu Gündoğdu ve Kahraman, 2019).

$$\xi(X_i) = \frac{D(X_i, X^*)}{D_i^* \min \frac{D(X_i, X^-)}{D_i^- \max}} \quad (9)$$

2.8 Compromise (Uzlaştırma)

Uzlaştırma (compromise), farklı kriterlerin bir arada değerlendirilmesi sonucu, ideal çözümün mümkün olmadığı durumlarda, mümkün olan en iyi dengeyi bulmayı hedefleyen karar verme yaklaşımıdır. Bu yöntem, farklı kriterler arasında oluşan çatışmayı azaltmak ve mümkün olan en uygun alternatifi seçmek amacıyla kullanılır (Lu ve Liu, 2024).

Uzlaştırma yöntemlerinin temel amacı, farklı kriterlerin karşılıklı ilişkilerini ve çelişkilerini göz önünde bulundurarak tek bir hedef yerine tüm kriterleri belirli ölçüde tatmin edecek dengeli çözümler üretmektir. Bu yöntemler özellikle kriterlerin çatıştığı, karmaşık ve çok boyutlu karar problemlerinde etkili olup, gerçek dünyada karşılaşılan karar verme durumlarında güvenilir ve uygulanabilir çözümler sağlamak için kullanılmaktadır (Rashid ve ark., 2025).

Bu çalışmada kullanılan Uzlaştırma yöntemlerinden biri olan ağırlıklı toplama yöntemi, farklı kaynaklardan veya yöntemlerden elde edilen skorların her birine belirli ağırlıklar verilerek bu skorların toplanmasıyla gerçekleştirilir. Her bir kaynağın veya yöntemin sonucu, önceden belirlenmiş sabit bir ağırlıkla çarpılarak nihai uzlaştırma skorlarına dönüştürülür. Bu yöntem, farklı kaynaklardan gelen bilgilerin ağırlıklı olarak birleştirilmesi ve nihai sonuçların daha dengeli bir biçimde elde edilmesini sağlar (Yamada, 2008).

2.9 Entropi Yöntemi

Entropi ağırlık yöntemi, kriterlerin önem derecelerini objektif olarak belirlemek amacıyla bilginin farklılığını ve iletim kapasitesini analiz eder. İlk adımda değerlendirme göstergeleri ile alternatiflerden oluşan Formül 10 kullanılarak karar matrisi kurulur ve farklı ölçü birimlerinin etkisini gidermek için Formül 11 kullanılarak normalize edilir.

$$X = [x_{ij}]_{m \times n}, \quad i = 1, \dots, m; j = 1, \dots, n \quad (10)$$

$$\sum_{i=1}^m p_{ij} = 1, p_{ij} \geq 0 \quad (11)$$

Normalleştirilmiş değerler kullanılarak her alternatifin göstergeler içindeki oranları hesaplanır; bu oranlardan Formül 12 kullanılarak ağırlık matrisi oluşturulur. Daha sonra her gösterge için Formül 13 ile entropi değeri hesaplanır entropi ne kadar küçükse, göstergenin bilgi içeriği ve ayırt ediciliği o kadar yüksektir. Bu değerlerden bilgi payı elde edilir ve nihai kriter ağırlıkları belirlenir. Böylece subjektif yargılar en aza indirilir, değerlendirme sonuçlarının tutarlılığı ve güvenilirliği artırılır (Chen, 2020; Hezam ve ark., 2023).

$$\sum_{j=1}^n w_j = 1 \quad (12)$$

Entropi yöntemi (ağırlıklandırma yöntemi) içinde kullanılan bir katsayıyı ifade eder. m alternatif sayısını (her karar seçeneğinin sayısını) ifade eder.

$$k = \frac{1}{\ln m} \quad (13)$$

2.10. Fuzzy–Makine Öğrenimi Entegrasyonu

Makine öğrenmesi, veri analizi ve tahminleme konularında güçlü çıktılar üretse de belirsizlik barındıran karmaşık problemlerde bazı sınırlamalar sergilemektedir. Bu nedenle, çoklu kriterlerin dikkate alındığı ve belirsizlikle başa çıkabilen yöntemlerle entegre edilmesi gereklilik hâline gelmiştir.

Makine öğrenmesi ile ÇKKV yöntemleri genel olarak iki temel yaklaşımla birleştirilmektedir. İlk yaklaşımda, makine öğrenmesi algoritmaları farklı performans ölçütlerine göre değerlendirilmekte ve en uygun model, ÇKKV teknikleri yardımıyla seçilmektedir. Bu yapı, makine öğrenmesi ve ÇKKV yöntemlerinin birbirini tamamladığı sıralı bir süreç oluşturur. İkinci yaklaşım ise, bulanık mantığın doğrudan makine öğrenmesi algoritmasının yapısına entegre edilmesidir. Böylece belirsizlikler ve dilsel değişkenler modelleme sürecine doğrudan dâhil edilerek Fuzzy Machine Learning (FML) yaklaşımı ortaya çıkmaktadır. FML, özellikle yüksek belirsizlik içeren durumlarda daha esnek, anlaşılır ve insan benzeri karar verme yapıları sunmaktadır.

Makine öğrenmesi algoritmalarının farklı veri kümeleri ve değerlendirme ölçütleri altında değişken performanslar göstermesi, model seçimini karmaşıktırmakta ve yalnızca doğruluk gibi tek bir ölçüte dayanarak karar vermeyi yetersiz kılmaktadır. Bu noktada Fuzzy-TOPSIS gibi bulanık temelli ÇKKV yöntemleri, belirsizlik altında karar vericilerin öznel yargılarını sayısal hale getirerek önemli avantajlar sağlamaktadır. Makine öğrenmesi ile ÇKKV yöntemlerinin birlikte kullanımı, özellikle yalnızca tek bir başarı ölçütüyle karar vermenin yetersiz kaldığı çok boyutlu problemlerde önem kazanmaktadır. Bu entegrasyon sayesinde hem algoritma seçimi hem de model değerlendirme süreçlerinde daha tutarlı ve kapsamlı sonuçlar elde edilmektedir (Akinsola ve ark., 2019).

Makine öğrenmesi algoritmalarının Fuzzy-TOPSIS, ARAS, MOORA ve EDAS gibi farklı ÇKKV yöntemleri ile değerlendirilmesi, sınıflama ve örüntü tanıma gibi güçlü yönlerinin çok boyutlu karar yapıları ile birleştirilmesine olanak sağlamaktadır. Bu sayede esnek, açıklanabilir ve daha etkili karar süreçleri ortaya çıkmaktadır (Yılmaz, 2020).

Makine öğrenmesi, teorik gelişmeleri ve farklı disiplinlerden beslenen yönüyle önemli bir gelişme göstermiştir. Ancak, belirsiz gözlemler, eksik veri ve karmaşık ilişkilerin yer

aldığı ortamlarda klasik yöntemler yetersiz kalabilmektedir. Bu nedenle araştırmacılar, makine öğrenmesi algoritmalarını bulanık sistemler (örneğin bulanık kümeler, ilişkiler, mantık ve ölçüler) ile bütünleştirme yoluna gitmiştir. Bu birleşim “bulanık makine öğrenmesi” (FML) olarak adlandırılmakta olup, gerçek yaşamda karşılaşılan farklı belirsizlik türlerine daha dirençli modeller sunmaktadır (Lu ve ark., 2024).

FML yaklaşımı, özellikle değişken ve öngörülemeyen ortamlarda geleneksel makine öğrenmesi yöntemlerine göre üstün performans sergilemektedir. Bulanık küme teorisine dayalı bu modeller, farklı senaryolardaki belirsizlikleri etkili biçimde temsil ederek değişen örüntülere uyum sağlayabilir. Ayrıca, karar süreçlerinin şeffaf ve anlaşılabilir olmasını mümkün kılan yorumlanabilirlik avantajı da sunmaktadır. Bulanık kümeler, eksik ya da dilsel nitelikli verileri temsil etmekte kullanılırken, robotik ve otonom sistemler gibi gerçek zamanlı belirsizliklerin söz konusu olduğu uygulamalarda karar süreçlerini daha etkili hâle getirmektedir (Wu, 1999). Takagi ve Sugeno (1985) tarafından geliştirilen bulanık kural tabanlı sistemler, özellikle tıbbi teşhis gibi alanlarda yorumlanabilir ve güvenilir tahmin altyapısı sağlamaktadır. Benzer şekilde, bulanık kümeleme yöntemleri görüntü işleme gibi alanlarda klasik yöntemlerle yakalanamayan veri desenlerini ortaya çıkarmakta etkilidir. Bulanık ilişkiler ise çok kaynaklı ya da çok boyutlu veriler arasındaki doğrusal olmayan ve esnek ilişkileri modelleyerek, makine öğrenmesi algoritmalarının kapsamını genişletmektedir (Lu ve ark., 2024).

SF-TOPSIS yöntemi, belirsizlik içeren ortamlarda karar vericilerin değerlendirmelerini daha hassas şekilde işleyebilmek için geliştirilmiş ÇKKV yöntemlerinden biridir (Kumar et al., 2021). Bu yöntemin kullanımı, özellikle İSG gibi karmaşık değişkenlerin yer aldığı alanlarda, uzman görüşlerinin sezgisel olarak modellenmesine olanak tanır. Öte yandan, Random Forest ve Extra Trees gibi topluluk (ensemble) makine öğrenmesi algoritmaları, yüksek doğruluk ve genelleme kapasitesiyle öne çıkmaktadır (Breiman, 2001; Geurts et al., 2006).

Bu çalışmada SF-TOPSIS yöntemi ile uzman görüşlerine dayalı kriter ağırlıkları hesaplanmış ve bu ağırlıklarla modellenen veri seti, Random Forest ve Extra Trees algoritmalarıyla analiz edilmiştir. Uzman temelli sezgisel karar süreci ile veri temelli öğrenme yöntemleri birleştirilerek hem açıklayıcı hem de tahminleyici güçlü bir model elde edilmiştir. Bu tür hibrit yaklaşımlar, literatürde öngörü performansını artırmak ve karar desteğini güçlendirmek amacıyla sıklıkla tercih edilmektedir (Vazirani et al., 2020; Manjunatha, 2023).

BÖLÜM 3. ARAŞTIRMA METODOLOJİSİ

Bu çalışmanın akışı, Grafik 1’de sunulan süreç akış diyagramında sistematik bir şekilde özetlenmiştir. Süreç, gemi inşa sektörüne ait iş kazası verilerinin toplanmasıyla başlamış ve bu verilerin analiz için uygun hale getirilmesi amacıyla ön işleme adımlarının gerçekleştirilmesiyle devam etmiştir. Ardından, çeşitli makine öğrenmesi algoritmaları kullanılarak veri üzerinden modelleme yapılmış ve en yüksek sınıflandırma başarımını gösteren modeller üzerinden değişken önem düzeyleri belirlenmiştir. Elde edilen önemli değişkenler, uzman görüşleriyle birlikte değerlendirilmek üzere SF-Fuzzy TOPSIS ÇKKV yöntemine girdi olarak sunulmuş ve aynı zamanda yöntemsel bütünlüğü sağlamak amacıyla kriter ağırlıkları entropi yöntemiyle hesaplanmıştır. Farklı analiz yöntemlerinden elde edilen skorlar, uzlaştırma (compromise) yaklaşımı ile bütünleştirilerek, iş kazalarına etki eden faktörlerin öncelik düzeyleri ortaya konmuştur. Sürecin sonunda hem veri temelli hem uzman görüşü tabanlı yöntemlerin çıktılarının birleştirilmesiyle nihai değerlendirme gerçekleştirilmiştir.

Tez kapsamında gemi inşa sektöründe meydana gelen iş kazalarının nedenlerini ve bu kazalara etki eden faktörleri belirlemek amacıyla hem makine öğrenmesi tabanlı tahminleme hem de ÇKKV teknikleri entegre bir yaklaşımla kullanılmıştır. Öncelikle, Türkiye’deki tersanelerde 10 yıllık süreçte meydana gelen 2020 iş kazasına ait veri kümesi üzerinden çeşitli veri ön işleme ve sınıflandırma işlemleri gerçekleştirilmiş, ardından dokuz farklı makine öğrenmesi algoritması uygulanmıştır. Bu modellerle kaza sebebini etkileyen faktörler araştırılmıştır. Başarı oranı en yüksek iki makine öğrenmesi modelinin sonuçlarına göre kaza sebebini etkileyen faktörlerden en önemlileri tespit edilmeye çalışılmıştır. Böylece tüm faktörler arasından önem dereceleri yüksek olan dört faktör dikkate alınmıştır.



Grafik 1. Süreç Akış Diyagramı

3.1 Veri Toplama Yöntemleri

Bu çalışmanın temelini oluşturan veri kümesi, gemi inşa sanayisinde faaliyet gösteren birden fazla tersanede meydana gelen iş kazalarına ilişkin ayrıntılı bilgilerden oluşmaktadır. Özellikle İstanbul ili Tuzla Tersaneler Bölgesi'nde bulunan ve çeşitli boyutlarda faaliyet gösteren tersanelerde gerçekleşen olaylar incelenmiş, toplamda 2020 adet iş kazasına ait veri analiz kapsamına alınmıştır. Bu kazalar yaklaşık on yıllık bir zaman dilimine yayılmakta olup, 2008-2018 yılları arasını kapsamaktadır. Araştırmada hem birincil hem de ikincil veri toplama yöntemleri kullanılmıştır. Birincil veriler, tersanelerdeki İSG birimlerinden temin edilen resmi kaza formları, raporlar ve olay bildirim dokümanlarından derlenmiştir. Bu belgeler, kazaların oluş şekli, kazazedenin görevi, yaranın vücuttaki yeri, kaza türü ve kaza sonucu gibi değişkenleri kapsamaktadır. Bu bilgiler, doğrudan olay gerçekleştirildikten sonra doldurulan formlar ve sahadaki gözlemlere dayalı olarak toplanmıştır. İkincil veriler ise İSG literatürü, sektörel istatistikler, daha önce yayınlanmış bilimsel çalışmalar ve kamu kurumlarının (TÜİK, SGK vb.) yayımladığı resmi raporlardan sağlanmıştır. Bu çift yönlü veri toplama stratejisi hem sahaya özgü durumu anlamaya hem de çalışmanın bilimsel geçerliliğini artırmaya hizmet etmiştir. Veri toplama sürecinde etik ilkelere dikkat edilmiş, kişisel veriler gizli tutulmuş ve anonimleştirilmiştir. Toplanan bu veriler hem makine öğrenmesi tabanlı modelleme hem de ÇKKV yöntemleri için temel girdi olarak kullanılmıştır. Sonuç olarak, elde edilen veri kümesi sektör gerçekliğini yansıtan, çok boyutlu ve analize uygun bir yapıya sahiptir.

3.1.1 Birincil Veri Kaynakları

Birincil veriler, doğrudan saha uygulamalarından ve olay sonrası oluşturulan resmî belgelerden elde edilmiştir. Bu kapsamda, Tuzla Tersaneler Bölgesi'nde faaliyet gösteren tersanelerin İSG birimlerinden temin edilen iş kazası kayıtları, kaza bildirim formları, İSG raporları ve olay değerlendirme tutanakları temel veri kaynaklarını oluşturmuştur. Her bir kaza olayı, olayın meydana geliş şekli, kazazedenin görev tanımı, kazanın gerçekleştiği yer, kaza türü, yaranın vücuttaki yeri, olayın gerçekleşme zamanı, kazalı sayısı, olay günü ve ayı gibi birçok boyutta detaylandırılmıştır. Bu bilgiler, olayın hemen ardından yetkili personel tarafından sahada yapılan gözlemler doğrultusunda kayıt altına alınmış ve dijital ortama aktarılmıştır. Verilerin güvenilirliğini sağlamak amacıyla çapraz kontrol yöntemleri uygulanmış, aynı kazaya ait çoklu kayıtların tutarlılığı incelenmiş ve eksik ya da çelişkili veriler İSG uzmanlarının görüşleri doğrultusunda netleştirilmiştir.

Ayrıca veriler, bireylerin kimlik bilgileri ile ilişkilendirilmeden, anonimleştirilmiş biçimde analiz sürecine dâhil edilmiştir. Bu süreç, çalışmanın saha gerçekliğine uygunluğunu artırmakla kalmamış, aynı zamanda makine öğrenmesi modelleri ve ÇKKV yaklaşımları için yüksek kalitede, güvenilir ve çok boyutlu bir veri kümesinin oluşturulmasına zemin hazırlamıştır.

3.1.2 İkincil Veri Kaynakları

Bu çalışmada kullanılan ikincil veri kaynakları, birincil kaynaklardan elde edilen saha verilerini desteklemek ve analizlere teorik bir zemin kazandırmak amacıyla kapsamlı biçimde değerlendirilmiştir. İkincil veriler, başta TÜİK, SGK ve Çalışma ve Sosyal Güvenlik Bakanlığı olmak üzere kamu kurum ve kuruluşlarının yayımladığı istatistiksel raporlar, yıllık faaliyet dökümleri ve iş kazalarına dair veri tabanlarından elde edilmiştir. Bu kurumların sağladığı istatistiksel bilgiler, Türkiye genelindeki iş kazalarının sektörel dağılımı, yıllara göre kaza sayıları, ölümcül ve yaralanmalı kaza oranları gibi kritik göstergeleri içermektedir. Ayrıca, akademik literatürde yayımlanmış ulusal ve uluslararası makaleler, yüksek lisans ve doktora tezleri, sektörel raporlar ve konferans bildirileri de çalışmanın ikincil veri kaynakları arasında yer almaktadır. Bu kaynaklar, gemi inşa sanayinde iş güvenliği, risk faktörleri, kaza önleme yöntemleri ve iş kazalarının analizi gibi konularda zengin bir kuramsal çerçeve sunmuş; modelleme sürecinde izlenen yöntemin sağlam temellere oturtulmasını sağlamıştır. Böylece hem veri kümesinin geçerliliği hem de kullanılan analitik yöntemlerin uygunluğu açısından literatüre dayalı destek sağlanmıştır. İkincil kaynaklar ayrıca, birincil verilerin yorumlanmasında karşılaştırmalı analiz olanağı sunarak, çalışmanın bulgularının daha geniş bir bağlamda değerlendirilmesine katkıda bulunmuştur.

3.2 Veri kümesi

Bu araştırmada kullanılan veri kümesi, İstanbul Tuzla Tersaneler Bölgesi'nde faaliyet gösteren çeşitli tersanelerde meydana gelen 2020 iş kazasına ilişkin özgün ve detaylı bilgileri içermektedir. Veriler, İSG birimlerinden temin edilen resmi kayıt formları ve olay bildirim dokümanları doğrultusunda yapılandırılmıştır. Tablo 1.'de gösterildiği gibi her bir kayıt; kazanın meydana geldiği tarih ve saat, kazazedenin yaşı, görevi, bağlı olduğu işveren, uzmanlık sınıfı, kazanın gerçekleştiği yer, yaranın vücuttaki yeri, kaza türü, kaza nedeni, sonucu ve önlenebilirlik düzeyi olmak üzere 17 değişken içermektedir. Ayrıca

olayın gerçekleştiği ay, gün ve işe başlama saatleri gibi zamansal veriler ile tersane tipi gibi çevresel faktörler de kapsam dâhilindedir.

Tablo 1. Değişken Türleri ve Kategorik Dağılım Tablosu

	Değişkenin Adı	Değişkenin Türü	Kategori Sayısı
1	Kaza Saati	Kategorik	28
2	Yaş	Sayısal	9
3	Bağlı Olunan işveren	Kategorik	4
4	Uzmanlık Sınıfı	Kategorik	8
5	Kazazedenin Görevi	Kategorik	42
6	Kazalı Sayısı	Sayısal	5
7	Kaza Yeri	Kategorik	8
8	Yaranın Vücuttaki Yeri	Kategorik	11
9	Kaza Türü	Kategorik	24
10	Kaza Sebebi	Kategorik	16
11	Önlenebilirlik	Kategorik	2
12	Kaza Sonucu	Kategorik	10
13	Kazanın Oluşması	Kategorik	2
14	Tersane Tipi	Kategorik	2
15	İşe Başlama Saati	Kategorik	11
16	Olay Günü	Kategorik	7
17	Olay Ayı	Kategorik	12

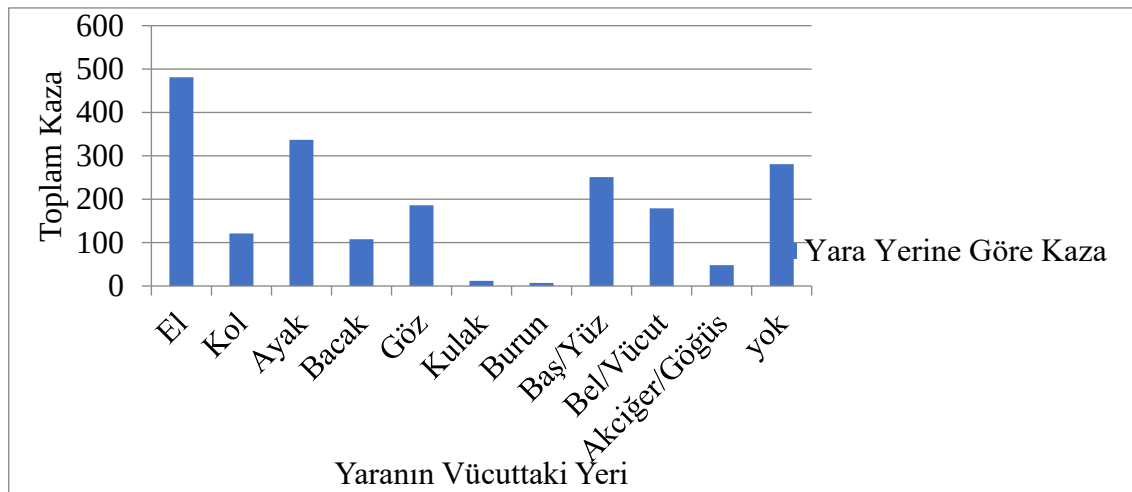
Araştırmada kullanılan veri kümesi, yapılandırılmış ve analiz odaklı bir formatta hazırlanmış olup toplamda 17 adet niteliksel ve niceliksel değişken içermektedir. Ek 1b’de verilen tabloda değişkenlere ait kodların işlendiği örneklem yer almaktadır. Bu tablo Ek 1b.’de yer almaktadır. Çalışmada kullanılan veri kümesinin toplam hücre sayısı 34.340’tır. Her bir kaydın tek bir iş kazasına karşılık geldiği gözlem birimlerinden oluşan bu set, kazanın fiziksel zamanı (kaza saati, olay günü, olay ayı), zamansal çevresi (işe başlama saati), bireysel özellikler (kazazedenin yaşı, görevi, uzmanlık sınıfı), mekânsal unsurlar (kazanın gerçekleştiği yer, tersane tipi), olayın niteliği (kaza türü, yaralanmanın vücuttaki yeri, kaza sonucu, önlenebilirlik durumu) ve yönetsel bilgiler (bağlı olunan işveren) gibi birçok boyutu kapsamaktadır.

3.2.1 Tanımsal İstatistikler

Bu tez kapsamında sunulan grafikler, veri kümesinde yer alan ham verilerin doğrudan görselleştirilmesi yoluyla elde edilmiştir. Grafiklerin her biri, belirli bir kategorik değişkenin-“kaza türü”, “yaranın vücuttaki yeri”, “kazazedenin görevi” ve “kaza ayı” gibi-sınıf içi frekans dağılımını temsil etmektedir. Bu görselleştirmeler herhangi bir istatistiksel işlem, modelleme veya dönüştürme uygulanmadan, veri kümesinin mevcut yapısı temel alınarak üretilmiştir.

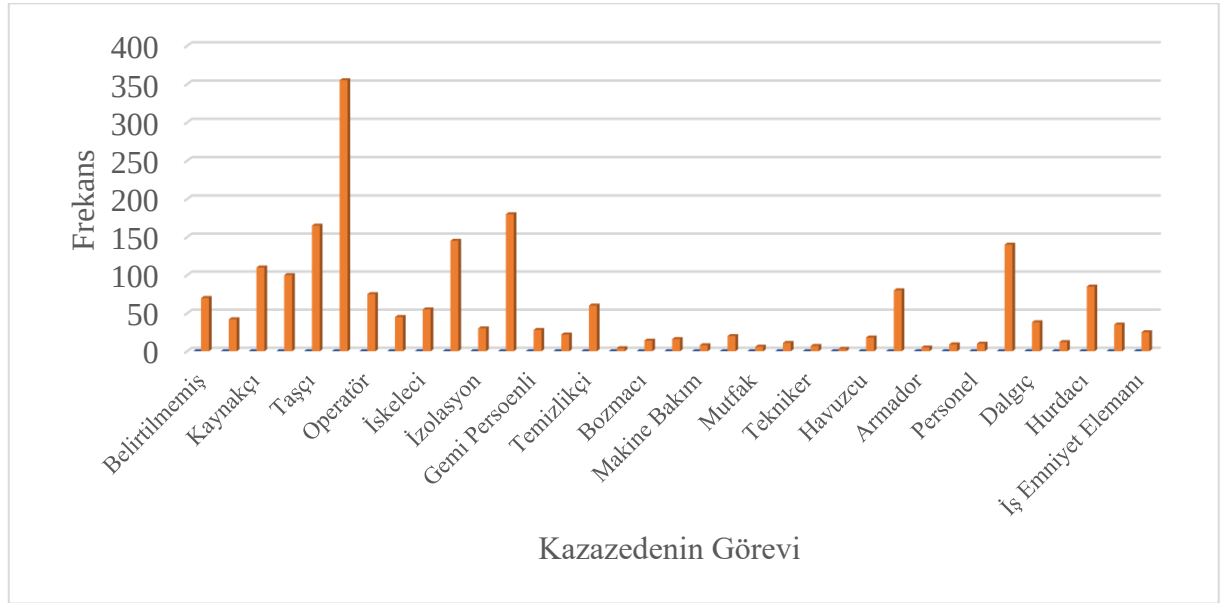
Grafikler, veri kümesinin içeriğinden kaza türleri, yaralanma bölgelerini veya görevlerin risk seviyesini özellikle hangi kaza türlerinin daha sık yaşandığını, hangi görev gruplarının daha riskli olduğunu veya hangi vücut bölgelerinin daha fazla etkilendiğini anlaşılır bir biçimde ortaya koymaktır. Bu görselleştirme süreci, aynı zamanda sınıf dengesizliklerinin saptanmasını sağlayarak, makine öğrenmesi modelleri için önemli olan veri dağılım yapısının önceden değerlendirilmesine olanak tanımaktadır. Bazı sınıfların gözlem sayısı bakımından diğerlerine göre belirgin şekilde fazla olması, model performansı üzerinde doğrudan etkili olabilecek bir unsur olarak öne çıkmaktadır. Bu istatistikler, İSG açısından kritik olabilecek ön risk sinyallerinin belirlenmesine de katkı sunmaktadır. El yaralanmalarının diğer bölgelere kıyasla daha sık görülmesi, el koruyucularının kullanımında eksiklikler olabileceğine işaret etmekte; benzer şekilde kazazede profillerinin dağılımı, iş güvenliği eğitimlerinin hangi gruplara öncelik verilerek planlanması gerektiği konusunda fikir vermektedir.

Grafik 2’de en sık yaralanan vücut bölgeleri arasında baş, el ve ayak öne çıkmaktadır. Özellikle el yaralanmalarının frekansı diğerlerine göre dikkat çekici şekilde yüksektir.



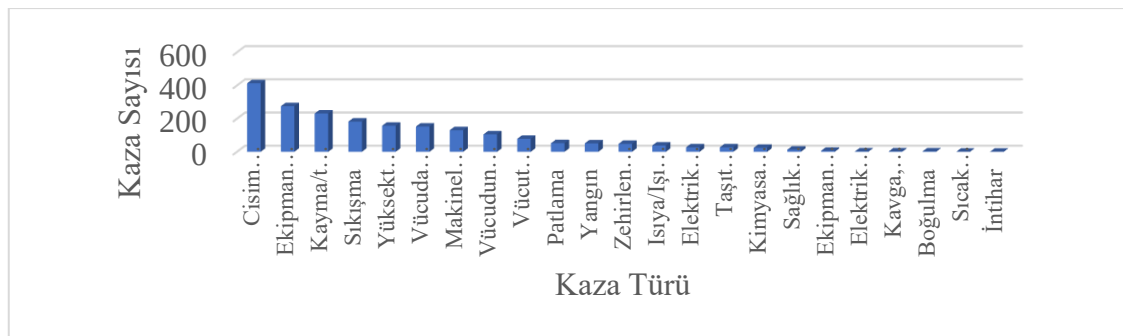
Grafik 2. Yaranın Vücuttaki Yeri Sınıf Dağılımı

Grafik 3'te yer alan kazaya karışan personel görevleri incelendiğinde, en fazla kazaya maruz kalanların "kaynakçı", "montaj işçisi" ve "temizlik personeli" olduğu gözlenmektedir. Bu görevler genellikle fiziksel temasın yoğun olduğu ve çalışma ortamı risklerinin yüksek olduğu alanlarda yer almaktadır.



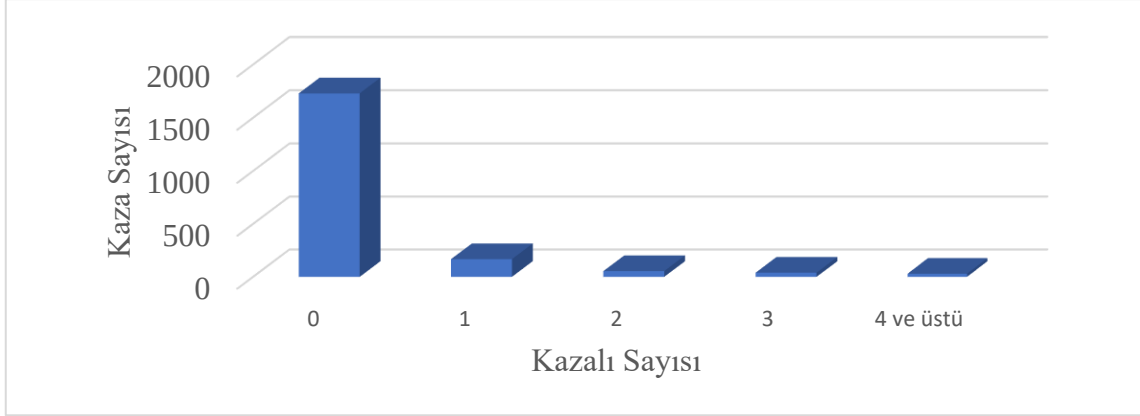
Grafik 3. Kazazedenin Görevi Sınıf Dağılımı

Grafik 4'te görüldüğü gibi en sık karşılaşılan kaza türleri başında “cisim çarpması, malzemelerin düşmesi, batması, ezmesi, kesmesi, yakması” gelmektedir. Bu, iş sahasında malzeme taşıma, üstten yük düşmesi, kontrolsüz ekipman hareketi gibi durumların ciddi bir risk oluşturduğunu gösterir. “Ekipmanların kesmesi, çizmesi, delmesi, çarpması, yakması”, “kayma/takılma düşme”, “sıkışma”, “yüksekten düşme”, “vücuda cisim/çapak kaçması”, “makinelerin sebep olduğu kaza” şeklinde sıralanmaktadır. Bu gruptaki kazalar, iş yerinde mekanik hareketler, dengesizlik ve donanım güvenliği eksikliğinden kaynaklanmaktadır. Önleyici tedbirlerin ekipman koruyucuları, bariyer sistemleri ve güvenli çalışma prosedürleriyle güçlendirilmesi gerekir.



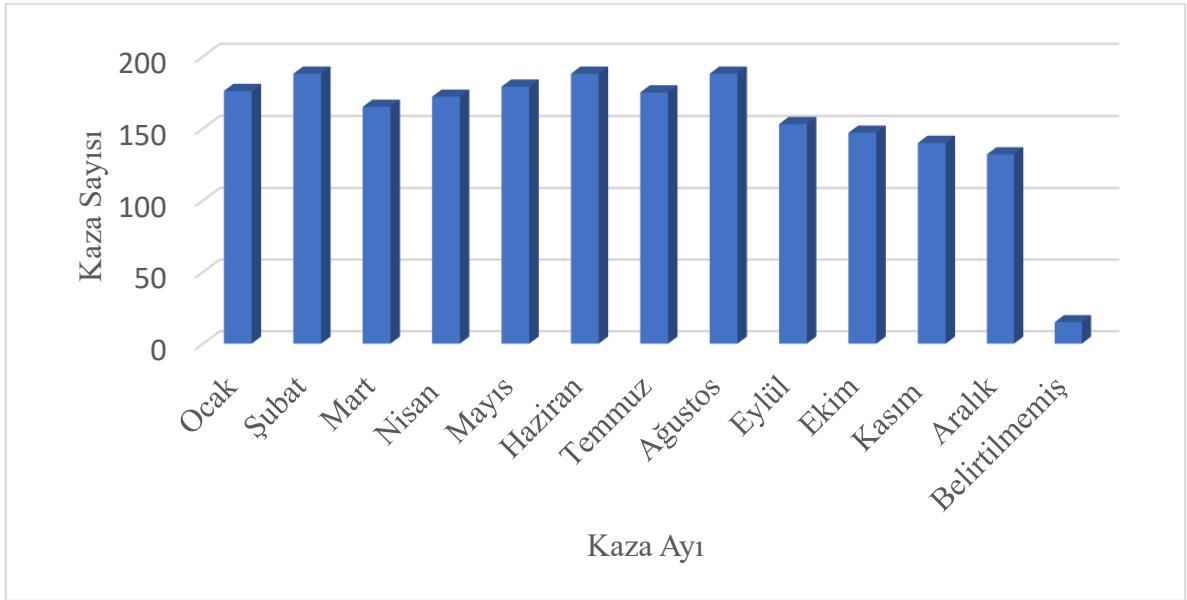
Grafik 4. Kaza Sayısına Göre Kaza Türleri

Grafik 5'te çoğu kazanın tek kişi ile sınırlı olduğu görülmektedir. Bununla birlikte, birden fazla kişinin aynı anda yaralandığı kazalar da raporlanmıştır. Bu tür kazalar genellikle sistematik hatalardan ya da koordinasyon eksikliğinden kaynaklanmaktadır.



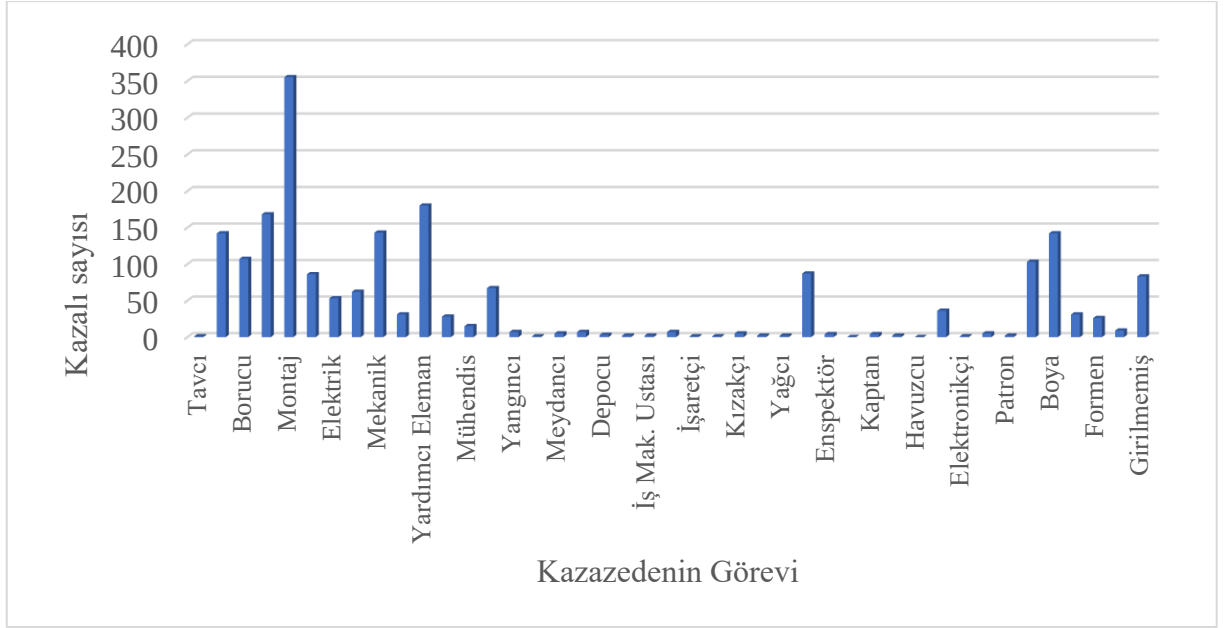
Grafik 5. Kazalı Sayısı Dağılımı

Grafik 6'te kazaların yıl içinde dağılımı incelendiğinde, yaz aylarında (özellikle Temmuz ve Ağustos) artış gözlenmektedir. Bu artışın nedeni olarak sıcak havanın fiziksel performansı düşürmesi, dikkatsizliğe yol açması veya iş yükünün mevsimsel olarak artması düşünülebilir.



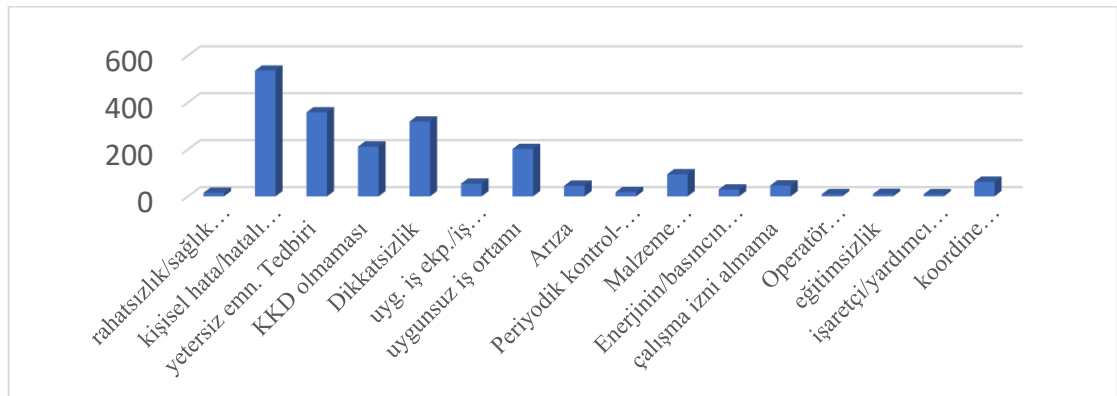
Grafik 6. Aylara Göre Kaza Dağılımı

Grafik 7'da kazaya karışan kişilerin büyük çoğunluğu "vasıfsız işçi" ve "yardımcı personel" sınıfında yer almaktadır. Bu durum, eğitim düzeyi ile kaza sıklığı arasında anlamlı bir ilişki olabileceğini ve iş güvenliği eğitimlerinin bu gruplara öncelikli olarak verilmesi gerektiğini göstermektedir.



Grafik 7.Görev Türüne Göre Kazalı Sayısı Dağılımı

Grafik 8’de kişisel hata/hatalı kullanım toplam 536 vaka ile en sık karşılaşılan kaza sebebidir. 358 vaka ile yetersiz emniyet tedbiri, 319 vaka ile dikkatsizlik, 212 vaka ile KKD (baret, eldiven, gözlük vb.) olmaması, eksikliği veya yanlış kullanımı, 202 vaka ile Uygun olmayan iş ortamı Aydınlatma, havalandırma, düzen gibi fiziksel şartların uygunsuzluğu 94 vaka ile Malzeme hasarı/bozukluğu kaza sebebi olarak görülmekte 62 vaka ile Koordine eksikliği/takipsizlik gelmektedir. 54 vaka ile Uygun iş ekipmanı/KKD kullanılmaması ekipman eksikliği veya koruyucu tedbirlerin kasıtlı ihmal edilmesi önemli bir risk faktörüdür. Nadir ama önemli sebepler 46 vaka ile çalışma izni almama, 45 vaka ile arıza, 18 vaka ile periyodik bakım eksikliği kazalarla ilişkilidir. 29 vaka ile enerji/basınç kesilmemesi makine ve sistem kaynaklı ciddi kazalar oluşturabileceğini göstermektedir. Çok az görülen diğer sebepler 10 vaka ile eğitimsizlik, 9 vaka ile operatör belgesi olmaması, 9 vaka ile işaretçi/yardımcı eksikliği olduğu görülmektedir.



Grafik 8. Kaza Türü Dağılımı

3.2.2 Veri Kümesi Özellikleri

Bu veri yapısı Python ortamında sayısallaştırılmış ve makine öğrenimi modellerine aktarılmıştır. Bu çok boyutlu yapı hem makine öğrenmesi algoritmalarının eğitilmesinde hem de SF-Fuzzy TOPSIS yöntemiyle yapılan çok kriterli karar analizlerinde yüksek düzeyde detay sunarak analizlerin doğruluk ve açıklayıcılığını artırmıştır. Veri kümesinde yer alan her bir gözlem birimi (iş kazası), numerik ve kategorik değişkenlerin dengeli dağılımını içerecek şekilde hazırlanmış; verilerin analiz öncesinde sayısal formatlara dönüştürülmesi sağlanmıştır. Veri kümesindeki değişkenlerin yaklaşık %70'i kategorik, %30'u sayısal niteliktedir. Kategorik değişkenler etiketlenmiş, sayısal değişkenler ise standartlaştırılmıştır. Kodlama sisteminde, her değişkene özgü numerik karşılıklar verilmiş ("Kazazedenin Görevi" için gemi personeli = 60, mühendis = 62), bu yapı sayesinde hem makine öğrenimi hem de bulanık mantık tabanlı analizler için uygun bir veri temeli oluşturulmuştur. Dolayısıyla bu veri kümesi, sektörel gerçekliği yansıtan ve araştırma sorularının istatistiksel ve karar analizi perspektiflerinden cevaplanmasına olanak tanıyan güçlü bir temel sağlamaktadır. Modellemelerde kullanılan veri kümesinin yapısı Ek 1'da yer alan tabloda gösterilmektedir. Bu tablo, SHAP, LIME ve Fuzzy-TOPSIS gibi ileri analizlerin temelini oluşturmaktadır.

3.2.3 Veri kümesi Hazırlama Süreci

Veri kümesi hem dengesiz sınıfları barındırmakta hem de eksik veya hatalı gözlemler içerebilmektedir; bu nedenle analiz öncesi ön işleme adımlarına tabi tutulmuştur. Verinin hacmi ve yapısı, istatistiksel modelleme, sınıflandırma algoritmaları ve ÇKKV yöntemleri için yeterli çeşitlilik ve kapsayıcılığı sağlamaktadır. Ayrıca, veri kümesinin gerçek saha koşullarına dayanması, yapılan analizlerin akademik ve sektörel geçerliliğini artırmakta; kazaların önlenmesi ve iş güvenliği stratejilerinin geliştirilmesine katkı sağlayarak değerli bir kaynak sunmaktadır.

Veri kümesinin analize hazır hale getirilme süreci, sistematik ve çok aşamalı bir yaklaşım çerçevesinde gerçekleştirilmiştir. İlk olarak, tersanelerden elde edilen ham veriler dijital ortama aktarılmış ve analiz için uygun bir veri tabanı oluşturulmuştur. Bu ilk aşamada, tüm değişkenlerin biçimsel uyumu sağlanmış, veri girişindeki tutarsızlıklar giderilmiş ve metin formatındaki kategorik veriler tutarlı bir biçimde standardize edilmiştir. Kodlama şemaları önceden oluşturularak, "Kazazedenin Görevi" gibi kategorik değişkenler belirli sayısal kodlarla temsil edilecek şekilde yapılandırılmıştır. Böylece, makine öğrenmesi modelleri için gerekli olan sayısal formatın temeli atılmıştır.

Kategorik deęişkenlerin sayısallaştırılması, modelleme aşamasına geçmeden önce gerçekleştirilmiştir. Sayısal olmayan tüm deęişkenler, veri tiplerine ve model gereksinimlerine uygun olarak etiketleme (Label Encoding) ya da gerektiğinde One-Hot Encoding gibi tekniklerle dönüştürülmüştür. Bu işlem, makine öğrenmesi algoritmalarının doğru ve etkin şekilde çalışabilmesi için zorunlu bir adımdır.

Sınıf dengesizliği problemi, hedef deęişkenin bazı sınıflarında yeterli temsil olmaması nedeniyle gözlemlenmiştir. Bu dengesiz dağılım, sınıflar arası öğrenme başarımını olumsuz etkileyeceęi için SMOTE (Synthetic Minority Over-sampling Technique) yöntemi kullanılarak dengelenmiştir. Böylece azınlık sınıflar yapay örneklerle zenginleştirilmiş, modelin her sınıfı öğrenme yetisi artırılmıştır.

Model geliştirme öncesi son adım olarak, veri kümesi eğitim ve test alt kümelerine ayrılmış, bu işlem sırasında **train_test_split** fonksiyonu kullanılmıştır. Model başarımını güvenilir şekilde değerlendirebilmek amacıyla ayrıca K-Fold Cross Validation gibi çapraz doğrulama yöntemleri için uygun yapılandırmalar oluşturulmuştur. Bu kapsamlı ön işleme süreci, hem makine öğrenimi algoritmalarının performansını artırmış hem de SF-Fuzzy TOPSIS analizine uygun, tutarlı ve homojen bir veri yapısı ortaya koymuştur.

Sonraki adımda, veri kümesi, .xlsx formatında olup analiz süreçleri Python programlama dili aracılığıyla Google Colab ortamında gerçekleştirilmiştir. Veri kümesinin Pandas Kütüphanesi kullanılarak Python ortamına okunmasına ilişkin kod Ek 1c'de gösterilmiştir.

Tüm bu ön işleme adımları, model doğruluęunu ve genelleme kapasitesini yükseltmekle birlikte, aynı zamanda verinin ÇKKV sürecine entegre edilebilirliğini de sağlamıştır. Bu sayede, makine öğrenmesi ve SF-Fuzzy TOPSIS gibi yöntemlerin uygulanmasında istatistiksel tutarlılığı yüksek, homojen ve anlamlı bir veri yapısı oluşturulmuştur.

3.2.4 Eksik Veri Yönetimi

Veri kümesinde yer alan eksik deęerlerin uygun şekilde yönetilmesi, analizlerin güvenilirliği ve model performansının sürdürülebilirliği açısından oldukça önemlidir. Bu çalışmada, eksik verilerin analize olumsuz etkisini en aza indirmek amacıyla sistematik bir eksik veri tamamlama yaklaşımı benimsenmiştir. Öncelikle, veri kümesi üzerinde kapsamlı bir inceleme yapılarak eksik veri içeren gözlemler ve sütunlar tespit edilmiştir. Eksikliklerin özellikle bazı kategorik deęişkenlerde yoğunlaştığı görülmüştür. Bu

durumda, eksik veri oranı yüksek olan değişkenlerin çıkarılması yerine, eksik alanların doldurulması tercih edilmiştir.

Eksik veriler, Python programlama dili içerisinde yer alan SimpleImputer sınıfı yardımıyla doldurulmuştur. Kullanılan strateji, “**most_frequent**” yani ilgili bölümde yer alan en sık tekrar eden değeri eksik değerlere atanarak aşağıda belirtilen şekilde yapılandırılmıştır.

Bu yöntem, özellikle kategorik değişkenler için uyumlu olup, veri dağılımını bozmadan eksikliklerin giderilmesini sağlamaktadır. Sayısal verilerde ise eksiklikler çok düşük oranlarda olduğundan, aynı strateji ile doldurma işlemi gerçekleştirilmiş ve verinin istatistiksel yapısı korunmuştur. Ayrıca, veri kümesi üzerinde yapılan analizlerde eksik verilerin rastgele dağılmış olduğu varsayılmış ve bu nedenle herhangi bir ileri düzey çok değişkenli tahmin tekniğine gerek duyulmamıştır.

Bu eksik veri yönetimi süreci sayesinde, modelleme aşamasında yaşanabilecek performans kayıpları önlenmiş, veri kümesi tam ve analize uygun hale getirilmiştir. Aynı zamanda, eksik veri kaynaklı sapmaların önüne geçilerek, özellikle sınıflandırma algoritmalarının eğitimi sırasında kararsızlık yaratabilecek durumların bertaraf edilmesi sağlanmıştır. Bu bağlamda uygulanan yöntem, veri bütünlüğünü korurken, model başarımına da doğrudan katkı sunmuştur.

Eksik veriler tamamlandıktan sonra, veri kümesindeki tüm değişkenler türlerine göre gruplandırılmıştır. Sayısal ve kategorik değişkenler ayrılarak, her biri için uygun dönüşüm ve modelleme stratejilerinin uygulanmasına zemin hazırlanmıştır.

3.2.5 Veri Normalizasyonu ve Dönüştürme

Veri kümesinin analizde kullanılabilir hale gelmesinde önemli adımlardan biri de normalizasyon ve dönüştürme süreçleridir. Özellikle makine öğrenmesi algoritmalarının etkinliği, veri formatının bu algoritmaların beklentilerine uygun olup olmamasıyla doğrudan ilişkilidir. Bu nedenle, çalışmada kullanılan veri kümesi üzerinde çeşitli dönüştürme ve ölçekleme işlemleri uygulanmıştır. Öncelikle, tüm kategorik veriler Python kütüphaneleri aracılığıyla “Label Encoding” yöntemiyle sayısal değerlere çevrilmiş; gerek görülen durumlarda “One-Hot Encoding” ile sınıflar ayrı sütunlara ayrılarak daha hassas bir veri temsili sağlanmıştır.

Sayısal değişkenler için ise “Min-Max Normalizasyonu” yöntemi benimsenmiş; tüm sayısal değerler 0 ile 1 arasında yeniden ölçeklendirilmiştir. Bu işlem, özellikle mesafe tabanlı algoritmalarda (KNN, SVM) veriler arası benzerlik ölçümlerinin sağlıklı yapılabilmesini mümkün kılmıştır. Ayrıca, değişkenler arasında ölçek farkının model performansını olumsuz etkilemesini önlemek amacıyla, bazı senaryolarda “Z-Score Standardizasyonu” da uygulanmıştır.

Veri dönüştürme sürecinde dikkat edilen bir diğer husus da dönüşüm sonrası bilgi kaybının yaşanmaması olmuştur. Kategorik verilerin anlamlılığını koruyacak biçimde kodlanması, istatistiksel analizlerin yanı sıra açıklanabilir yapay zekâ yaklaşımlarında da (SHAP ve LIME) yorumlanabilirliği artırmıştır. Bu işlemler, yalnızca makine öğrenimi için değil, aynı zamanda SF-Fuzzy TOPSIS gibi çok kriterli karar destek sistemlerine uygunluğu da güçlendirmiştir. Sonuç olarak, uygulanan veri normalizasyonu ve dönüştürme adımları, verinin istatistiksel tutarlılığına katkı sunarken, algoritmaların veriyi daha verimli işlemesini mümkün kılmıştır.

3.3. Makine Öğrenimi Modelleri

Makine öğrenimi, verilerden öğrenen ve bu verilerden elde edilen bilgilere dayalı olarak tahminlerde bulunan algoritmaları ve istatistiksel modelleri kullanan yapay zekânın alt dalıdır. Makine öğrenimi algoritmaları, özellikle karmaşık ve doğrusal olmayan ilişkilerin bulunduğu durumlarda büyük veri setlerinden anlamlı bilgiler çıkarmak için yaygın olarak kullanılmaktadır (Martín-Martín ve ark., 2023).

Makine öğrenimi, büyük veri kümelerinden bilgi çıkarma, tahmin yapma ve karar verme amacıyla veri temelli modelleri oluşturmak ve eğitmek için kullanılan bir yapay zekâ alanıdır. Bu alan, özellikle karmaşık yapıya sahip olan ve geleneksel yöntemlerle modellenmesi güç olan sorunlarda güçlü bir alternatif sunmaktadır. Makine öğrenimi teknikleri, verilerdeki örüntüleri ve ilişkileri belirleyerek bu bilgileri gelecekteki olayları tahmin etmekte kullanır. Makine öğrenimi yöntemleri temel olarak üç ana kategoriye ayrılır: denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli öğrenme. Denetimli öğrenmede, veriler etiketlidir ve algoritma giriş-çıkış ilişkisini öğrenir. Denetimsiz öğrenmede, veriler etiketsizdir ve algoritma verilerin altında yatan yapıyı keşfetmeye çalışır. Pekiştirmeli öğrenmede ise algoritma, ödül ve cezalarla belirli bir hedefe doğru ilerleyerek öğrenme gerçekleştirir (Matías ve ark., 2008). Denetimsiz öğrenme, veri kümesindeki gizli yapı ve örüntüleri etiketleme olmadan keşfetmeyi amaçlar (Martín-

Martín et al., 2023). Pekiştirmeli öğrenme (RL), çevreden alınan geri bildirimlerle karar verme süreçlerini geliştirerek özellikle siber güvenlikte tehdit tespiti ve otonom savunma mekanizmaları oluşturur (Louati ve ark., 2024; Olutimehin, 2025)

Bu araştırmada, gemi inşa sektöründe meydana gelen iş kazalarının nedenlerini tahmin etmek amacıyla çeşitli denetimli makine öğrenmesi algoritmaları uygulanmıştır. Sınıflandırma problemini çözmek üzere; Naive Bayes, K-En Yakın Komşu (KNN), Karar Ağaçları, Lojistik Regresyon, Destek Vektör Makineleri (SVM), Random Forest, Extra Trees, AdaBoost ve Gradient Boosting algoritmalarından yararlanılmıştır. Model seçiminde hem literatürdeki başarı düzeyleri hem de veri kümesinin yapısal özellikleri göz önünde bulundurulmuştur.

3.3.1. Sınıflandırma Problemlerinde Başarı Metrikleri

Literatürde doğruluğun “en yaygın değerlendirme metriği” olduğu vurgulanır; özellikle sınıf dağılımı dengeli ve yanlış-pozitif ile yanlış-negatif değişken dağılımı simetrik olan senaryolarda model başarısını net biçimde yansıttığı belirtilir (Xu ve ark., 2022). Duyarlılık gerçek kaza kayıtlarının ne kadarının doğru yakalandığını, kesinlik ise modelimin “kaza var” dediği durumların ne kadarının gerçekten kazayla sonuçlandığını göstermektedir. Harmonik ortalamaya dayalı F1 skoru, bu iki ölçütü dengeli biçimde birleştirerek özellikle “kaçırılan kaza” maliyetinin yüksek olduğu senaryolarda daha güvenilir bir özet sunmaktadır (Powers, 2020).

3.4. Küresel Fuzzy-TOPSIS Yönteminin Uygulanması (SF-TOPSIS)

Bu çalışmada, makine öğrenmesi algoritmaları ile elde edilen bulguların ÇKKV süreçleriyle desteklenmesi amacıyla SF-Fuzzy TOPSIS yöntemi uygulanmıştır. Bu yöntem, belirsizlik içeren ve niteliksel değerlendirmelere dayanan karar problemlerinde etkili sonuçlar sunabilen bir analiz tekniğidir. Çalışmada SF-Fuzzy TOPSIS yaklaşımı, özellikle makine öğrenmesi modellerinin belirlediği en önemli faktörlerin insan yargıları içeren bir önceliklendirilmesinin yapılması amacıyla kullanılmıştır.

Küresel bulanık mantık tabanlı SF-TOPSIS uygulaması için öncelikle uzmanlardan her bir kriterin önem düzeyine ilişkin değerlendirmeler alınmıştır. Bu değerlendirmeler, küresel bulanık sayılar (Spherical Fuzzy Numbers - SFN) şeklinde ifade edilerek bulanık karar matrisine dönüştürülmüştür. Uzman görüşleri, sektörde doğrudan deneyimi olan İSG uzmanları, mühendisler ve yöneticilerden toplanmıştır. Her bir uzman, her kriterin

önemini "düşük", "orta" ve "yüksek" gibi sözel terimlerle değerlendirmiş; bu terimler daha sonra literatürde kabul gören bulanık sayı karşılıklarına dönüştürülmüştür.

Elde edilen bulanık karar matrisi doğrultusunda, SF-Fuzzy TOPSIS adımları izlenerek normalize matris, ağırlıklı normalize matris, ideal ve negatif ideal çözümler belirlenmiştir.

Her bir alternatifin bu ideal çözümlere uzaklıkları hesaplanarak, tercih edirlilik skoru elde edilmiştir. Bu skorlar, faktörlerin görece önem düzeylerini yansıtmaktadır.

Bu yöntem, sadece nicel sonuçlar sunmakla kalmamış, aynı zamanda uzman bilgisi ve deneyimini karar süreçlerine entegre etme fırsatı da vermiştir. SF-Fuzzy TOPSIS'in kullanımı, klasik sayısal analizlerin ötesine geçerek daha bütüncül ve açıklanabilir bir karar süreci sunmuş; elde edilen sıralama ise entegre modelde daha sonraki analizlerin temelini oluşturmuştur.

3.4.1 Kriterlerin Belirlenmesi

SF-Fuzzy TOPSIS yönteminin uygulanabilmesi için ilk adım olarak analizde kullanılacak karar kriterlerinin belirlenmesi gerekmektedir. Bu çalışmada kriter seçimi hem makine öğrenmesi algoritmalarının çıktılarına hem de sektörel uzman görüşlerine dayanarak gerçekleştirilmiştir. Random Forest ve Extra Trees algoritmaları ile yapılan değişken önem derecelendirmesi sonucunda Tablo 4'te gösterilen en yüksek etkili dört faktör belirlenmiştir: kazazedenin görevi, kaza türü, yaranın vücuttaki yeri ve Olay ayı. Bu faktörler, iş kazalarının oluşumuna etki eden başlıca unsurlar olarak hem istatistiksel hem de mesleki deneyim açısından desteklenmektedir.

Tablo 2. SF-Fuzzy TOPSIS'te Kullanılan En Etkili Dört Faktör

Faktör Adı
Kazanın Türü
Yaranın Vücuttaki Yeri
Kazazedenin Görevi
Olay ayı

Uzmanlardan alınan nitel değerlendirmelerle desteklenen bu kriterler, SFN yapısı kullanılarak karar matrisine dönüştürülmüştür. Her bir uzmana, bu kriterlerin iş kazası risk düzeyini belirlemedeki görece önem derecesi sorulmuş; yanıtlar "çok düşük" ten "çok

yüksek"e kadar değişen linguistik ifadelerle alınmıştır. Bu ifadeler daha sonra literatürde tanımlanmış küresel bulanık sayı karşılıklarına dönüştürülerek sayısal analiz sürecine dâhil edilmiştir.

Kriterlerin belirlenmesinde temel amaç, karar verme sürecinde subjektif uzman görüşü ile nesnel veri analizinin bütünleştirilmesidir. Böylece SF-Fuzzy TOPSIS uygulamasında kullanılan her bir kriter hem analitik olarak doğrulanmış hem de sektörel bilgiyle desteklenmiştir. Bu yaklaşım, karar sürecinin şeffaflığını ve güvenilirliğini artırmakta; elde edilen sıralamanın sektördeki pratik koşulları yansıtmasını mümkün kılmaktadır.

3.4.2 Fuzzy Mantık ve TOPSIS Süreci

SF-Fuzzy TOPSIS süreci, küresel bulanık mantık çerçevesinde ÇKKV adımlarının sistematik biçimde uygulanmasını kapsamaktadır. Bu yöntemin temel amacı, alternatiflerin küresel bulanık sayılarla ifade edilmiş kriter ağırlıklarına göre sıralanmasını sağlamaktır. Uygulamada ilk adım, Tablo 3'te yer alan uzman değerlendirmeleriyle oluşturulan küresel bulanık karar matrisinin normalize edilmesidir.

Tablo 3. Uzman Değerlendirme Matrisi

	Etkilenme		Kontrol		
	Olasılık	Şiddet	Süresi	Edilebilirlik	Sıklık
KARAR VERİCİ-1	Kriter 1	Kriter 2	Kriter 3	Kriter 4	Kriter 5
Kazanın Türü	EI	AMI	VHI	HI	AMI
Yaranın Vücuttaki Yeri	SMI	VHI	HI	SLI	EI
Kazazedenin Görevi	VHI	VHI	LI	VHI	VHI
Olay Ayı	VLI	ALI	VLI	ALI	VLI
KARAR VERİCİ-2					
Kazanın Türü	AMI	VHI	VHI	AMI	AMI
Yaranın Vücuttaki Yeri	HI	HI	HI	VLI	VHI
Kazazedenin Görevi	HI	EI	SLI	SMI	SMI
Olay Ayı	ALI	LI	LI	VLI	ALI
KARAR VERİCİ-3					
Kazanın Türü	AMI	VHI	AMI	VHI	VHI
Yaranın Vücuttaki Yeri	VHI	VLI	HI	HI	VHI
Kazazedenin Görevi	HI	VHI	LI	SMI	LI
Olay Ayı	ALI	VLI	LI	ALI	ALI

Normalize işleminde, her kriter için değerler belirli bir ölçeğe indirgenerek karşılaştırılabilir hâle getirilmiştir.

Tablo 4. Dilsel Terimler ve Bunlarla İlişkili Küresel Bulanık Sayılar

Dilsel Terimler	(μ, ν, π)
Kesinlikle Daha Önemli (AMI)	(0.9, 0.1, 0.1)
Çok Yüksek Önem (VHI)	(0.8, 0.2, 0.2)
Yüksek Önem (HI)	(0.7, 0.3, 0.3)
Biraz Daha Önemli (SMI)	(0.6, 0.4, 0.4)
Eşit Önem (EI)	(0.5, 0.5, 0.5)
Biraz Düşük Önem (SLI)	(0.4, 0.6, 0.4)
Düşük Önem (LI)	(0.3, 0.7, 0.3)
Çok Düşük Önem (VLI)	(0.2, 0.8, 0.2)
Kesinlikle Düşük Önem (ALI)	(0.1, 0.9, 0.1)

Tablo 5. Karar Vericilerin Önemini Derecelendirmek İçin Kullanılan Dilsel Terimler

Dilsel Terimler	(μ, ν, π)
Çok Önemli (VI)	(0.9, 0.1, 0.0)
Önemli (I)	(0.75, 0.2, 0.05)
Orta (M)	(0.5, 0.45, 0.05)
Önemsiz (U)	(0.35, 0.6, 0.05)
Çok Önemsiz (VU)	(0.1, 0.9, 0.0)

İkinci aşamada, normalize değerler önceden belirlenen kriter ağırlıklarıyla çarpılarak ağırlıklı normalize karar matrisi elde edilmiştir. Bu aşamada kullanılan ağırlıklar, Tablo 6'da yer alan uzmanların küresel bulanık değerlendirmelerinden türetilmiştir. Daha sonra, ideal ve negatif ideal çözümler belirlenmiştir. İdeal çözüm, her kriter için en yüksek performansı yansıtan değerlerden; negatif ideal çözüm ise en düşük performansı temsil eden değerlerden oluşmaktadır.

Tablo 6. Uzmanların Küresel Bulanık Değerlendirme Matrisi

KARAR VERİCİ-1	Olasılık			Şiddet		
	Kriter 1			Kriter 2		
Kazanın Türü	0,5	0,5	0,5	0,9	0,1	0,1
Yaranın Vücuttaki Yeri	0,6	0,4	0,4	0,8	0,2	0,2

Kazazedenin Görevi	0,8	0,2	0,2	0,8	0,2	0,2
Olay Ayı	0,2	0,8	0,2	0,1	0,9	0,1
KARAR VERİCİ-2						
Kazanın Türü	0,9	0,1	0,1	0,8	0,2	0,2
Yaranın Vücuttaki Yeri	0,7	0,3	0,3	0,7	0,3	0,3
Kazazedenin Görevi	0,7	0,3	0,3	0,5	0,5	0,5
Olay Ayı	0,1	0,9	0,1	0,3	0,7	0,3
KARAR VERİCİ-2						
Kazanın Türü	0,9	0,1	0,1	0,8	0,2	0,2
Yaranın Vücuttaki Yeri	0,8	0,2	0,2	0,2	0,8	0,2
Kazazedenin Görevi	0,7	0,3	0,3	0,8	0,2	0,2
Olay Ayı	0,1	0,9	0,1	0,2	0,8	0,2

	Etkilenme Süresi			Kontrol Edilebilirlik			Sıklık		
KARAR VERİCİ-1	Kriter 3			Kriter 4			Kriter 5		
Kazanın Türü	0,8	0,2	0,2	0,7	0,3	0,3	0,9	0,1	0,1
Yaranın Vücuttaki Yeri	0,7	0,3	0,3	0,4	0,6	0,4	0,5	0,5	0,5
Kazazedenin Görevi	0,3	0,7	0,3	0,8	0,2	0,2	0,8	0,2	0,2
Olay Ayı	0,2	0,8	0,2	0,1	0,9	0,1	0,2	0,8	0,2
KARAR VERİCİ-2									
Kazanın Türü	0,8	0,2	0,2	0,9	0,1	0,1	0,9	0,1	0,1
Yaranın Vücuttaki Yeri	0,7	0,3	0,3	0,2	0,8	0,2	0,8	0,2	0,2
Kazazedenin Görevi	0,4	0,6	0,4	0,6	0,4	0,4	0,6	0,4	0,4
Olay Ayı	0,3	0,7	0,3	0,2	0,8	0,2	0,1	0,9	0,1
KARAR VERİCİ-2									
Kazanın Türü	0,9	0,1	0,1	0,8	0,2	0,2	0,8	0,2	0,2
Yaranın Vücuttaki Yeri	0,7	0,3	0,3	0,7	0,3	0,3	0,8	0,2	0,2
Kazazedenin Görevi	0,3	0,7	0,3	0,6	0,4	0,4	0,3	0,7	0,3
Olay Ayı	0,3	0,7	0,3	0,1	0,9	0,1	0,1	0,9	0,1

Bir sonraki adımda, her alternatifin ideal ve negatif ideal çözümlere olan küresel bulanık uzaklıkları hesaplanmıştır. Bu uzaklıklar, küresel bulanık vektörler arasındaki benzerlik ölçütleri kullanılarak elde edilmiştir. Uzaklıklar dikkate alınarak her alternatif için bir tercih skoru (closeness coefficient) hesaplanmış; bu skor, alternatifin ideal çözüme ne kadar yakın olduğunu göstermektedir.

Son adımda, elde edilen tercih skorlarına göre Tablo 9' belirtilen etkenler sıralanmış ve analiz edilen faktörlerin göreceli önem dereceleri ortaya konmuştur.

Bu süreçte elde edilen bulgular, makine öğrenmesi sonuçlarıyla karşılaştırılarak doğrulama sağlanmış ve entegre karar desteği güçlendirilmiştir. SF-Fuzzy TOPSIS süreci, özellikle belirsizlik ve öznel değerlendirmelerin yoğun olduğu İSG gibi alanlarda kararların daha sağlam temellere oturtulmasına katkı sağlamaktadır.

Küresel Fuzzy kümeler, geleneksel bulanık kümelere kıyasla karar vericilerin belirsizlik, üyelik ve üyelik dışı derecelerini aynı anda tanımlayabilmelerine olanak tanır ve bu yönüyle karar verme problemlerinde belirsizlik ve belirsizliğin üstesinden gelmede etkili bir araçtır. Gerçek dünya senaryolarında, karar verme süreci sıklıkla içgüdüsel değerlendirmeler nedeniyle kesin olmayan ve muğlak olabilir. Bu nedenle, karar verme sürecinde risk ve belirsizlik koşullarını ele almak için teoriler ve yöntemler geliştirilmiştir. Bu çerçevede, Fuzzy-ÇKKV teknikleri, bilgi eksikliği, insan sezgilerine dayalı değerlendirmeler ve sübjektif görüşlerin nesnel hale getirilmesi gibi alanlarda avantaj sağlamaktadır. Fuzzy entegrasyonunun bazı sınırlamaları da vardır: "Karar verici tercihleri ve uzmanların öznel yargıları, modelin karar verme sürecini etkileyebilir. SF-Delphi yöntemi uzman görüş birliği sağlasa da karar vericilerin sübjektifliği modelin nesnelliğini sınırlandırabilir." (Nguyen, 2024).

BÖLÜM 4. ANALİZ VE BULGULAR

Bu bölümde, üçüncü bölümde açıklanan yöntemsel çerçeve doğrultusunda uygulamalar gerçekleştirilmiş; makine öğrenmesi modelleri ile SF-Fuzzy TOPSIS yöntemine ilişkin analiz bulguları ayrıntılı olarak sunulmuştur. Öncelikle makine öğrenmesi algoritmalarının doğruluk oranları, hata analizleri ve değişken önem düzeylerine ilişkin sonuçlar paylaşılmış; ardından SF-Fuzzy TOPSIS yöntemi ile belirlenen kriter ağırlıkları ve sıralamaları değerlendirilmiştir. Son olarak bu iki yöntemin entegrasyonundan elde edilen birleşik bulgular ve faktörlerin sektörel açıdan etkileri yorumlanmıştır.

4.1 Modelleme Süreci

Makine öğrenimi modellerinin belirlenmesinde hem literatürdeki başarı düzeyleri hem de veri kümesinin yapısal özellikleri temel alınmıştır. Bu bağlamda, klasik yöntemlerden gelişmiş topluluk modellerine kadar geniş bir algoritma yelpazesi tercih edilmiştir. Naive Bayes ve Lojistik Regresyon gibi geleneksel algoritmalar, özellikle kategorik veri setlerinde etkili sonuçlar üretmeleri nedeniyle seçilmiştir. Öte yandan, Karar Ağaçları ve KNN algoritmaları, yorumlana bilirlilik avantajı ve düşük hesaplama maliyeti nedeniyle dâhil edilmiştir. SVM, doğrusal olmayan sınırları tanıyabilme kapasitesiyle öne çıkarken, AdaBoost ve Gradient Boosting gibi topluluk öğrenme yöntemleri sınıflar arası ayrımı güçlendirme amacıyla değerlendirilmiştir. Bu algoritmaların sonuçlarına göre yüksek doğruluk oranına sahip ilk iki model seçilmiştir.

Bununla birlikte, özellikle Random Forest ve Extra Trees algoritmaları veri kümesi üzerinde sağladıkları yüksek doğruluk oranları, gürültüye karşı dayanıklılık ve değişken önemini hesaplayabilme kapasiteleri sayesinde öne çıkmıştır. Sonuçlar, hedef değişkende pek çok kategori bulunmasına karşın kullanılan modellerin tatmin edici başarı oranlarına ulaştığını ve çok boyutlu veri yapısını yönetirken ek kolaylık sağladığını göstermektedir.

Model seçim süreci yalnızca teknik başarı ölçütlerine değil, aynı zamanda modellerin açıklana bilirlilik seviyelerine de dayanmıştır. Bu kapsamda, SHAP ve LIME gibi açıklanabilir yapay zekâ (XAI) araçlarının uygulanabilirliği göz önünde bulundurularak, özellikle karar ağaçları tabanlı modellerin tercih edilmesine özen gösterilmiştir. Sonuç olarak, seçilen modeller yalnızca sınıflandırma doğruluğu açısından değil, aynı zamanda iş güvenliği stratejileri açısından yorumlanabilir çıktılar üretebilmeleri bakımından da değerli bulunmuştur.

Veri kümesinde “Kaza Sebebi” hedef değişken, geriye kalan 16 değişken ise öz nitelik olarak dikkate alınmıştır. 'Kaza Sebebi' hedef değişkeni diğer değişkenlerden ayrılmıştır.

4.2 Makine Öğrenimi Modellerinin Sonuçları

Uygulanan makine öğrenimi algoritmaları (Random Forest, Extra Trees, Gradient Boosting, KNN, SVM, Naive Bayes vb.) farklı sınıflandırma başarıları göstermiştir. Modellerin genel doğruluk oranları %70 ile %94 arasında değişmiştir. Özellikle Random Forest ve Extra Trees algoritmaları %90'ın üzerinde doğruluk oranı ile öne çıkmıştır. Bu modeller, aynı zamanda değişken önem derecelerinin belirlenmesinde kullanılmış ve kazazedenin görevi, yaranın vücuttaki yeri, olay ayı ve kaza türü gibi değişkenlerin ön plana çıktığını ortaya koymuştur. Model performansları sadece doğruluk üzerinden değil, aynı zamanda hassasiyet, duyarlılık, Denge skoru ve Confusion matrisi gibi metriklerle de değerlendirilmiştir. Tablo 10'da görüldüğü üzere Random Forest modelinde denge skoru 0.91 iken optimize Random Forest denge skoru 0.93 seviyesine ulaşmıştır. Extra Trees modeli denge skoru 0.9260 seviyesinde iken optimize edilmiş extra trees modeli denge skoru ise 0.9388 seviyesine ulaşmıştır. Makine öğrenmesi modellerinin değerlendirilmesinde sonucunda optimize Random Forest denge skoru 0.93 seviyesine, optimize edilmiş extra trees modeli denge skoru ise 0.9388 seviyesinde bir başarı sergilerken, KNN ve Naive Bayes gibi daha basit algoritmaların doğruluk oranı %75-80 aralığında kalmış; bu da bu modellerin veri kümesinin karmaşıklığı karşısında sınırlı performans sergilediğini göstermiştir. Gradient Boosting algoritması da yüksek performans göstermiş; özellikle overfitting eğilimi düşük ve genel doğruluk seviyesi yüksektir. Bu algorithmada skoru 0.89, AUC ise 0,91 seviyelerine ulaşmıştır. Hassasiyet ve duyarlılık oranları açısından ise SVM modeli dengeli sonuçlar sunmuştur. Ancak modelin doğruluk oranı Random Forest'a göre daha düşüktür.

Tablo 7. Tüm Model Performansları

Model	Doğruluk	Hassasiyet	Duyarlılık	F1 Skoru	ROC AUC
Extra Trees	0,94	0,94	0,94	0,94	0,95
Extra Trees (Optimize)	0,94	0,94	0,94	0,94	0,95
Random Forest (Optimize)	0,93	0,93	0,93	0,93	0,94
Random Forest	0,91	0,91	0,91	0,91	0,93

Gradient Boosting	0,89	0,88	0,89	0,89	0,91
AdaBoost	0,88	0,87	0,88	0,88	0,89
SVC	0,86	0,85	0,86	0,85	0,88
Logistic Regression	0,85	0,84	0,85	0,84	0,86
KNN	0,83	0,82	0,83	0,82	0,84
Decision Tree	0,82	0,81	0,82	0,81	0,83
Naive Bayes	0,81	0,79	0,8	0,79	0,83

Bu skorlar Random Forest ve Extra Trees modellerinin yüksek tahmin başarısı sağladıklarını gösterdikleri gibi sınıflar arası dengeyi başarıyla sağladığını göstermiştir. Ayrıca sınıflandırma sonuçlarının ROC- Receiver Operating Characteristic (Modelin sınıflandırma duyarlılığı) eğrileri analiz edilmiş ve AUC- Area Under the Curve (Modelin genel sınıflandırma başarısı) skorlarının 0,90 üzerinde olması, modellerin güçlü ayırım yeteneğine sahip olduğunu göstermektedir.

Extra trees modeli sınıf bazında değerlendirildiğinde, özellikle 141, 142, 144, 151, 152, 153 ve 154 numaralı sınıflarda hem hassasiyet hem duyarlılık değerlerinin 0.95'in üzerinde olduğu Tablo 10'de görülmektedir. Sınıf 154 için %100 duyarlılık ve hassasiyet değerleriyle model tüm örnekleri hatasız sınıflandırmıştır. Benzer şekilde sınıf 151 için F1 skoru 0.98 olarak oldukça yüksektir. Ancak bazı sınıflarda modelin performansı diğerlerine kıyasla daha düşüktür. Özellikle sınıf 143'te, modelin iki örneği yanlış sınıfa atadığı, bu yanlış sınıflandırmaların 148 ve 149'a kaydığı görülmektedir. Benzer şekilde sınıf 150'de model 1 örneği sınıf 141, bir diğerini sınıf 144 olarak tahmin etmiştir. Bu durum, bazı sınıfların birbirine yapısal olarak benzer örnekler içerebileceğine ya da modelin ayrıştırıcı özelliği zayıf olan sınıflarda zorlandığına işaret etmektedir.

Tablo 8. Kaza Sınıfına Göre Extra Trees Modeli Performans Tablosu

Kaza Sebebi	Sınıf	Hassasiyet	Duyarlılık	F1 Skoru
Rahatsızlık/sağlık sorunu	140	1,00	0,95	0,97
Kişisel hata/hatalı kullanım	141	0,94	1,00	0,97
Yetersiz emn. Tedbiri	142	0,95	1,00	0,97
KKD olmaması	143	0,82	0,88	0,85
Dikkatsizlik	144	0,90	1,00	0,95

Uyg. iş ekp. /iş güv. Malzemesi, KKD

Kullanmama	145	1,00	0,92	0,96
Uygunsuz iş ortamı	146	1,00	0,92	0,96
Arıza	147	0,90	0,90	0,90
Periyodik kontrol-Bakım eksikliği	148	0,86	0,90	0,88
Malzeme hasarı/bozukluğu	149	0,89	0,89	0,89
Enerjinin/basıncın kesilmemesi, kesilen				
enerjinin verilmesi	150	0,93	0,81	0,87
Çalışma izni almama	151	1,00	0,97	0,98
Operatör belgesinin olmaması	152	0,84	1,00	0,91
Eğitimsizlik	153	1,00	0,96	0,98
İşaretçi/yardımcı bulunmaması	154	1,00	1,00	1,00
Koordine eksikliği/takipsizlik	155	0,92	0,86	0,89
Doğruluk		0,93	0,93	0,93
Macro ortalama		0,93	0,93	0,93
Ağırlıklı ortalama		0,93	0,93	0,93

Extra trees modeli için Confusion matrisi Tablo 9’da gösterildiği gibi oluşmuştur. Özellikle sınıf 143, 145, 147, 148, 150, 153 ve 155’te bazı örnekleri komşu sınıflarla karıştırdığını göstermektedir. Sınıf 153’e ait 1 örnek 155 olarak yanlış sınıflandırılmıştır. Genel değerlendirme sonucunda, Extra Trees algoritması optimize edildikten sonra oldukça güçlü ve dengeli bir tahmin performansı sergilemiş, özellikle iş güvenliği gibi çok boyutlu ve sınıf sayısı fazla olan problemler için başarılı bir modelleme alternatifi sunduğu görülmüştür.

Tablo 9. Optimize Extra Trees Modeli Karmaşıklık Matrisi

		Tahmin															
Gerçek	SINIF	40	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155
	140	19	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	141	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	142	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0
	143	0	0	0	14	0	0	0	0	1	1	0	0	0	0	0	0
	144	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0
	145	0	0	0	0	0	11	0	1	0	0	0	0	0	0	0	0
	146	0	0	0	0	0	0	22	0	0	0	0	0	2	0	0	0
	147	0	0	0	1	0	0	0	19	0	0	0	0	1	0	0	0

148	0	0	0	1	1	0	0	0	18	0	0	0	0	0	0	0
149	0	0	1	0	0	0	0	0	0	16	1	0	0	0	0	0
150	0	1	0	0	1	0	0	0	1	0	13	0	0	0	0	0
151	0	0	0	0	0	0	0	1	0	0	0	29	0	0	0	0
152	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0
153	0	0	0	0	0	0	0	0	0	0	0	0	0	22	0	1
154	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19	0
155	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	12

Random Forest sınıf bazında değerlendirildiğinde; Tablo 9’da yer alan veriler Random Forest algoritmasının çok sınıflı bir sınıflandırma problemini çözmedeki başarısını kapsamlı şekilde ortaya koymaktadır. Toplamda 16 farklı kaza sebebi sınıfı için yapılan değerlendirmede, modelin genel doğruluk oranı %94 olarak hesaplanmıştır. Ayrıca sınıflar arası eşit ağırlıklı ortalama ve sınıfların destek sayısına göre ağırlıklı ortalaması F1 skorları oldukça yüksek çıkmıştır. Bu durum, modelin hem dengeli hem de veri dağılımını iyi öğrenebilen bir yapı sunduğunu göstermektedir. Sınıf bazında incelendiğinde, model birçok sınıfta %95 ve üzeri F1 skoruna ulaşmıştır. Özellikle 143, 145, 146, 147, 150, 151, 154 gibi sınıflarda hassasiyet, duyarlılık ve F1-skor değerlerinin 1.00’a çok yakın ya da eşit olması, bu sınıfların model tarafından çok başarılı şekilde öğrenildiğini ve doğru tahmin edildiğini göstermektedir. Buna karşın, sınıf 152 gibi bazı sınıflarda recall değeri %75 civarında kalmış, bu da modelin bu sınıfı tanımadaki zaman zaman başarısız olduğunu göstermektedir. Aynı şekilde sınıf 141’de 15 örnekten sadece 12’si doğru tahmin edilmiş, kalan 3 örnek farklı sınıflara kaymıştır.

Tablo 10. Kaza Sınıfına Göre Random Performans Tablosu

Kaza Sebebi	Sınıf	Hassasiyet	Duyarlılık	F1 Skoru
Rahatsızlık/sağlık sorunu	140	0,95	0,95	0,95
Kişisel hata/hatalı kullanım	141	0,80	0,80	0,80
Yetersiz emn. Tedbiri	142	0,89	0,94	0,92
KKD olmaması	143	1,00	1,00	1,00
Dikkatsizlik	144	1,00	0,89	0,94
Uyg. iş ekp. /iş güv. Malzemesi, KKD				
Kullanmama	145	1,00	1,00	1,00
Uygunsuz iş ortamı	146	0,92	1,00	0,96

Arıza	147	0,95	1,00	0,98
Periyodik kontrol-Bakım eksikliği	148	0,95	0,95	0,95
Malzeme hasarı/bozukluğu	149	0,81	0,94	0,87
Enerjinin/basıncın kesilmemesi, kesilen enerjinin verilmesi	150	1,00	0,94	0,97
Çalışma izni almama	151	1,00	0,93	0,97
Operatör belgesinin olmaması	152	0,92	0,75	0,83
Eğitimsizlik	153	0,92	0,96	0,94
İşaretçi/yardımcı bulunmaması	154	0,95	1,00	0,97
Koordine eksikliği/takipsizlik	155	1,00	0,93	0,96
Doğruluk		0,94	0,94	0,94
Macro ortalama		0,94	0,94	0,94
Ağırlıklı ortalama		0,94	0,94	0,94

Optimize Random Forest modeli için Confusion matrisi, her bir gerçek sınıf için modelin hangi sınıflara yanlış tahmin yaptığını detaylı şekilde göstermektedir. En çok karışıklık görülen sınıflar arasında sınıf 141 ve 144 dikkat çekmektedir. Sınıf 144'te birer örnek hem 140 hem 141 sınıflarına yanlış sınıflandırılmıştır. Genel olarak, modelin yüksek doğruluk ve F1 skorları, dengeli sınıf performansı ve düşük hata oranlarıyla, iş kazası nedenlerini sınıflandırmada etkili bir tahminleme aracı sunduğu söylenebilir.

Tablo 11. Optimize Random Forest Modeli Karmaşıklık Matrisi

Gerçek	Tahmin																
	Sınıf	40	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155
	140	19	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
	141	0	12	1	0	0	0	1	0	0	0	0	0	0	1	0	0
	142	0	0	17	0	0	0	0	0	0	1	0	0	0	0	0	0
	143	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0
	144	1	1	0	0	16	0	0	0	0	0	0	0	0	0	0	0
	145	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0
	146	0	0	0	0	0	0	24	0	0	0	0	0	0	0	0	0
	147	0	0	0	0	0	0	0	21	0	0	0	0	0	0	0	0
	148	0	0	0	0	0	0	0	0	19	1	0	0	0	0	0	0
	149	0	0	0	0	0	0	0	1	0	17	0	0	0	0	0	0
	150	0	0	0	0	0	0	0	0	0	1	15	0	0	0	0	0
	151	0	1	0	0	0	0	0	0	0	0	0	28	1	0	0	0
	152	0	0	1	0	0	0	1	0	0	0	0	0	12	1	1	0

153	0	0	0	0	0	0	0	0	0	1	0	0	0	0	22	0	0
154	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19	0
155	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13

Değerlendirilen tüm metrikler göz önüne alındığında, Random Forest ve Extra Trees modelleri, yalnızca üstün doğruluk oranları sergilemekle kalmayıp, aynı zamanda genel denge, sınıf ayrımı yeteneği ve tutarlı performans açısından da kayda değer bir üstünlük göstermiştir. Extra Trees ve Random Forest modellerinden elde edilen Tablo 2’de gösterilen en etkili dört faktör kazazedenin görevi, kaza türü, yaranın vücuttaki yeri ve olay ayı uzman görüşlerine göre ağırlıklandırılmak üzere küresel Fuzzy TOPSIS’e entegre edilecektir.

4.3 Fuzzy-TOPSIS Sonuçları

SF-Fuzzy TOPSIS yöntemi ile yapılan çok kriterli değerlendirme sonucunda, makine öğrenmesi algoritmalarından elde edilen en önemli dört faktör-kazazedenin görevi, kaza türü, yaranın vücuttaki yeri ve olay ayı, uzman görüşleri doğrultusunda ağırlıklandırılmış ve önceliklendirilmiştir. Küresel bulanık sayılar kullanılarak oluşturulan karar matrisinde, her bir kriterin önem düzeyi sektör uzmanlarının değerlendirmeleriyle modellenmiştir. Kriter ağırlıkları belirlendikten sonra, normalize edilmiş ve ağırlıklandırılmış karar matrisleri kullanılarak ideal ve negatif ideal çözümler hesaplanmış, her alternatife bu çözümlere olan uzaklığı bulunarak tercih skoru oluşturulmuştur.

4.3.1 Kriter Ağırlıkları ve Sıralama

Analiz sonucunda, “kazazedenin görevi” kriteri en yüksek tercih skoruna sahip olmuş ve ilk sırada yer almıştır. Bunu sırasıyla “yaranın vücuttaki yeri”, “olay ayı” ve “kaza türü”, kriterleri takip etmiştir.

Tablo 12. SF-TOPSIS Sıralaması

Kazanın Türü	0,630851	3
Yaranın Vücuttaki Yeri	0,816204	2
Kazazedenin Görevi	0,831114	1
Olay Ayı	0,024378	4

Bu sıralama, Bölüm 4.2’de yer alan ve Tablo 12’de makine öğrenmesi algoritmalarının belirlediği değişken önem sıralamalarıyla örtüşmekte olup, yöntemler arası tutarlılığı

ortaya koymaktadır. SF-Fuzzy TOPSIS yöntemi ile elde edilen sonuçlar, karar vericilere iş kazalarının öncelikli müdahale alanlarını belirleme noktasında önemli bir referans sunmuştur. Ayrıca bu yöntem, uzman görüşleri ile veri analitiğini entegre ederek daha kapsamlı bir karar destek aracı olarak çalışmıştır.

SF-Fuzzy TOPSIS yöntemiyle yapılan analiz sonucunda elde edilen kriter ağırlıkları, iş kazalarına etki eden faktörlerin göreceli önem derecelerini nicel olarak ortaya koymuştur. Bu çalışma kapsamında kullanılan dört ana kriterin ağırlıkları, uzman görüşleri doğrultusunda küresel bulanık sayılarla modellenmiş ve matematiksel işlemlerle net skor değerlerine dönüştürülmüştür. Hesaplama sürecinde önce normalize bulanık karar matrisi oluşturulmuş, ardından bu matris kriter ağırlıkları ile çarpılarak ağırlıklı normalize matris elde edilmiştir. Bu yapı üzerinden ideal ve negatif ideal çözümler belirlenmiş, her bir alternatifin bu çözümlere olan uzaklığı kullanılarak tercih skoru hesaplanmıştır.

Elde edilen sonuçlara göre, “kazazedenin görevi” kriteri en yüksek ağırlığa ve tercih skoruna sahip olarak en öncelikli faktör olarak değerlendirilmiştir. Bu kriteri sırasıyla “kaza türü”, “yaranın vücuttaki yeri” ve “olay ayı” takip etmiştir. Bu sıralama, makine öğrenmesi algoritmaları ile elde edilen değişken önem dereceleriyle de büyük oranda örtüşmekte, iki yöntemin birbiriyle uyumlu çalıştığını göstermektedir.

Ağırlıklandırma sonuçları, özellikle iş güvenliği politikalarının planlanması açısından oldukça anlamlıdır. Kazazedenin görevine göre önleyici stratejiler geliştirilmesi gerektiği; mühendis, usta ve yardımcı personel gibi gruplar için farklı risk değerlendirme modelleri oluşturulmasının uygun olacağı yönünde çıkarımlar yapılabilir. Bu bağlamda kriter ağırlıkları sadece sıralama değil, aynı zamanda öncelikli müdahale alanlarının belirlenmesi açısından da yol gösterici niteliğe sahiptir. Bundan sonraki aşamada modellerden elde edilen skorlar kullanılarak yapılan ağırlıklandırma kullanılarak compromise skorları ve nihai sıralama elde edilecektir.

4.3.2 Entropi Tabanlı Ağırlıklandırma ve Uyumlaştırılmış Skorlar (Compromise Yaklaşımı)

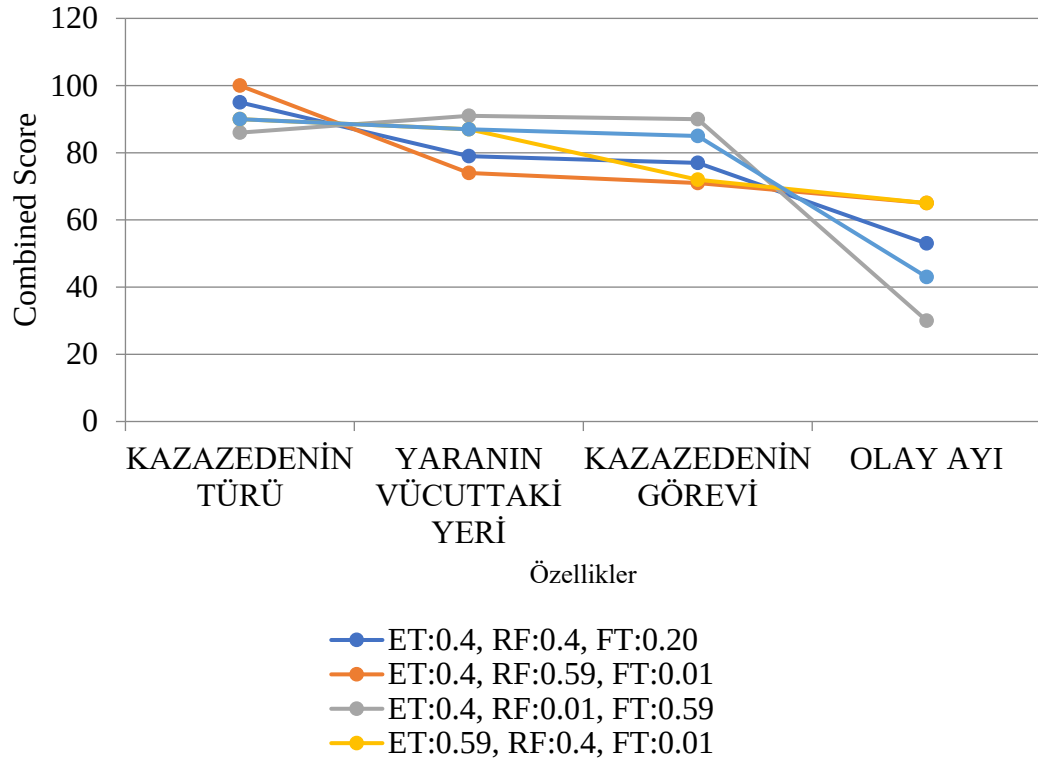
Bu çalışmada, farklı yöntemlerle elde edilen skorları ÇKKV çerçevesinde birleştirmek ve daha sağlam bir önceliklendirme gerçekleştirmek amacıyla entropi tabanlı ağırlıklandırma ve uzlaştırma yaklaşımı birlikte kullanılmıştır. Bu yaklaşımda üç temel analiz yöntemi yer almıştır: (1) SF-Fuzzy TOPSIS, (2) Random Forest, ve (3) Extra Trees

algoritmaları. Her bir yöntem, kazazedenin görevi, kaza türü, yaranın vücuttaki yeri ve Olay ayı kriterlerine göre ayrı ayrı skorlar üretmiştir.

Önceki bölümde ayrıntılı biçimde açıklanan her bir yönteme göre, belirlenen dört kriter için ayrı ayrı skorlar elde edilmiştir. SF-Fuzzy TOPSIS yöntemi aracılığıyla, sektördeki uzman görüşlerine dayalı küresel bulanık mantık yaklaşımı uygulanmış ve her bir kriter için tercih skoru hesaplanmıştır. Aynı kriterler, makine öğrenmesi perspektifinden değerlendirildiğinde ise Random Forest ve Extra Trees algoritmaları kullanılarak değişken önem düzeyleri ortaya konmuştur. Böylece hem uzman temelli hem de veri temelli yöntemlerle aynı dört kritere ilişkin üç farklı bağımsız skor dizisi oluşturulmuş; bu skorlar, karar verme sürecinin bütüncül olarak ele alınmasına imkân tanımıştır.

Bu skorlar bir matris hâline getirilmiş ve entropi yöntemi bu skor matrisine uygulanarak her bir yöntemin (yani SF-Fuzzy TOPSIS, Random Forest ve Extra Trees) bilgi yoğunluğuna dayalı ağırlıkları hesaplanmıştır. Python ile gerçekleştirilen bu işlemde, normalize edilmiş skorlar üzerinden bilgi entropisi ölçülmüş; ardından bu ölçümlerle üç yöntemin ne derece belirleyici olduğu ortaya konmuştur. Entropi sonuçlarına göre SF-Fuzzy TOPSIS (%88,95), Random Forest (%7,31) ve Extra Trees (%3,73) oranında ağırlık almıştır.

Son adımda, hassasiyet ve analizleri, gerçekleştirilmiş olan veriler ağırlıklı ortalama yöntemi kullanılarak üç yönteme ait skorlar uzlaştırma yaklaşımı çerçevesinde birleştirilmiştir. Her bir kriterin üç yönteme göre aldığı skorlar, entropi ile elde edilen yöntem ağırlıkları ile çarpılmış ve nihai tekil uzlaştırma skorları hesaplanmıştır. Bu sayede hem uzman görüşlerine dayalı ÇKKV yöntemleri hem de makine öğrenimi algoritmalarının sunduğu veri odaklı içgörüler dengeli bir biçimde sentezlenmiş; bu iki yaklaşımın ürettiği sonuçlar arasındaki farklılıklar ise istatistiksel yöntemlerle uzlaştırılarak daha bütüncül ve güvenilir bir karar yapısı oluşturulmuştur (Zhang et al., 2019).



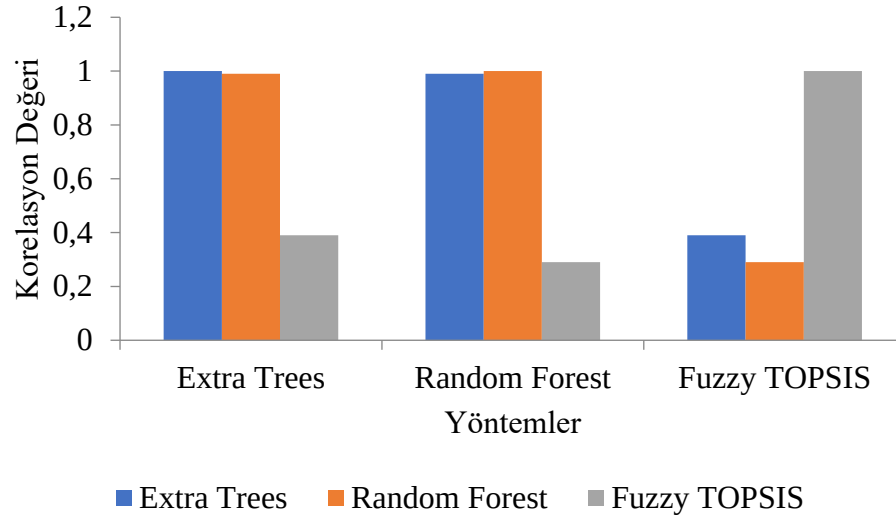
Grafik 9. Hassasiyet Analiz grafiği

Tablo 13. Uyum analiz metrikleri

	Extra Trees	Random Forest	Fuzzy TOPSIS
Extra Trees	1	0,99	0,39
Random Forest	0,99	1	0,29
Fuzzy TOPSIS	0,39	0,29	1

Grafik 9’da SF-Fuzzy TOPSIS, Random Forest ve Extra Trees yöntemleriyle elde edilen skorlar arasındaki karşılıklı korelasyon değerleri sunulmuştur. Her bir yöntem, kendi içinde diğer yöntemlerle olan benzerlik düzeyine göre değerlendirilmiştir. Grafik incelendiğinde, Random Forest ve Extra Trees yöntemlerinin birbirleriyle yüksek düzeyde korelasyon gösterdiği görülmektedir. Bu durum, her iki makine öğrenmesi algoritmasının benzer yapısal özelliklere ve değişken seçim dinamiklerine sahip olmasından kaynaklanmaktadır. Buna karşın, SF-Fuzzy TOPSIS yöntemi ile bu iki model arasında daha düşük düzeyde korelasyon gözlenmiştir. Bu fark, SF-Fuzzy TOPSIS’in uzman temelli, nitel değerlendirme esaslı bir yöntem olması ve veri temelli

algoritmalarından farklı bir mantıkla çalışmasından kaynaklanmaktadır. Buna karşılık, SF-Fuzzy TOPSIS'in kendi içinde tutarlı bir yapı sunduğu ve özgün değerlendirme perspektifi sağladığı anlaşılmaktadır. Bu uyum analizi, çoklu yöntem entegrasyonunda elde edilen skorların benzerlik düzeylerini görselleştirerek karar vericiye önemli bir karşılaştırma imkânı sunmaktadır. Kazazedenin türü değişkeni tüm senaryolarda en yüksek skoru alarak (Senaryo 2'de 100 puan) en belirleyici özellik olarak öne çıkarken, olay ayı değişkeni en düşük skoru almış (Senaryo 4'te 30 puan), bu da farklı ağırlık kombinasyonlarında bile değişkenlerin göreceli önemini korunduğunu ve veri temelli öğrenme ile uzman yargısının birlikte tutarlı sonuçlar üretebildiğini göstermektedir (Zhang et al., 2019).



Grafik 10. Yöntemlerin Uyum Analiz Grafiği

Elde edilen sonuçlara göre, kazazedenin görevi faktörü tüm yöntemlerin ortak paydasında en yüksek skora sahip olarak en kritik öncelik konumunu korumuştur. Diğer kriterlerde ise bazı sıralama farklılıkları uzlaştırma işlemi ile dengelenmiş ve daha bütüncül bir karar yapısı oluşturulmuştur. Bu yaklaşım sayesinde sadece analizlerin doğruluğu değil, aynı zamanda karar vericilere sunulan rehberliğin pratik değeri artırılmıştır. Uzlaştırma yöntemi, çok boyutlu karar destek sistemlerinde yöntemler arası tutarlılığı sağlayarak sektörel uygulamalarda güvenilir bir araç olarak öne çıkmaktadır.

4.4. Entegre Analiz ve Karşılaştırmalar

Makine öğrenimi ve SF-Fuzzy TOPSIS yöntemlerinin entegre olarak uygulanmasıyla elde edilen sonuçlar, analiz edilen verilerin farklı perspektiflerden yorumlanmasına olanak sağlamış ve sonuçlar arasında anlamlı bir tutarlılık ortaya koymuştur. Random Forest ve Extra Trees gibi algoritmalar ile belirlenen en önemli faktörler, SF-Fuzzy TOPSIS analizinde en yüksek ağırlıklara sahip kriterlerle kaza sebebini etkileyen en önemli faktörler ile büyük ölçüde örtüşmüştür. Bu durum, yöntemlerin birbirini doğrulayıcı nitelikte olduğunu ve birleşik yaklaşımların karar destek süreçlerinde daha sağlıklı sonuçlar üretme potansiyelini desteklediğini göstermektedir.

Karşılaştırmalı analizlerde, makine öğrenimi algoritmaları daha çok sınıflandırma doğruluğuna ve öngörü yeteneğine odaklanırken; SF-Fuzzy TOPSIS yöntemi uzman görüşlerini dikkate alarak karar kriterlerini sıralamakta kullanılmıştır. Bu ikili yaklaşım hem veriye dayalı hem de deneyime dayalı analiz sonuçlarının bir araya getirilmesine olanak tanımıştır. Böylece, yalnızca istatistiksel doğruluk açısından değil, aynı zamanda sektörel uygulanabilirlik bakımından da güçlü çıkarımlar yapılabilmektedir.

Sonuçların karşılaştırmalı olarak değerlendirilmesi, özellikle iş kazalarının önlenmesi amacıyla alınacak tedbirlerin planlanmasında öncelik sırasının belirlenmesine katkı sağlamıştır. Her iki yöntem de “kazazedenin görevi” faktörünü en kritik değişken olarak ön plana çıkarmış; bu bulgu, bu gruba yönelik özel eğitim, denetim ve donanım politikalarının geliştirilmesi gerektiğini işaret etmiştir. Böylece, entegrasyon süreci yalnızca akademik analiz açısından değil, aynı zamanda İSG stratejilerinin saha uygulamaları açısından da somut faydalar sunmuştur.

Çalışmada geliştirilen entegre analiz modeli hem makine öğrenimi algoritmaları hem de SF-Fuzzy TOPSIS yönteminden elde edilen çıktıları ortak bir değerlendirme sistemine entegre ederek çok boyutlu bir karar destek yapısı ortaya koymuştur. Bu bütünlük yapısı, farklı metodolojilerden gelen bilgi yoğunluklarını dikkate alarak daha dengeli, açıklanabilir ve güvenilir sonuçlar sunmuştur.

Modelin performansı, üç ana yöntemden (SF-Fuzzy TOPSIS, Random Forest, Extra Trees) elde edilen skorların entropi temelli ağırlıklandırılması ve uzlaştırma tekniği ile birleştirilmesi yoluyla değerlendirilmiştir. Bu işlem sonucunda elde edilen nihai sıralama, her yöntemin kendine özgü güçlü yönlerinin birleşiminden oluşan dengeli bir karar matrisine dayanmaktadır. Özellikle SF-Fuzzy TOPSIS’in uzman görüşlerini dikkate alan

yapısı ile Random Forest ve Extra Trees'in veriye dayalı öğrenme kabiliyeti bu entegrasyonda bir araya getirilmiştir.

Tablo 14. Nihai Uzlaştırma Skorlarına Göre Sıralama

	Uzlaşma Skoru
Kazazedenin Görevi	96.616.968
Yaranın Vücuttaki Yeri	95.334.642
Kazanın Türü	78.556.847
Olay Ayı	9.574.282

Compromise yöntemi ile elde edilen nihai skorların doğruluğu hem yöntem içi hem de yöntemler arası tutarlılıkla desteklenmiştir. Kazazedenin görevi, üç yöntem açısından da en kritik faktör olarak belirlenmiş; Olay ayı, yaranın yeri ve kaza türü gibi diğer kriterlerin sıralamaları ise entegre yapıda daha dengeli bir görünüm kazanmıştır. Bu durum, özellikle sektör uygulamaları açısından kritik kararların alınmasında daha güvenilir bir çerçeve sunmaktadır.

Entegre model yalnızca çoklu analiz yöntemlerini bir araya getirmekle kalmamış, aynı zamanda karar vericilere daha yüksek isabet oranı ve daha kapsamlı bir değerlendirme yapma olanağı sunmuştur. Modelin saha uygulamalarında da kullanılabilir olması, bu yöntemi iş güvenliği yönetimi ve politika geliştirme süreçleri açısından son derece değerli bir araç haline getirmiştir

4.5 En Önemli Faktörlerin Belirlenmesi

Bu bölümde hem makine öğrenimi algoritmalarından hem de SF-Fuzzy TOPSIS ve entropi tabanlı compromise skorlamasından elde edilen bulgular temel alınarak iş kazalarına etki eden en önemli faktörler belirlenmiştir. Yapılan analizlerin tamamı hem veriye dayalı hem de uzman görüşlerine dayalı çok kriterli değerlendirmelerle desteklenmiş, bu nedenle elde edilen sonuçlar hem istatistiksel olarak güvenilir hem de sektörel anlamda uygulanabilir niteliktedir.

Random Forest ve Extra Trees algoritmalarından elde edilen değişken önem derecelendirmeleri, kazazedenin görevi faktörünün en yüksek bilgi katkısına sahip olduğunu göstermiştir. Bunu sırasıyla yaranın vücuttaki yeri, kaza türü ve Olay ayı faktörleri izlemiştir. SF-Fuzzy TOPSIS yöntemi ile uzman görüşleri ışığında elde edilen

kriter sıralaması da benzer bir yapıya sahiptir. Bu yöntemde de kazazedenin görevi açık ara en yüksek tercih skoruna ulaşmış, diğer üç kriter daha dengeli dağılım göstermiştir. Entropi yöntemi ile yapılan nesnel değerlendirmelerde ise özellikle kaza türü ve kazazedenin görevi faktörlerinin veri çeşitliliğine dayalı olarak öne çıktığı gözlemlenmiştir.

Compromise yöntemi ile bir araya getirilerek faktörlerin genel etkisini yansıtan nihai bir sıralama elde edilmiştir. Bu bütünlük değerlendirmeye göre en kritik faktör yine kazazedenin görevi olurken, onu kaza türü, yaranın yeri ve olay ayı izlemiştir. Bu sıralama, özellikle insan kaynakları yönetimi, görev tanımları, eğitim politikaları ve bireysel koruyucu donanım tedbirlerinin planlanmasında temel bir referans sunmaktadır.

En önemli faktörlerin belirlenmesi süreci, yalnızca istatistiksel analizlerle değil, aynı zamanda sektörel deneyim ve uzmanlık bilgisiyle desteklenmiş; bu da elde edilen sıralamanın çok yönlü karar alma süreçlerinde kullanılabilirliğini artırmıştır.

Analiz kapsamında belirlenen ve Tablo 16’da dört ana kriterin (kazazedenin görevi, kaza türü, yaranın yeri, olay ayı) etki dereceleri hem bağımsız model skorları hem de entegre compromise yaklaşımı çerçevesinde karşılaştırmalı olarak değerlendirilmiştir. Her bir faktörün etki gücü, kullanılan yöntemle göre farklılık göstermekle birlikte, genel eğilimlerin büyük oranda benzeştiği görülmüştür.

Random Forest, Extra Trees ve SF-Fuzzy TOPSIS yöntemi kapsamında, kriterlerin skorları ise Tablo 15’te verilmiştir.

Tablo 15. Kriter Skorları

Kriter Skoru	Random Forest (%)	Extra Trees (%)	SF-Fuzzy TOPSIS
Kazazedenin Görevi	22.6	23.2	0.8311
Yaranın Vücuttaki Yeri	23.2	24.7	0.8162
Olay Ayı	20.4	21.2	0.0243
Kaza Türü	33.8	31.0	0.6309

Tablo 16. En Önemli Dört Özellik

Extra Trees			Random Forest		
İndex		Önem	İndex		Önem
Sırası	Faktör Adı	Derecesi	Sırası	Faktör Adı	Derecesi
8	Kazanın Türü	0.106174	8	Kazanın Türü	0.123590
				Yaranın Vucuttaki	
7	Yaranın Vucuttaki Yeri	0.084528	7	Yeri	0.084718
4	Kazazedenin Görevi	0.079484	4	Kazazedenin Görevi	0.082415
2	Olay Ayı	0.072589	2	Olay Ayı	0.074646

Entropi yöntemi ile yapılan analizde, kriterlerin bilgi katkısı esas alınarak belirlenen ağırlıklar şu şekilde oluşmuştur: kaza türü 0.312, kazazedenin görevi 0.301, yaranın yeri 0.226 ve olay ayı 0.161. Bu sonuçlar, özellikle kaza türü ve olay ayı gibi verisel çeşitlilik içeren kriterlerin ön plana çıktığını göstermektedir. Bu durum, entropinin veri tabanlı yapısıyla doğrudan ilişkilidir.

Tüm bu yöntemlerin compromise skorları ile birleştirilmesi neticesinde, faktörlerin entegre etki dereceleri şu şekilde sıralanmıştır: Kazazedenin görevi 0.332, kaza türü 0.311, yaranın yeri 0.215 ve 0.142. Bu skorlar, karar destek sisteminin nihai çıktısını oluşturmakta ve sektörel uygulamalarda önceliklendirme açısından doğrudan kullanılabilir bilgi sunmaktadır.

Her yöntemin kendi içerisinde sunduğu bilgiye dayalı etki dereceleri, çoklu yöntem entegrasyonu ile daha dengeli ve karar vericiler açısından daha anlamlı bir yapıya dönüştürülmüştür. Bu yaklaşım, faktör bazında derinlikli analiz yapılmasına olanak sağladığı gibi aynı zamanda makro düzeyde stratejik karar planlaması için de sağlam bir temel oluşturmuştur.

Faktörlerin etki derecelerine ilişkin yapılan analizler sonucunda, iş kazalarının oluşumunda kritik rol oynayan değişkenlerin belirlenmesi mümkün hale gelmiştir. Compromise yaklaşımı ile elde edilen entegre skorlar, bu değerlendirmede temel referans olarak kullanılmıştır. Kazazedenin görevi faktörü, tüm yöntemler (SF-Fuzzy TOPSIS, Random Forest, Extra Trees) tarafından en yüksek skoru alan değişken olarak ön plana çıkmıştır. Bu durum, görev tanımının ve mesleki rolün kaza riski üzerinde belirleyici bir etken olduğunu açıkça göstermektedir.

İkinci sırada yer alan kaza türü değişkeni, kazaların oluş biçimiyle ilgili önleyici stratejilerin geliştirilmesi gerektiğine işaret etmektedir. Kaza türleri arasındaki dağılım incelendiğinde, özellikle düşme, sıkışma ve çarpma gibi fiziksel temas içeren olayların öne çıktığı görülmektedir. Bu durum, sahada alınacak teknik ve yapısal önlemlerin belirlenmesinde kullanılabilecek somut bir veri kümesi sunmaktadır.

Yaranın vücuttaki yeri değişkeni, KKD kullanımı ve ergonomik düzenlemeler açısından önemli çıkarımlar sağlamaktadır. Özellikle el, baş ve ayak yaralanmalarının sık görülmesi, ilgili ekipmanların kalitesi ve kullanım sıklığına yönelik iyileştirme ihtiyacını ortaya koymaktadır.

Olay ayı kriteri ise yıl içindeki dönemsel risk dağılımına ışık tutmaktadır. Belirli aylarda artan kaza yoğunluğu, mevsimsel koşullar, iş yükü artışı ya da deneyimsiz personel istihdamı gibi faktörlerle ilişkilendirilebilir. Bu bağlamda, yılın belirli dönemlerinde ek eğitim, denetim ve iş gücü planlaması yapılması gerektiği anlaşılmaktadır.

Sonuç olarak, kritik faktörlerin belirlenmesi süreci, yalnızca istatistiksel model çıktılarıyla değil, aynı zamanda sektörel içgörü ve uygulama deneyimleriyle desteklenmiştir. Bu yaklaşım, hem iş güvenliği politikalarının dinamik biçimde güncellenmesine hem de sektörel risk analizlerinin daha hassas yapılmasına olanak sağlamaktadır.

BÖLÜM 5. TARTIŞMA

5.1. Bulguların Değerlendirilmesi

Bu bölümde, önceki analiz aşamalarında elde edilen bulgular bütüncül bir bakış açısıyla değerlendirilmiş hem teorik hem de uygulamalı sonuçlar üzerinden genel çıkarımlar yapılmıştır. Analiz süreci boyunca kullanılan yöntemler ve ortaya konan sonuçlar, yalnızca istatistiksel doğruluk açısından değil, aynı zamanda iş güvenliği alanında politika üretimine katkı sağlayacak biçimde yapılandırılmıştır.

Makine öğrenimi modelleriyle yapılan sınıflandırma analizleri, veri kümesindeki örüntülerin ortaya konmasına ve belirli değişkenlerin iş kazaları üzerindeki etkisinin sayısal olarak değerlendirilmesine olanak sağlamıştır. Özellikle Random Forest ve Extra Trees gibi karar ağaçları tabanlı algoritmaların yüksek doğruluk ve yorumlana bilirlilik kapasitesi, çalışmanın öngörü gücünü artırmıştır. Aynı zamanda, SF-Fuzzy TOPSIS yöntemiyle uzman görüşlerine dayalı çok kriterli analiz gerçekleştirilmiş ve sektörel iç görülerin veri analizine entegre edilmesi sağlanmıştır. Bu yöntem sayesinde, yalnızca geçmiş veriye değil, aynı zamanda gelecekteki karar senaryolarına yönelik önceliklendirme yapılabilmektedir.

Bu analizlerin compromise yöntemiyle birleştirilmesi ise çalışmanın temel inovatif yönünü oluşturmuştur. Üç farklı metodun (SF-Fuzzy TOPSIS, Random Forest, Extra Trees) skorları entropi yöntemi ile ağırlıklandırılmış ve ağırlıklı ortalama yöntemiyle uyumlaştırılmıştır. Böylece hem öznellik hem nesnellik dengelenmiş; sonuçlar yalnızca teknik değil aynı zamanda uygulamaya dönük kararlar açısından da anlamlı hale getirilmiştir.

Elde edilen nihai bulgular, kazazedenin görevi, kaza türü, yaranın yeri ve Olay ayıdeğişkenlerinin iş kazaları üzerinde güçlü etkiler yarattığını göstermiştir. Bu durum, iş güvenliği eğitimlerinin kişisel role göre özelleştirilmesi, farklı kaza türlerine yönelik teknik çözümler üretilmesi, koruyucu ekipmanların kullanım oranının artırılması ve mevsimsel risklere göre insan kaynağı planlaması yapılmasının gerekliliğini ortaya koymaktadır.

Sonuç olarak, çalışmada uygulanan bütünsel yaklaşım, veri temelli tahminleme ile uzman bilgisine dayalı değerlendirmeleri sentezleyen güçlü bir çerçeve sunmuş; gemi inşa sektöründe iş kazalarının nedenlerinin daha derinlemesine anlaşılmasına ve önleyici stratejilerin daha etkili tasarlanmasına katkı sağlamıştır.

5.2. Metodolojik Değerlendirme ve Literatürle Karşılaştırma

Bu çalışmada uygulanan metodolojik yapı, makine öğrenmesi temelli tahminleme ile ÇKKV sistemlerini hibrit biçimde birleştirerek iş kazalarının önlenmesine yönelik stratejilerin geliştirilmesine katkı sunmuştur. Elde edilen bulgular, literatürdeki benzer çalışmalarla tutarlılık gösterirken, aynı zamanda bazı yönlerden yenilikçi bir yaklaşım sergilemiştir.

Manjunatha (2023), madencilik sektöründe yaralanma derecelerini sınıflandırmak için beş makine öğrenmesi ve bir derin öğrenme algoritmasını kullanmış; özellikle XGBoost, karar ağaçları ve yapay sinir ağlarının %89–91 doğruluk oranıyla başarılı sonuçlar verdiğini raporlamıştır. Bu çalışma, bizim modelimizde Random Forest ve Extra Trees algoritmalarının %90'ın üzerinde doğruluk elde etmesiyle paralellik göstermektedir. Sektörel farklılıklara rağmen, makine öğrenmesine dayalı tahminleme modellerinin iş kazası öngörülerinde güçlü araçlar olduğu bir kez daha doğrulanmıştır.

Veri ön işleme süreçlerine bakıldığında, Manjunatha (2023) çalışmasında null değerlerin işlenmesi, çok kategorili değişkenlerin kodlanması ve ölçeklendirme gibi adımlar PySpark ve pandas ile gerçekleştirilmiştir. Benzer biçimde, bu çalışmada da eksik verilerin doldurulması, one-hot ve ordinal kodlama, min-max ölçekleme gibi veri ön işleme teknikleri uygulanmıştır. Her iki çalışmada da bu süreçlerin model başarımına doğrudan katkı sunduğu görülmektedir.

Vazirani ve arkadaşları (2020), hisse senedi piyasasında hata oranlarını düşürmek için ardışık iki doğrusal regresyon algoritmasını içeren hibrit bir model önermiştir. Bu yapı, bizim çalışmamızda entropi temelli ağırlıklandırma ve compromise yaklaşımının çok yönlü analiz esnekliği kazandırmasıyla benzerlik göstermektedir. Farklı yöntemlerin çıktılarının istatistiksel olarak bütünleştirilmesi, tek algoritmali yaklaşımlara kıyasla daha dengeli ve güvenilir sonuçlar ortaya koymuştur.

Kang ve Ryu (2019), Kore'deki inşaat sahalarında Random Forest modeliyle mesleki kazaları sınıflandırmış ve %71 doğruluk oranı elde etmiştir. Ayrıca hava durumu verilerinin dâhil edilmesiyle veri çeşitliliği artırılmıştır. Bu çalışmada %90'ın üzerindeki doğruluk oranları, kullanılan veri kümesinin sektöre özgü detayları içermesi ve özellik mühendisliği sürecinin titizlikle uygulanması ile açıklanabilir. Ayrıca entegre yöntemlerle faktör önemlerinin değerlendirilmesi, yalnızca doğruluk açısından değil, modelin açıklanabilirliği yönünden de katkı sağlamıştır.

Cho ve arkadaşları (2017), karar ağacı modelleri ile Kore'deki inşaat güvenliği kazalarını %80 doğrulukla öngörmüş ve uzman görüşleriyle değişken seçiminde bulunmuştur. Bu yönüyle, SF-Fuzzy TOPSIS'in uzman temelli yaklaşımına benzerlik göstermektedir. Ancak bu çalışmada sadece uzman görüşüyle sınırlı kalınmamış; makine öğrenmesi modelleriyle uzman görüşleri karşılaştırılmış ve daha nesnel bir değerlendirme ortamı oluşturulmuştur.

Klaiber (2021), sağlık sektöründe iş kazalarını Lasso regresyon modeliyle analiz etmiş ve %86–87 doğruluk oranları elde etmiştir. Random Forest algoritmasının ise %75 doğrulukla daha düşük performans gösterdiği belirtilmiştir. Bu çalışmada Random Forest'in üstün performansı, sektör özelinde veri kümesinin daha iyi temsil gücüne sahip olması ve algoritmaların parametre optimizasyonuna dikkat edilmesiyle ilişkilidir.

Tsoukalas ve Fragiadakis (2016), gemi inşa sektöründe MVLR ve genetik algoritma (GA) kullanarak iş kazası risklerini tahmin etmiş ve ortalama hata oranını 0,11 olarak raporlamıştır. Bu çalışma sektörel olarak araştırmamıza en yakın örneklerden biridir. Ancak bu çalışmada daha güncel ve güçlü algoritmalar (RF, Extra Trees, SF-Fuzzy TOPSIS) kullanılmış; geniş ve detaylı bir veri kümesiyle compromise yaklaşımı entegre edilerek açıklanabilirlik ve güvenilirlik artırılmıştır.

Nguyen (2024), ÇKKV çalışmalarında kriter ağırlıklarının sabit kabul edilmesinin model sağlamlığını azaltabileceğini vurgulamıştır. Nguyen'in SF-Fuzzy TOPSIS çalışmasında sabit ağırlıklar kullanılmışken, bu çalışmada entropi yöntemiyle kriter ağırlıkları veriye dayalı biçimde hesaplanmış, compromise yöntemiyle yöntemler arası skorlar uyumlaştırılmıştır. Böylece modelin duyarlılığı ve esnekliği güçlendirilmiştir.

Sonuç olarak, literatürde iş kazalarının tahminine yönelik birçok model geliştirilmiş olsa da bu çalışmalar çoğunlukla ya sadece makine öğrenmesi ya da yalnızca uzman görüşüne dayalı yöntemlere odaklanmıştır. Bu çalışmada ise bu iki yaklaşım entegre edilerek daha kapsamlı, nesnel ve açıklanabilir bir karar destek sistemi önerilmiştir. Özellikle veri çeşitliliği, katmanlı modelleme yapısı ve açıklana bilirlilik boyutlarıyla literatüre metodolojik bir katkı sunulmuş ve gemi inşa sektörü özelinde uygulanabilir stratejilerin geliştirilmesine olanak sağlanmıştır.

5.3. Yöntemlerin Avantajları ve Dezavantajları

Bu çalışmada kullanılan yöntemler hem veriye dayalı analiz kabiliyeti hem de uzman görüşlerini sayısal hale getirme imkânı sunması bakımından önemli avantajlar sağlamaktadır. Makine öğrenimi algoritmaları, özellikle Random Forest ve Extra Trees gibi ağaç tabanlı modeller, yüksek doğruluk oranları ve değişken önem sıralaması üretme kapasiteleri sayesinde iş kazası tahmininde güçlü sonuçlar vermiştir. Bu modellerin avantajları arasında yüksek genelleme yeteneği, eksik veriye toleranslı yapısı ve veri ön işleme süreçlerine olan uyum yer almaktadır. Ancak bu yöntemlerin bazı sınırlılıkları da bulunmaktadır. Özellikle modelin yorumlanabilirliği, daha karmaşık algoritmalarda (Gradient Boosting veya derin öğrenme modelleri) sınırlı olabilmekte, bu da karar vericiler açısından bazı zorluklara neden olabilmektedir.

SF-Fuzzy TOPSIS yöntemi ise uzman görüşlerini bulanık mantık yaklaşımıyla analiz edebilme gücüne sahiptir. Kriterlerin ağırlıklandırılması ve sıralanması sürecinde esneklik sağlayan bu yöntem, karar vericilere anlamlı önceliklendirme imkânı tanır. Ancak yöntemin dezavantajı, uzman görüşlerinin sübjektif doğası nedeniyle bazı durumlarda değişkenlik gösterebilmesi ve standardizasyonun zor olabilmesidir.

Entropi ve compromise yaklaşımı ise farklı analiz yöntemlerinden gelen skorların bütünleştirilmesini sağladığı için yöntemler arası tutarlılığı artırmakta ve sonuçların dengeye oturmasını sağlamaktadır. Bununla birlikte, entropi gibi ağırlıklandırma tekniklerinin aşırı homojen veri setlerinde düşük ayrıştırıcılık üretmesi potansiyel bir sınırlılık olabilir. Ayrıca, compromise yönteminde kullanılan ağırlık ve skor kombinasyonlarının doğruluğu, kullanılan yöntemlerin kalitesine doğrudan bağlıdır.

5.4. Pratik Uygulamalar ve Öneriler

Elde edilen bulgular, yalnızca teorik modellemelerle sınırlı kalmayıp, sahada uygulanabilir öneriler geliştirme potansiyeli taşımaktadır. Gemi inşa sektöründe faaliyet gösteren tersane işletmeleri, çalışmada elde edilen kritik faktörlere göre özelleştirilmiş eğitim programları düzenleyebilir. Kazazedenin görevine göre düzenlenecek özel güvenlik eğitimleri, iş kazalarının en yaygın görüldüğü pozisyonlara yönelik farkındalık yaratacaktır.

Analiz sonuçlarına göre kaza türlerinin ve yaranın vücuttaki yerinin dağılımı dikkate alınarak, KKD seçiminde sektör ve görev bazlı bir optimizasyon yapılabilir. Mevsimsel

olarak artış gösteren kaza yoğunlukları, iş planlamalarında denetim sıklığını ve güvenlik brifinglerinin zamanlamasını belirlemede kullanılabilir.

İş güvenliği uzmanları ve saha mühendisleri için geliştirilecek bir karar destek sistemi, bu analizlerden beslenerek kaza riskini anlık olarak değerlendirebilecek yazılım tabanlı çözümler sunabilir. Özellikle Python tabanlı modellerin otomasyona uyarlanarak iş güvenliği izleme sistemlerine entegre edilmesi, anlık raporlama ve uyarı sistemlerinin kurulmasına olanak tanıyacaktır.

Son olarak, kamu otoriteleri ve politika yapıcılar açısından da bu tür çok katmanlı analizlerin rehberlik edici yönü bulunmaktadır. Tersanelerde uygulanacak denetim politikaları, yatırım teşvikleri ve sektörel düzenlemeler, bu tür bilimsel analizlerle desteklendiğinde daha hedef odaklı ve etkili sonuçlar doğurabilir.

BÖLÜM 6. SONUÇ VE ÖNERİLER

Gemi inşa sektöründe meydana gelen iş kazalarının nedenlerini daha kapsamlı şekilde analiz edebilmek amacıyla, makine öğrenimi algoritmaları ile ÇKKV yaklaşımlarını (özellikle SF-Fuzzy TOPSIS yöntemini) entegre eden çok katmanlı bir model geliştirilerek geniş kapsamlı veri kümesi üzerinden yürütülen analizler sonucunda, hem yüksek doğruluk oranlarıyla çalışan öngörücü modeller elde edilmiş hem de kaza oluşumunu etkileyen faktörlerin göreceli önem dereceleri çok yönlü biçimde değerlendirilmiştir.

Elde edilen bulgular, kazazedenin görevinin iş kazalarının oluşumunda en kritik belirleyici olduğunu göstermektedir. Bununla birlikte, kaza türü, yaranın vücuttaki yeri ve Olay ayı gibi değişkenlerin de önemli risk faktörleri arasında yer aldığı belirlenmiştir. Makine öğrenmesi algoritmalarıyla elde edilen değişken önem düzeyleri, SF-Fuzzy TOPSIS yöntemiyle yapılan uzman görüşlerine dayalı değerlendirmelerle büyük ölçüde örtüşerek analizlerin geçerliliğini ve tutarlılığını desteklemiştir. Entropi tabanlı ağırlıklandırma yöntemiyle her bir analiz yaklaşımının katkı oranı nesnel biçimde belirlenmiş, bu oranlar compromise yöntemiyle bütünleştirilerek nihai faktör sıralaması oluşturulmuştur. Böylece hem nesnel hem de öznelliğe dayalı çok kriterli bir karar destek sistemi ortaya konmuştur. Modellerin %90'ın üzerinde doğrulukla sınıflandırma başarısı sergilemesi, geliştirilen yapının sektörel uygulamalarda güvenle kullanılabileceğini göstermektedir.

Bu doğrultuda, gemi inşa sektöründe iş güvenliğini artırmak için bazı temel öneriler geliştirilebilir. Öncelikle kazazedenin görevine özel risk temelli eğitimlerin ve önleyici politikaların oluşturulması gereklidir. Kaza türlerine göre teknik önlemler, mühendislik kontrolleri ve uygun iş süreçleri tasarlanmalı; mevsimsel risk yoğunlukları göz önüne alınarak işe alım, eğitim ve denetim süreçleri revize edilmelidir. Ayrıca, kişisel koruyucu donanım kullanımının artırılması, ergonomik iş tasarımı stratejilerinin benimsenmesi, kazalara ilişkin sürekli veri toplanması ve bu verilerin ileri analitik modellerle değerlendirilmesi de iş güvenliğini güçlendirecek diğer önemli adımlardır. Geliştirilen yöntem ve analiz yapısı yalnızca akademik düzeyde katkı sunmakla kalmamış, aynı zamanda iş güvenliği uzmanları, tersane yöneticileri ve politika yapıcılar için pratik ve uygulanabilir karar destek çıktıları üretmiştir. Bu bağlamda, gelecekteki çalışmaların daha büyük ve çeşitli veri setleriyle genişletilerek sektörel farkların da analiz edileceği kapsamlı uygulamalara yönelmesi önerilmektedir.

Bu çalışmada yürütülen analizlerde, makine öğrenimi algoritmalarının iş kazalarının öngörülmesinde oldukça yüksek performans sergilediği görülmüştür. Random Forest ve Extra Trees gibi algoritmalar, %90'ın üzerinde doğruluk oranı ile kazaların sınıflandırılmasında başarılı olmuş; bu yöntemlerin iş kazalarının meydana gelme koşullarını tahmin etmede etkili araçlar olduğu anlaşılmıştır. SF-Fuzzy TOPSIS yöntemi ise uzman görüşlerini sistematik hale getirerek faktörlerin önceliklendirilmesinde önemli katkı sunmuştur. Özellikle kazazedenin görevi hem uzman değerlendirmelerinde hem de makine öğrenimi sonuçlarında öne çıkarak yöntemin güvenilirliğini pekiştirmiştir.

Farklı analiz yöntemlerinin bütünleşik biçimde kullanılması için entropi tabanlı ağırlıklandırma yaklaşımı kullanılmış ve her yöntemin katkı düzeyi bu şekilde belirlenmiştir. Ardından, compromise yöntemiyle bu skorlar bir araya getirilerek daha dengeli ve güvenilir bir sıralama yapılmıştır. Bu yapı, nesnel ve öznel kaynakları entegre ederek güçlü bir karar destek mekanizması oluşturmuştur. Sonuçlar, kazazedenin görevinin kazaların oluşumunda en kritik faktör olduğunu ortaya koymuş; bunu sırasıyla kaza türü, yaranın vücuttaki yeri ve olay ayı izlemiştir. Bu sıralama, farklı yöntemlerle elde edilen skorların sentezlenmesi ile güvenilir şekilde belirlenmiş ve faktörler arasında yüksek düzeyde tutarlılık sağlanmıştır.

Elde edilen sonuçlar, makine öğrenimi algoritmalarının iş kazalarının tahmininde kullanılabileceğini güçlü şekilde göstermektedir. Random Forest ve Extra Trees gibi algoritmalar yüksek doğruluk oranları ile etkili sınıflandırma performansı sergilemiştir. Bu durum, ilerleyen çalışmalarda XGBoost, LightGBM veya derin öğrenme tabanlı modellerin de değerlendirilmesiyle öngörü kapasitesinin daha da güçlendirilebileceğini ortaya koymaktadır. Özellikle yapay zekâ temelli sistemler, geleneksel yöntemlerin ötesine geçerek karar vericilere hızlı, doğru ve veri odaklı çözümler sunabilir.

Modelin başarısı, gelecek araştırmalar açısından birçok yeni fırsatı da beraberinde getirmiştir. Mevcut veri kümesi belirli bir sektöre ve döneme özgü olduğundan, modelin farklı endüstrilere uygulanarak genellenebilirliği test edilebilir. İmalat, inşaat, enerji ve sağlık sektörleri gibi alanlarda benzer yöntemlerle yapılacak karşılaştırmalı analizler, sektörel farklılıkların anlaşılmasına katkı sağlayacaktır. Ayrıca, veri kümesine yeni değişkenlerin (psikososyal faktörler, eğitim düzeyi, vardiya süresi, ekipman tipi vb.) eklenmesiyle modelin açıklayıcılığı artırılabilir. Zamana bağlı değişkenlerle trend analizi yapılması da modelin dinamik hale getirilmesini sağlayabilir.

Bu çalışmada ağırlıklandırma entropi yöntemiyle yapılmış olsa da, CRITIC, Gini, PCA gibi alternatif nesnel yöntemlerle elde edilecek sonuçların kıyaslanması ilerideki araştırmalara katkı sağlayacaktır. Ağırlık hassasiyet analizleri yapılarak modelin sağlamlığı da test edilebilir. Ayrıca geliştirilen bu hibrit yaklaşımın karar destek sistemlerine entegre edilerek kullanıcı dostu yazılım ya da web tabanlı platformlar şeklinde uygulanması, saha koşullarında anlık risk değerlendirmesi yapılmasını mümkün kılabilir. Bu sistemlerin test edilmesi, geribildirimlerle iyileştirilmesi ve yaygınlaştırılması, araştırma çıktılarının pratik faydaya dönüşmesini sağlayacaktır.

İleride gerçekleştirilecek araştırmalarda, yalnızca akademik değil uygulamalı düzeyde de çalışanlar, işverenler ve denetçiler gibi farklı aktörleri içeren daha kapsayıcı modellerin geliştirilmesi önerilmektedir. Yapay zekâ teknolojilerindeki gelişmeler, iş güvenliği alanında daha proaktif yaklaşımlar geliştirmeyi mümkün kılmaktadır. Özellikle derin öğrenme ve görüntü işleme sistemlerinin entegrasyonu ile gerçek zamanlı risk izleme sistemleri oluşturulabilir. Açıklanabilir yapay zekâ (XAI) uygulamaları sayesinde modellerin karar verme süreçleri şeffaflştırılabilir. Gelecekte sadece geçmiş veriye dayalı değil, öngörücü, uyarıcı ve öğrenen sistemlerin geliştirilmesi; kazaların nedenleri yanında davranışsal kalıplar, erken uyarı sinyalleri ve iş ortamı dinamiklerinin de dikkate alındığı bütüncül güvenlik yönetim sistemlerinin hayata geçirilmesine katkı sağlayacaktır. Bu doğrultuda, yaş grubu veya cinsiyet gibi faktörlere göre özel stratejilerin (eğitim, koruyucu önlem, ergonomik düzenleme vb.) geliştirilmesi mümkün olabilecektir. Elde edilen bulguların, gemi inşa sektöründe uygulanacak İSG politikalarının geliştirilmesine katkı sağlaması beklenmektedir.

KAYNAKLAR

- Ahmed, T., Ahsan, A., Khan, M. H. R. B., Nahian, T. K., Antar, R. H., Hasan, A., Karim, M. R., Shafiquzzaman, M., & Imteaz, M. (2024). Comprehensive study on the selection and performance of the best electrode pair for electrocoagulation of textile wastewater using multi-criteria decision-making methods (TOPSIS, VIKOR and PROMETHEE II). *Journal of Environmental Management*, 363, 121337. <https://doi.org/10.1016/j.jenvman.2024.121337>
- Akinsola, J., Awodele, O., Kuyoro, S., & Kasali, F. (2019). Performance evaluation of supervised machine learning algorithms using multi-criteria decision making techniques. Proceedings of the International Conference on Information Technology in Education and Development (ITED),
- Ayvaz, T., & Ergör, O. A. (2023). TERSANELERDE İŞ SAĞLIĞI VE GÜVENLİĞİNİN GÖSTERGELER ÜZERİNDEN İRDELENMESİ. *Mühendis ve Makina*, 64(711), 259-296.
- Aziz, N. A. A., Nordin, R. M., Ismail, Z., & Yunus, J. (2018). Identifying Factors and Impacts of Occupational Safety and Health (OSH) towards Work Accident in Achieving Sustainable Construction. In.
- Barlas, B. (2012). Shipyard fatalities in Turkey. *Safety science*, 50(5), 1247-1252.
- Barlas, B., & Izci, F. B. (2018). Individual and workplace factors related to fatal occupational accidents among shipyard workers in Turkey. *Safety science*, 101, 173-179.
- Bevilacqua, M., Ciarapica, F. E., & Giacchetta, G. (2008). Industrial and occupational ergonomics in the petrochemical process industry: A regression trees approach. *Accident Analysis & Prevention*, 40(4), 1468-1479. <https://doi.org/10.1016/j.aap.2008.03.012>
- Bullejos, M., Cabezas, D., Martín-Martín, M., & Alcalá, F. J. (2022). A Python Application for Visualizing the 3D Stratigraphic Architecture of the Onshore Llobregat River Delta in NE Spain. *Water*, 14(12), 1882.
- Celebi, U. B., Ekinci, S., Alarcin, F., & Unsalan, D. (2010). The risk of occupational safety and health in shipbuilding industry in Turkey. Proceedings of the 3rd Int. Conf. Maritime and Naval Science and Engineering,
- Chen, C.-H. (2020). A novel multi-criteria decision-making model for building material supplier selection based on entropy-AHP weighted TOPSIS. *Entropy*, 22(2), 259.

- Cho, Y., Kim, Y.-C., & Shin, Y. (2017). Prediction model of construction safety accidents using decision tree technique. *Journal of the Korea Institute of Building Construction*, 17(3), 295-303.
- DeCamp, W., & Herskovitz, K. (2015). The theories of accident causation. In *Security supervision and management* (pp. 71-78). Elsevier.
- Fu, G., Xie, X., Jia, Q., Li, Z., Chen, P., & Ge, Y. (2020). The development history of accident causation models in the past 100 years: 24Model, a more modern accident causation model. *Process Safety and Environmental Protection*, 134, 47-82.
- Guldenmund, F. W. (2000). The nature of safety culture: A review of theory and research. *Safety Science*, 34(1-3), 215–257. [https://doi.org/10.1016/S0925-7535\(00\)00014-X](https://doi.org/10.1016/S0925-7535(00)00014-X)
- Hämäläinen, P., Takala, J., & Saarela, K. L. (2006). Global estimates of occupational accidents. *Safety science*, 44(2), 137-156. <https://doi.org/https://doi.org/10.1016/j.ssci.2005.08.017>
- Hezam, I. M., Mishra, A. R., Rani, P., Saha, A., Smarandache, F., & Pamucar, D. (2023). An integrated decision support framework using single-valued neutrosophic-MASWIP-COPRAS for sustainability assessment of bioenergy production technologies. *Expert Systems with Applications*, 211, 118674.
- Izci, F. B., Gökyay, O., & Barlas, B. (2023). Investigation of non-fatal occupational accidents and their causes in Turkish shipyards. *International journal of occupational safety and ergonomics*, 1-8.
- Jeong, B. Y. (2021). Prevalence of occupational accidents and factors associated with deaths and disabilities in the shipbuilding industry: Comparisons of novice and skilled workers. *Work*, 69(3), 997-1005.
- Kang KyungSu, K. K., & Ryu HangUk, R. H. (2019). Predicting types of occupational accidents at construction sites in Korea using random forest model.
- Khairuddin, M. Z. F., Hasikin, K., Abd Razak, N. A., Lai, K. W., Osman, M. Z., Aslan, M. F., Sabanci, K., Azizan, M. M., Satapathy, S. C., & Wu, X. (2022). Predicting occupational injury causal factors using text-based analytics: A systematic review. *Frontiers in public health*.
- Kim, K. W., Park, S. J., Lim, H. S., & Cho, H. H. (2017). Safety climate and occupational stress according to occupational accidents experience and employment type in shipbuilding industry of korea. *Safety and health at work*, 8(3), 290-295.

- Klaiber, M. (2021). *Predictive Analytics in Occupational Safety for Health Care in Ohio to Reduce Injuries* University of Cincinnati].
- Kutlu Gündoğdu, F., & Kahraman, C. (2019). Spherical fuzzy sets and spherical fuzzy TOPSIS method. *Journal of intelligent & fuzzy systems*, 36(1), 337-352.
- Kutlu Gündoğdu, F., & Kahraman, C. (2020). Spherical fuzzy sets and decision making applications. *Intelligent and Fuzzy Techniques in Big Data Analytics and Decision Making: Proceedings of the INFUS 2019 Conference, Istanbul, Turkey, July 23-25, 2019*,
- Li, Y., He, X., Martínez, L., Zhang, J., Wang, D., & Liu, X. A. (2024). Comparative analysis of three categories of multi-criteria decision-making methods. *Expert Systems with Applications*, 238, 121824. <https://doi.org/https://doi.org/10.1016/j.eswa.2023.121824>
- Li, Y., Yang, Z., Chen, X., Yuan, H., & Liu, W. (2019). A stacking model using URL and HTML features for phishing webpage detection. *Future Generation Computer Systems*, 94, 27-39.
- Louati, F., Ktata, F. B., & Amous, I. (2024). Enhancing Intrusion Detection Systems with Reinforcement Learning: A Comprehensive Survey of RL-based Approaches and Techniques. *SN Computer Science*, 5(6), 665.
- Lu, J., Ma, G., & Zhang, G. (2024). Fuzzy machine learning: A comprehensive framework and systematic review. *IEEE Transactions on Fuzzy Systems*.
- Lu, Q., & Liu, M. (2024). A multi-criteria compromise ranking decision-making approach for analysis and evaluation of community-integrated energy service system. *Energy*, 306, 132439.
- Manjunatha, A. (2023). *Injury Prediction in Mining Industry through Applied Machine Learning Approaches* Dublin, National College of Ireland].
- Martín-Martín, M., Bullejos, M., Cabezas, D., & Alcalá, F. J. (2023). Using python libraries and k-Nearest neighbors algorithms to delineate syn-sedimentary faults in sedimentary porous media. *Marine and Petroleum Geology*, 153, 106283.
- Matías, J. M., Rivas, T., Martín, J., & Taboada, J. (2008). A machine learning methodology for the analysis of workplace accidents. *International Journal of Computer Mathematics*, 85(3-4), 559-578.
- Murè, S., & Demichela, M. (2009). Fuzzy Application Procedure (FAP) for the risk assessment of occupational accidents. *Journal of Loss Prevention in the Process*

<https://doi.org/https://doi.org/10.1016/j.jlp.2009.05.007>

- Nguyen, P.-H. (2024). A data-driven MCDM approach-based spherical fuzzy sets for evaluating global augmented reality providers in education. *IEEE Access*, 13, 6102-6119.
- Olutimehin, A. T. (2025). The Synergistic Role of Machine Learning, Deep Learning, and Reinforcement Learning in Strengthening Cyber Security Measures for Crypto Currency Platforms. *Deep Learning, and Reinforcement Learning in Strengthening Cyber Security Measures for Crypto Currency Platforms (February 11, 2025)*.
- Oubellouch, H., & Aziz, S. (2024). Risk assessment for the liquefied petroleum gas filling industry using fuzzy logic and hazard and operability. *Salud, Ciencia y Tecnología*, 4, 749-749.
- Pietilä, J., Räsänen, T., Reiman, A., Ratilainen, H., & Helander, E. (2018). Characteristics and determinants of recurrent occupational accidents. *Safety science*, 108, 269-277. <https://doi.org/https://doi.org/10.1016/j.ssci.2017.12.020>
- Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*.
- Rashid, M. Z., AlQadi, H., Aljohani, H. M., & Akhter, A. S. (2025). NEW WEIGHTED PARETO DISTRIBUTION WITH APPLICATIONS TO MEDICAL AND ENGINEERING DATA. *Advances and Applications in Statistics*, 92(5), 701-725.
- Remeshevskaya, I., & Gurets, N. (2023). Identification of occupational accident relations of shipyard labour in terms of individual and workplace factors.
- Sachdeva, M., Lehal, R., Gupta, S., & Gupta, S. (2023). Influence of contextual factors on investment decision-making: a fuzzy-AHP approach. *Journal of Asia Business Studies*, 17(1), 108-128.
- Salmon, P. M., Stanton, N. A., Lenné, M., Jenkins, D. P., Rafferty, L., & Walker, G. H. (2017). *Human factors methods and accident analysis: practical guidance and case study applications*. CRC Press.
- Sarkar, S., Vinay, S., Raj, R., Maiti, J., & Mitra, P. (2019). Application of optimized machine learning techniques for prediction of occupational accidents. *Computers & Operations Research*, 106, 210-224.
- Shabani, T., Jerie, S., & Shabani, T. (2024). A comprehensive review of the Swiss cheese model in risk management. *Safety in Extreme Environments*, 6(1), 43-57.

- Shameli, P. (2024). *Cause and Sequence of Accident/Incidents in Shipyards* Høgskulen på Vestlandet].
- Sharaf, I. M. (2023). A new approach for spherical fuzzy TOPSIS and spherical fuzzy VIKOR applied to the evaluation of hydrogen storage systems. *Soft Computing*, 27(8), 4403-4423.
- Sikalo, M., Arnaut-Berilo, A., & Delalic, A. (2023). A Combined AHP-PROMETHEE Approach for Portfolio Performance Comparison. *International Journal of Financial Studies*, 11(1), 46.
- Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE transactions on systems, man, and cybernetics*(1), 116-132.
- Tsoukalas, V., & Fragiadakis, N. (2016). Prediction of occupational risk in the shipbuilding industry using multivariable linear regression and genetic algorithm analysis. *Safety science*, 83, 12-22.
- Vazirani, S., Sharma, A., & Sharma, P. (2020). Analysis of various machine learning algorithm and hybrid model for stock market prediction using python. 2020 International conference on smart technologies in computing, electrical and electronics (ICSTCEE),
- Wiegmann, D. A., & Shappell, S. A. (2001). *Human error analysis of commercial aviation accidents using the human factors analysis and classification system (HFACS)*.
- World Health Organization. (1995). Global Strategy on occupational health for all. The way to health at work. http://www.who.int/occupational_health/globstrategy/en.
- Wu, H.-C. (1999). Probability density functions of fuzzy random variables. *Fuzzy sets and Systems*, 105(1), 139-158.
- Xu, S., Liu, X., Cai, P., Li, J., Wang, X., & Liu, B. (2022). Machine-learning-assisted accurate prediction of molecular optical properties upon aggregation. *Advanced Science*, 9(2), 2101074.
- Yamada, K. (2008). A new combination of evidence based on compromise. *Fuzzy sets and Systems*, 159(13), 1689-1708.
- Yılmaz, A. (2020). *Hemşirelerin Delici ve Kesici Alet kullanımına Yönelik Tutumları ve Etkileyen Faktörlerin Belirlenmesi* Maltepe University (Turkey)].
- Yorulmaz, M., & Öztürk, M. A. (2024). Evaluation of Occupational Health and Safety Management System Performance in Shipyard Businesses with FAHP and FPROMETHEE Methods. *Pomorstvo*, 38(2), 169-187.

- You, W. S., Jeong, B. Y., & Lee, D. K. (2019). Comparison of Occupational Accidents and Illnesses in Shipbuilding Industry. *Journal of the Ergonomics Society of Korea*, 38(5).
- Zhang, Y., Zhang, G., Wang, H., & Sun, J. (2019). An ensemble multi-criteria decision-making approach integrating data mining and decision analysis. *Expert Systems with Applications*, 122, 27–42. <https://doi.org/10.1016/j.eswa.2018.11.025>
- Zhao, C., Li, Q. q., Harris Adamson, C., & Nemati, A. (2025). A fuzzy logic approach to improve the sensitivity of the rapid entire body assessment method. *International journal of occupational safety and ergonomics*, 1-12.

Ek 1

Veri kümesi Şablonu

KO D	KAZA SAATİ	KO D	KAZAZEDENİN GÖREVİ	KOD	KAZANIN TÜRÜ
1	08:01-08:30	50	Tavcı	116	Elektrik çarpması
2	08:31-09:00	51	Kaynakçı	117	Yüksekten düşme
3	09:01-10:00	52	Borucu	118	Taşıt kazası
4	10:01-11:00	53	Taşçı	119	Zehirlenme/gaza maruziyet/kaynak, metal, mobilya tozuna maruziyet
5	11:01-11:30	54	Montaj	120	Makinelerin Sebep Olduğu Kaza
6	11:31-12:00	55	Operatör	121	Patlama
7	12:00-13:00	56	Elektrik	122	Isıya/ışımaya (kaynak, elektrik vb.) maruz kalma
8	13:01-13:30	57	İskeleci	123	Cisim çarpması, malzemelerin düşmesi, batması, ezmesi, kesmesi, yakması
9	13:31-14:00	58	Mekanik	124	Sıkışma
10	14:01-15:00	59	İzolasyon	125	Vücut incinmeleri/burkma
11	15:01-16:00	60	Yardımcı Eleman	126	Vücuda cisim/çopak kaçması
12	16:01-16:30	61	Gemi Personeli	127	İntihar
13	16:31-17:00	62	Mühendis	128	Ekipmanların kesmesi, çizmesi, delmesi, çarpması, yakması vb.
14	17:01-18:00	63	Temizlikçi	129	Kavga, dövüş etme
15	18:01-19:00	64	Yangıncı	130	Doğal afet
16	19:01-20:00	65	Bozmacı	131	Sağlık problemleri (beyin kanaması, kalp krizi, tansiyon, şeker, epilepsi, sara vb.) bayılma
17	20:01-21:00	66	Meydancı	132	Boğulma
18	21:01-22:00	67	Makine Bakım	133	Kimyasala/radyasyona maruziyet
19	22:01-23:00	68	Depocu	134	Yangın
20	23:01-24:00	69	Mutfak	135	Kayma/takılma düşme
21	00:01-01:00	70	İş Mak. Ustası	136	Ekipman parçalarının düşmesi, kopması,
22	01:01-02:00	71	Tekniker	137	Vücudunu bir yere çarpma
23	02:01-03:00	72	İşaretçi	138	Sıcak sıvı dökülmesi
24	03:01-04:00	73	Havuzcu	139	Elektrik kablolarının patlaması
25	04:01-05:00	74	Kızakçı	KOD	KAZA SEBEBİ

26	05:01-06:00	75	Armador	140	Rahatsızlık/sağlık sorunu
27	06:01-07:00	76	Yağcı	141	Kişisel hata/hatalı kullanım
28	07:01-08:00	77	Personel	142	Yetersiz emn. Tedbiri
KOD	YAŞ	78	Enspektör	143	KKD olmaması
D					
29	15-18	79	Dalgıç	144	Dikkatsizlik
30	19-21	80	Kaptan	145	Uyg. iş eksp. /iş güv. Malzemesi, KKD Kullanmama
31	22-25	81	Hurdacı	146	Uygunsuz iş ortamı
32	26-30	82	Havuzcu	147	Arıza
33	31-35	83	İş Emniyet Elemanı	148	Periyodik kontrol-Bakım eksikliği
34	36-40	84	Elektronikçi	149	Malzeme hasarı/bozukluğu
35	41-50	85	Destek Elemanı	150	Enerjinin/basıncın kesilmemesi, kesilen enerjinin verilmesi
36	51-60	86	Patron	151	Çalışma izni almama
37	61-70	87	Raspa	152	Operatör belgesinin olmaması
KOD	BAĞLI		Boya	153	Eğitimsizlik
D	OLUNAN				
	İŞVEREN				
38	Asıl İveren (Tersane)	89	Diğer	154	İşaretçi/yardımcı bulunmaması
39	Alt İşveren	90	Formen	155	Koordine eksikliği/takipsizlik
40	Gemi Personeli	91	Teçhiz	KOD	KAZA SONUCU
41	Tersane Dışı	KOD	YARANIN VUCUTTAKİ YERİ	156	Ramak Kaza
KOD	UZMANLIK SINIFI	105	El	157	İş günü kaybı
D					
42	Yardımcı Eleman	106	Kol	158	Uzuv kaybı/maluliyet
43	Formen/Tersane Çalışanı	107	Ayak	159	Ölüm
44	Usta	108	Bacak	160	İlk yardım
45	Mühendis	109	Göz	161	Maddi hasar
46	Enspektör	110	Kulak	162	Yaralanma
47	Kaptan	111	Burun	163	Bayılma/bilinç kaybı
48	Firma Sahibi	112	Baş/Yüz	164	Meslek hastalığı
49	Gemi Personeli	113	Bel/Vücut	165	Çevre kirliliği
KOD	KAZALI	114	Akciğer/Göğüs	KOD	ÖNLENEBİLİRLİK
D	SAYISI				
92	0	115	Belirtilmemiş	166	Önlenemez

93	1	KOD	İŞE BAŞLAMA SAATİ	167	Önlenebilir
94	2	172	1	KOD	KAZANIN OLUŞMASI
95	3	173	2	168	Tehlikeli Durum
96	4 ve üstü	174	3	169	Tehlikeli Davranış
KOD	KAZANIN YERİ	175	4	KOD	OLAY AYI
97	Gemi	176	5	190	Ocak 198 Eylül
98	Açık Saha	177	6	191	Şubat 199 Ekim
99	Atölye/Depo	178	7	192	Mart 200 Kasım
100	İdari Bina/ofisler	179	8	193	Nisan 201 Aralık
101	İş Makineleri	180	9	194	Mayıs
102	Havuz	181	10	195	Haziran
103	Deniz	182	11ve üstü	196	Temmuz
104	Tersane Dışı			197	Ağustos
				198	Eylül
				199	Ekim

Ek-2

Veri kümesi Örneklem Tablosu

YAŞ	UZMANLIK SINIFI	KAZA SAATİ	BAĞLI OLUNAN İŞVEREN	KAZAZEDEN İN GÖREVİ	KAZALI SAYISI	KAZANIN MEYDANA GELDİĞİ YER	YARANIN VUCUTTA Kİ YERİ
35	44	3	39	53	93	97	109
35	44	9	39	56	93	97	113
32	44	11	39	57	93	97	112
31	44	24	39	51	93	97	112
30	42	4	39	60	93	97	109
35	43	11	39	60	93	97	106
31	44	14	39	54	93	97	108
35	44	6	38	66	93	98	108
35	44	2	38	64	93	98	108
35	44	3	39	58	93	97	109
33	44	20	38	71	93	99	105
32	44	24	39	88	93	102	107
32	42	3	39	60	93	97	107
30	44	10	39	88	93	97	112
30	44	6	39	88	93	98	107
32	44	4	39	51	93	102	108
31	44	2	38	64	93	99	105
34	42	4	39	60	93	97	105
33	42	6	39	60	93	97	113
34	44	6	39	88	93	97	107
31	44	5	38	64	93	97	109
32	44	13	39	88	93	97	105
-	49	-	40	61	93	100	113
32	44	10	39	88	93	102	112
33	44	12	39	52	93	97	107
-	49	-	40	61	93	97	105
31	44	18	39	63	93	97	113
32	45	9	38	62	93	97	107
32	44	4	39	51	93	97	105
31	44	2	39	57	93	97	105
30	42	11	39	60	93	97	105
34	44	4	39	54	93	97	105
31	42	11	39	60	93	97	112
35	44	4	39	54	93	97	112
30	42	11	39	60	93	97	108
35	44	13	39	54	93	97	112
31	42	10	39	60	93	97	105

35	42	15	39	60	93	97	105
32	44	4	39	58	93	97	107
32	44	14	38	85	93	97	105
33	44	11	39	88	93	97	105
31	44	4	39	88	93	97	113
35	44	11	38	74	93	103	114
34	44	2	39	51	93	97	107
33	44	8	39	88	93	102	112

Ek 3

Python Kodları

1. Kütüphaneler ve Ayarlar

```
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.model_selection import train_test_split, cross_validate, StratifiedKFold

from sklearn.preprocessing import LabelEncoder, StandardScaler

from sklearn.impute import SimpleImputer

from sklearn.pipeline import Pipeline

from sklearn.compose import ColumnTransformer

from sklearn.metrics import classification_report, accuracy_score, precision_score,
recall_score, f1_score, confusion_matrix, ConfusionMatrixDisplay, roc_auc_score,
roc_curve, auc

from sklearn.ensemble import RandomForestClassifier, ExtraTreesClassifier,
GradientBoostingClassifier, AdaBoostClassifier

from sklearn.linear_model import LogisticRegression

from sklearn.tree import DecisionTreeClassifier

from sklearn.svm import SVC

from sklearn.naive_bayes import GaussianNB

from sklearn.neighbors import KNeighborsClassifier

from imblearn.over_sampling import SMOTE

import shap

!pip install lime

from lime.lime_tabular import LimeTabularExplainer

import warnings

warnings.filterwarnings('ignore')

pd.set_option('display.max_columns', None)
```

```

from google.colab import files

# Excel dosyasını yükle

uploaded = files.upload()

# Excel dosyasını oku

df = pd.read_excel("Kazalar .xlsx")

from sklearn.preprocessing import LabelEncoder

label_encoders = {}

# Sadece kategorik sütunları al

categorical_cols = df.select_dtypes(include=["object"]).columns

for col in categorical_cols:

    le = LabelEncoder()

    df[col] = le.fit_transform(df[col].astype(str)) # Zorla string'e çevir

    label_encoders[col] = le

print(" Label encoding tamamlandı. Kodlanmış veri örneği:")

print(df.head())

# Tüm Sütunları Stringe Çevirme

df = df.astype(str)

# Sayısal değerlere dönüştürülebilen sütunları otomatik sayısal yap

for col in df.columns:

    try:

        df[col] = pd.to_numeric(df[col])

    except ValueError:

        pass # dönüştürülemeyen sütunlar string kalır

# Eksik verileri ve "-" değerlerini doldurma

imputer = SimpleImputer(strategy='most_frequent')

df_imputed = pd.DataFrame(imputer.fit_transform(df), columns=df.columns)

# Kategorik değişkenleri ve sayısal değişkenleri ayırma

categorical_cols = df_imputed.select_dtypes(include=['object']).columns

```



```

numerical_cols = df_imputed.select_dtypes(include=['int64', 'float64']).columns

numerical_cols = df_imputed.columns[df_imputed.apply(lambda x:
pd.api.types.is_numeric_dtype(x))].tolist()

categorical_cols = [col for col in df_imputed.columns if col not in numerical_cols]

import matplotlib.pyplot as plt

import seaborn as sns

# Seçilen sütunlar

selected_columns = [

    "YARANIN VUCUTTAKİ YERİ",

    "KAZAZEDENİN GÖREVİ",

    "KAZANIN TÜRÜ",

    "KAZALI SAYISI",

    "OLAY AYI",

    "UZMANLIK SINIFI"

]

# Barplot ile görselleştir

for col in selected_columns:

    if col in df.columns:

        plt.figure(figsize=(10, 4))

        sns.countplot(x=col, data=df, order=df[col].value_counts().index)

        plt.title(f'{col} - Sınıf Dağılımı", fontsize=14)

        plt.xlabel(col)

        plt.ylabel("Frekans")

        plt.xticks(rotation=45)

        plt.tight_layout()

        plt.show()

    else:

        print(f"Kolon bulunamadı: {col}")

df = df.apply(pd.to_numeric, errors='coerce')

```

```

# Oluşabilecek eksik değerleri most_frequent ile tekrar doldur

imputer = SimpleImputer(strategy='most_frequent')

df = pd.DataFrame(imputer.fit_transform(df), columns=df.columns)

import matplotlib.pyplot as plt

import seaborn as sns

# Önce sütun gerçekten df'te var mı kontrol et

if 'KAZA SEBEBİ' in df.columns:

    plt.figure(figsize=(8, 5))

    sns.countplot(x='KAZA SEBEBİ', data=df)

    plt.title("Kaza Sebebi Dağılımı", fontsize=14)

    plt.xlabel("Kaza Sebebi (Sınıf)")

    plt.ylabel("Olay Sayısı")

    plt.xticks(rotation=45)

    plt.tight_layout()

    plt.show()

else:

    print(" 'KAZA SEBEBİ' sütunu bulunamadı.")

# Hedef değişken (KAZA SEBEBİ) ve özelliklerin ayrılması

X = df_imputed.drop(columns=['KAZA SEBEBİ'])

y = df_imputed['KAZA SEBEBİ']

if 'KAZA SEBEBİ' in df.columns:

    X = df.drop(columns=['KAZA SEBEBİ'])

    y = df['KAZA SEBEBİ']

else:

    print(" 'KAZA SEBEBİ' sütunu bulunamadı.")

from sklearn.preprocessing import OrdinalEncoder

from sklearn.compose import ColumnTransformer

from sklearn.preprocessing import StandardScaler

```

```

preprocessor = ColumnTransformer(
    transformers=[
        ('num', StandardScaler(), numerical_cols),
        ('cat', OrdinalEncoder(), categorical_cols)
    ])

from sklearn.compose import ColumnTransformer
from sklearn.preprocessing import OneHotEncoder, StandardScaler
from sklearn.pipeline import Pipeline

# Hedef değişken ve özellikler

X = df_imputed.drop(columns=['KAZA SEBEBİ'])
y = df_imputed['KAZA SEBEBİ']

# Güncel veri tipi kontrolleri

categorical_cols = X.select_dtypes(include=['object', 'category']).columns.tolist()
numerical_cols = X.select_dtypes(include=['int64', 'float64']).columns.tolist()

print("Kategorik sütunlar:", categorical_cols)
print("Sayısal sütunlar:", numerical_cols)

# Preprocessing pipeline

preprocessor = ColumnTransformer(
    transformers=[
        ('num', StandardScaler(), numerical_cols),
        ('cat', OneHotEncoder(handle_unknown='ignore', sparse_output=False),
categorical_cols)
    ]
)

# Pipeline tanımı

pipeline = Pipeline(steps=[('preprocessor', preprocessor)])

# Özellikleri dönüştür

X = pipeline.fit_transform(X)

print(" X dönüşümü tamamlandı. Yeni şekil:", X.shape)

```

```

# Dengesiz veri setlerini dengeleme (SMOTE kullanarak)

smote = SMOTE(random_state=42)

X, y = smote.fit_resample(X, y)


# Veriyi eğitim ve test setlerine ayırma

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

from sklearn.metrics import classification_report, accuracy_score, roc_auc_score

def evaluate_model(model, X_train, y_train, X_test, y_test, model_name):

    model.fit(X_train, y_train)

    y_pred = model.predict(X_test)

    accuracy = accuracy_score(y_test, y_pred)

    # ROC AUC için kontrol

    try:

        if hasattr(model, "predict_proba"):

            y_proba = model.predict_proba(X_test)

            if y_proba.shape[1] == 2:

                roc_auc = roc_auc_score(y_test, y_proba[:, 1])

            else:

                roc_auc = roc_auc_score(y_test, y_proba, multi_class='ovr')

        else:

            roc_auc = None

    except:

        roc_auc = None

    print(f"\n {model_name} Skorları:")

    print(f'Accuracy: {accuracy:.4f}')

    if roc_auc is not None:

        print(f'ROC AUC: {roc_auc:.4f}')

    print("\n", classification_report(y_test, y_pred))

```

```

    return accuracy, roc_auc, y_pred

# Modellerin eğitilmesi ve değerlendirilmesi

rf_model = RandomForestClassifier(random_state=42)
gb_model = GradientBoostingClassifier(random_state=42)
logistic_model = LogisticRegression()
decision_tree_model = DecisionTreeClassifier()
svm_model = SVC(probability=True)
ada_model = AdaBoostClassifier(random_state=42)
extra_trees_model = ExtraTreesClassifier(random_state=42)
naive_bayes_model = GaussianNB()
knn_model = KNeighborsClassifier()

models = {
    'Random Forest': rf_model,
    'Gradient Boosting': gb_model,
    'Logistic Regression': logistic_model,
    'Decision Tree': decision_tree_model,
    'SVM': svm_model,
    'AdaBoost': ada_model,
    'Extra Trees': extra_trees_model,
    'Naive Bayes': naive_bayes_model,
    'KNN': knn_model
}

results = []

all_predictions = {}

from sklearn.model_selection import StratifiedKFold # Import StratifiedKFold

# EXTRA TREES – Baseline model eğitimi ve değerlendirme

et_baseline_model = ExtraTreesClassifier(random_state=42)

et_acc, et_auc, et_pred = evaluate_model(

```

```

    et_baseline_model, X_train, y_train, X_test, y_test, "Baseline Extra Trees"
)

print(f' Baseline ET Accuracy: {et_acc:.4f}, ROC AUC: {et_auc:.4f}')

results.append(["Baseline Extra Trees", round(et_acc, 4), None, None, None])

all_predictions["Baseline Extra Trees"] = et_pred

# Define kf before using it

kf = StratifiedKFold(n_splits=5, shuffle=True, random_state=42) # Define kf with
StratifiedKFold

# EXTRA TREES – Hyperopt fonksiyonu

def et_hyperopt(params):

    model = ExtraTreesClassifier(**params)

    score = cross_val_score(model, X_train, y_train, cv=kf, scoring='accuracy').mean()

    return {'loss': -score, 'status': STATUS_OK}

param_space_et = {

    'n_estimators': hp.choice('n_estimators', [100, 200, 300]),

    'max_depth': hp.choice('max_depth', [None, 10, 20, 30]),

    'min_samples_split': hp.choice('min_samples_split', [2, 5, 10]),

    'min_samples_leaf': hp.choice('min_samples_leaf', [1, 2, 4])

}

trials_et = Trials()

best_params_et = fmin(fn=et_hyperopt, space=param_space_et, algo=tpe.suggest,
max_evals=20, trials=trials_et)

print(" Best ET Params:", best_params_et)

best_et_model = ExtraTreesClassifier(

    n_estimators=[100, 200, 300][best_params_et['n_estimators']],

    max_depth=[None, 10, 20, 30][best_params_et['max_depth']],

    min_samples_split=[2, 5, 10][best_params_et['min_samples_split']],

    min_samples_leaf=[1, 2, 4][best_params_et['min_samples_leaf']],

    random_state=42

```

```

)

et_opt_acc, et_opt_auc, et_opt_pred = evaluate_model(
    best_et_model, X_train, y_train, X_test, y_test, "Optimized Extra Trees"
)

print(f" Optimized ET Accuracy: {et_opt_acc:.4f}, ROC AUC: {et_opt_auc:.4f}")
results.append(["Optimized Extra Trees", round(et_opt_acc, 4), None, None, None])
all_predictions["Optimized Extra Trees"] = et_opt_pred

from hyperopt import fmin, tpe, hp, Trials, STATUS_OK

# RANDOM FOREST – Baseline model eğitimi ve değerlendirme

rf_baseline_model = RandomForestClassifier(random_state=42)

rf_acc, rf_auc, rf_pred = evaluate_model(
    rf_baseline_model, X_train, y_train, X_test, y_test, "Baseline Random Forest"
)

print(f" Baseline RF Accuracy: {rf_acc:.4f}, ROC AUC: {rf_auc:.4f}")
results.append(["Baseline Random Forest", round(rf_acc, 4), None, None, None])
all_predictions["Baseline Random Forest"] = rf_pred

# RANDOM FOREST – Hyperopt fonksiyonu

def rf_hyperopt(params):
    model = RandomForestClassifier(**params)
    score = cross_val_score(model, X_train, y_train, cv=kf, scoring='accuracy').mean()
    return {'loss': -score, 'status': STATUS_OK}

param_space_rf = {
    'n_estimators': hp.choice('n_estimators', [100, 200, 300]),
    'max_depth': hp.choice('max_depth', [None, 10, 20, 30]),
    'min_samples_split': hp.choice('min_samples_split', [2, 5, 10]),
    'min_samples_leaf': hp.choice('min_samples_leaf', [1, 2, 4])
}

trials_rf = Trials()

```

```

best_params_rf = fmin(fn=rf_hyperopt, space=param_space_rf, algo=tpe.suggest,
max_evals=20, trials=trials_rf)

print(" Best RF Params:", best_params_rf)

best_rf_model = RandomForestClassifier(
    n_estimators=[100, 200, 300][best_params_rf['n_estimators']],
    max_depth=[None, 10, 20, 30][best_params_rf['max_depth']],
    min_samples_split=[2, 5, 10][best_params_rf['min_samples_split']],
    min_samples_leaf=[1, 2, 4][best_params_rf['min_samples_leaf']],
    random_state=42
)

rf_opt_acc, rf_opt_auc, rf_opt_pred = evaluate_model(
    best_rf_model, X_train, y_train, X_test, y_test, "Optimized Random Forest"
)

print(f" Optimized RF Accuracy: {rf_opt_acc:.4f}, ROC AUC: {rf_opt_auc:.4f}")

# EXTRA TREES – Baseline model eğitimi ve değerlendirme
et_baseline_model = ExtraTreesClassifier(random_state=42)

et_acc, et_auc, et_pred = evaluate_model(
    et_baseline_model, X_train, y_train, X_test, y_test, "Baseline Extra Trees"
)

print(f" Baseline ET Accuracy: {et_acc:.4f}, ROC AUC: {et_auc:.4f}")

results.append(["Baseline Extra Trees", round(et_acc, 4), None, None, None])

all_predictions["Baseline Extra Trees"] = et_pred

# EXTRA TREES – Hyperopt fonksiyonu
def et_hyperopt(params):
    model = ExtraTreesClassifier(**params)
    score = cross_val_score(model, X_train, y_train, cv=kf, scoring='accuracy').mean()
    return {'loss': -score, 'status': STATUS_OK}

param_space_et = {

```



```

'n_estimators': hp.choice('n_estimators', [100, 200, 300]),
'max_depth': hp.choice('max_depth', [None, 10, 20, 30]),
'min_samples_split': hp.choice('min_samples_split', [2, 5, 10]),
'min_samples_leaf': hp.choice('min_samples_leaf', [1, 2, 4])
}

trials_et = Trials()

best_params_et = fmin(fn=et_hyperopt, space=param_space_et, algo=tpe.suggest,
max_evals=20, trials=trials_et)

print(" Best ET Params:", best_params_et)

best_et_model = ExtraTreesClassifier(
    n_estimators=[100, 200, 300][best_params_et['n_estimators']],
    max_depth=[None, 10, 20, 30][best_params_et['max_depth']],
    min_samples_split=[2, 5, 10][best_params_et['min_samples_split']],
    min_samples_leaf=[1, 2, 4][best_params_et['min_samples_leaf']],
    random_state=42
)

et_opt_acc, et_opt_auc, et_opt_pred = evaluate_model(
    best_et_model, X_train, y_train, X_test, y_test, "Optimized Extra Trees"
)

print(f" Optimized ET Accuracy: {et_opt_acc:.4f}, ROC AUC: {et_opt_auc:.4f}")
results.append(["Optimized Extra Trees", round(et_opt_acc, 4), None, None, None])
all_predictions["Optimized Extra Trees"] = et_opt_pred

# Logistic Regression

from sklearn.metrics import classification_report, roc_auc_score

logistic_accuracy, logistic_roc_auc, logistic_pred = evaluate_model(
    logistic_model, X_train, y_train, X_test, y_test, "Logistic Regression"
)

# Detaylı sınıf bazlı raporu yazdır

print("\n Logistic Regression - Sınıf Bazlı Performans Raporu:")

```

```

print(classification_report(y_test, logistic_pred))

# ROC AUC ve Accuracy çıktısı

print(f'\n Logistic Regression Accuracy: {logistic_accuracy:.4f}, ROC AUC:
{logistic_roc_auc:.4f}')

# Decision Tree değerlendirme

cv_result = cross_validate(decision_tree_model, X_train, y_train, cv=kf,
                           scoring=['accuracy', 'precision_macro', 'recall_macro', 'f1_macro'])

dt_acc = cv_result['test_accuracy'].mean()

dt_prec = cv_result['test_precision_macro'].mean()

dt_rec = cv_result['test_recall_macro'].mean()

dt_f1 = cv_result['test_f1_macro'].mean()

results.append(["Decision Tree", round(dt_acc, 4), round(dt_prec, 4), round(dt_rec, 4),
round(dt_f1, 4)])

# Ayrıca test seti sonuçları da ayrı saklanabilir

decision_tree_accuracy, decision_tree_roc_auc, decision_tree_pred = evaluate_model(
    decision_tree_model, X_train, y_train, X_test, y_test, "Decision Tree"
)

all_predictions["Decision Tree"] = decision_tree_pred

# Detaylı sınıf bazlı raporu yazdır

print("\n Decision Tree - Sınıf Bazlı Performans Raporu:")

print(classification_report(y_test, decision_tree_pred))

# ROC AUC ve Accuracy çıktısı

print(f'\n Decision Tree Accuracy: {decision_tree_accuracy:.4f}, ROC AUC:
{decision_tree_roc_auc:.4f}')

# SVM değerlendirme

cv_result = cross_validate(svm_model, X_train, y_train, cv=kf,
                           scoring=['accuracy', 'precision_macro', 'recall_macro', 'f1_macro'])

svm_acc = cv_result['test_accuracy'].mean()

svm_prec = cv_result['test_precision_macro'].mean()

```

```

svm_rec = cv_result['test_recall_macro'].mean()

svm_f1 = cv_result['test_f1_macro'].mean()

results.append(["SVM", round(svm_acc, 4), round(svm_prec, 4), round(svm_rec, 4),
round(svm_f1, 4)])

# Test setinde ayrıca skorların

svm_accuracy, svm_roc_auc, svm_pred = evaluate_model(
    svm_model, X_train, y_train, X_test, y_test, "SVM"
)

all_predictions["SVM"] = svm_pred

# Sınıf bazlı detaylı classification raporu

print("\n SVM - Sınıf Bazlı Performans Raporu:")

print(classification_report(y_test, svm_pred))

# ROC AUC ve Accuracy çıktısı

print(f"\n SVM Accuracy: {svm_accuracy:.4f}, ROC AUC: {svm_roc_auc:.4f}")

# AdaBoost değerlendirme

# 1. Çapraz doğrulama metrikleri

cv_result = cross_validate(ada_model, X_train, y_train, cv=kf,
                            scoring=['accuracy', 'precision_macro', 'recall_macro', 'f1_macro'])

ada_acc = cv_result['test_accuracy'].mean()

ada_prec = cv_result['test_precision_macro'].mean()

ada_rec = cv_result['test_recall_macro'].mean()

ada_f1 = cv_result['test_f1_macro'].mean()

# 2. Sonuçları tabloya ekle

results.append(["AdaBoost", round(ada_acc, 4), round(ada_prec, 4), round(ada_rec, 4),
round(ada_f1, 4)])

# 3. Test setinde değerlendirme

ada_accuracy, ada_roc_auc, ada_pred = evaluate_model(
    ada_model, X_train, y_train, X_test, y_test, "AdaBoost"
)

```

```
all_predictions["AdaBoost"] = ada_pred
```

4. Sınıf bazlı detaylı metrikler

```
print("\n AdaBoost - Sınıf Bazlı Performans Raporu:")
```

```
print(classification_report(y_test, ada_pred))
```

5. Genel skorlar

```
print(f"\n AdaBoost Accuracy: {ada_accuracy:.4f}, ROC AUC: {ada_roc_auc:.4f}")
```

Extra Trees değerlendirme

1. Çapraz doğrulama metrikleri

```
cv_result = cross_validate(extra_trees_model, X_train, y_train, cv=kf,  
                           scoring=['accuracy', 'precision_macro', 'recall_macro', 'f1_macro'])
```

```
et_acc = cv_result['test_accuracy'].mean()
```

```
et_prec = cv_result['test_precision_macro'].mean()
```

```
et_rec = cv_result['test_recall_macro'].mean()
```

```
et_f1 = cv_result['test_f1_macro'].mean()
```

2. Sonuçları sonuç tablosuna ekle

```
results.append(["Extra Trees", round(et_acc, 4), round(et_prec, 4), round(et_rec, 4),  
               round(et_f1, 4)])
```

3. Test setinde modelin değerlendirilmesi

```
et_accuracy, et_roc_auc, et_pred = evaluate_model(  
    extra_trees_model, X_train, y_train, X_test, y_test, "Extra Trees"  
)
```

```
all_predictions["Extra Trees"] = et_pred
```

4. Sınıf bazlı detaylı classification report

```
print("\n Extra Trees - Sınıf Bazlı Performans Raporu:")
```

```
print(classification_report(y_test, et_pred))
```

5. Genel doğruluk ve ROC AUC

```
print(f"\n Extra Trees Accuracy: {et_accuracy:.4f}, ROC AUC: {et_roc_auc:.4f}")
```

Naive Bayes değerlendirme

```
from scipy.sparse import issparse
```

1. Naive Bayes için veriyi uygun formata getir

```
if issparse(X_train):
```

```
    X_train_nb = X_train.toarray()
```

```
    X_test_nb = X_test.toarray()
```

```
else:
```

```
    X_train_nb = X_train
```

```
    X_test_nb = X_test
```

2. Çapraz doğrulama

```
cv_result = cross_validate(naive_bayes_model, X_train_nb, y_train, cv=kf,  
                           scoring=['accuracy', 'precision_macro', 'recall_macro', 'f1_macro'])
```

```
nb_acc = cv_result['test_accuracy'].mean()
```

```
nb_prec = cv_result['test_precision_macro'].mean()
```

```
nb_rec = cv_result['test_recall_macro'].mean()
```

```
nb_f1 = cv_result['test_f1_macro'].mean()
```

3. Sonuçları ekle

```
results.append(["Naive Bayes", round(nb_acc, 4), round(nb_prec, 4), round(nb_rec, 4),  
               round(nb_f1, 4)])
```

4. Test seti değerlendirmesi

```
nb_accuracy, nb_roc_auc, nb_pred = evaluate_model(  
    naive_bayes_model, X_train_nb, y_train, X_test_nb, y_test, "Naive Bayes"  
)
```

```
all_predictions["Naive Bayes"] = nb_pred
```

5. Sınıf bazlı classification report

```
print("\n Naive Bayes - Sınıf Bazlı Performans Raporu:")
```

```
print(classification_report(y_test, nb_pred))
```

6. Accuracy ve ROC AUC çıktısı

```
print(f"\n Naive Bayes Accuracy: {nb_accuracy:.4f}, ROC AUC: {nb_roc_auc:.4f}")
```

KNN değerlendirme

```
from scipy.sparse import issparse
```

```
# 1. Veriyi uygun formata getir
```

```
if issparse(X_train):
```

```
    X_train_knn = X_train.toarray()
```

```
    X_test_knn = X_test.toarray()
```

```
else:
```

```
    X_train_knn = X_train
```

```
    X_test_knn = X_test
```

2. Çapraz doğrulama

```
cv_result = cross_validate(knn_model, X_train_knn, y_train, cv=kf,
```

```
                           scoring=['accuracy', 'precision_macro', 'recall_macro', 'f1_macro'])
```

```
knn_acc = cv_result['test_accuracy'].mean()
```

```
knn_prec = cv_result['test_precision_macro'].mean()
```

```
knn_rec = cv_result['test_recall_macro'].mean()
```

```
knn_f1 = cv_result['test_f1_macro'].mean()
```

3. Sonuçları tabloya ekle

```
results.append(["KNN", round(knn_acc, 4), round(knn_prec, 4), round(knn_rec, 4),  
               round(knn_f1, 4)])
```

4. Test setinde değerlendirme

```
knn_accuracy, knn_roc_auc, knn_pred = evaluate_model(
```

```
    knn_model, X_train_knn, y_train, X_test_knn, y_test, "KNN"
```

```
)
```

```
all_predictions["KNN"] = knn_pred
```

5. Sınıf bazlı classification report

```
print("\n KNN - Sınıf Bazlı Performans Raporu:")
```

```
print(classification_report(y_test, knn_pred))
```

6. Accuracy ve ROC AUC çıktısı

```
print(f'\n KNN Accuracy: {knn_accuracy:.4f}, ROC AUC: {knn_roc_auc:.4f}')
```

```
# Random Forest değerlendirme
```

1. Çapraz doğrulama

```
cv_result = cross_validate(rf_model, X_train, y_train, cv=kf,  
                           scoring=['accuracy', 'precision_macro', 'recall_macro', 'f1_macro'])
```

```
rf_acc = cv_result['test_accuracy'].mean()
```

```
rf_prec = cv_result['test_precision_macro'].mean()
```

```
rf_rec = cv_result['test_recall_macro'].mean()
```

```
rf_f1 = cv_result['test_f1_macro'].mean()
```

2. Sonuçları tabloya ekle

```
results.append(["Random Forest", round(rf_acc, 4), round(rf_prec, 4), round(rf_rec, 4),  
round(rf_f1, 4)])
```

3. Test setinde değerlendirme

```
rf_accuracy, rf_roc_auc, rf_pred = evaluate_model(  
    rf_model, X_train, y_train, X_test, y_test, "Random Forest"  
)
```

```
all_predictions["Random Forest"] = rf_pred
```

4. Sınıf bazlı classification report

```
print("\n Random Forest - Sınıf Bazlı Performans Raporu:")
```

```
print(classification_report(y_test, rf_pred))
```

5. Accuracy ve ROC AUC çıktısı

```
print(f'\n Random Forest Accuracy: {rf_accuracy:.4f}, ROC AUC: {rf_roc_auc:.4f}')
```

```
# DataFrame'e çevir
```

```
results_df = pd.DataFrame(results_list)
```

Sırala

```
results_df = results_df.sort_values(by="Accuracy",  
ascending=False).reset_index(drop=True)
```

Göster

```

print(" Otomatik Oluşturulan Model Performans Tablosu:")

display(results_df)

# 4 metrikli tabloyu göster

print(" Modellerin Performans Skorları (F1, Accuracy, Precision, Recall):")

display(results_df.round(4))

import matplotlib.pyplot as plt

import seaborn as sns

# Plot stilleri

plt.style.use("seaborn-v0_8-whitegrid")

colors = sns.color_palette("viridis")

metrics = ["F1-score", "Accuracy", "Precision", "Recall"]

titles = {

    "F1-score": "F1-Score Karşılaştırması",

    "Accuracy": "Accuracy Karşılaştırması",

    "Precision": "Precision Karşılaştırması",

    "Recall": "Recall Karşılaştırması"

}

# Her metrik için ayrı grafik

for metric in metrics:

    plt.figure(figsize=(10, 6))

    sns.barplot(data=results_df.sort_values(by=metric, ascending=False),

                y="Model", x=metric, palette=colors)

    plt.title(f' {titles[metric]}', fontsize=14)

    plt.xlabel(metric)

    plt.ylabel("Model")

    plt.xlim(0, 1)

    plt.tight_layout()

    plt.show()

```



```

import numpy as np

import pandas as pd

from collections import defaultdict

from sklearn.preprocessing import OneHotEncoder

# Model ve isim eşlemesi (optimize edilmiş modeller kullanılıyor)

feature_importance_models = {

    'Random Forest (Optimized)': best_rf_model,

    'Extra Trees (Optimized)': best_et_model

}

# Özellik adlarını al

num_features = numerical_cols

# Kategorik özellikler parçalanmış isimlerle gelir → orijinal isimleri gruplayacağız

cat_encoder = pipeline.named_steps['preprocessor'].transformers_[1][1]

cat_feature_names = cat_encoder.get_feature_names_out(categorical_cols)

# Tüm özellik isimleri

all_feature_names = num_features + list(cat_feature_names)

# Orijinal isim düzeyinde gruplanmış önemleri tutmak için

for model_name, model in feature_importance_models.items():

    if hasattr(model, 'feature_importances_'):

        print(f"\n {model_name} için en önemli özellikler:")

        importances = model.feature_importances_

        importance_df = pd.DataFrame({'Feature': all_feature_names, 'Importance': importances})

        importance_df['OriginalFeature'] = importance_df['Feature'].apply(lambda x: x.split('_')[0] if '_' in x else x)

# Gruplandır ve toplam önemleri yüzdeye çevir

grouped = importance_df.groupby('OriginalFeature').sum(numeric_only=True)

grouped['Importance (%)'] = (grouped['Importance'] / grouped['Importance'].sum())

* 100

```

```

# En önemli 4 özelliği seç

top_4 = grouped.sort_values(by='Importance (%)', ascending=False).head(4)

# Tabloyu göster

display(top_4[['Importance (%)']].round(2))

# Grafikle göster

plt.figure(figsize=(8, 5))

ax = sns.barplot(x=top_4['Importance (%)'], y=top_4.index, palette='Blues_d')

plt.title(f'{model_name} - En Önemli 4 Özellik')

plt.xlabel("Önem (%)")

plt.ylabel("Özellik")

for container in ax.containers:

    ax.bar_label(container, fmt="%.2f", label_type="edge")

plt.tight_layout()

plt.show()

else:

    print(f'{model_name} modelinde 'feature_importances_' özelliği yok.")

from sklearn.metrics import classification_report, confusion_matrix,
ConfusionMatrixDisplay

for model_name, model in {

    "Random Forest (Optimized)": best_rf_model,

    "Extra Trees (Optimized)": best_et_model

}.items():

    print(f'\n {model_name} - Tahmin Raporu')

# Tahmin et

y_pred = model.predict(X_test)

# Confusion Matrix

cm = confusion_matrix(y_test, y_pred)

disp = ConfusionMatrixDisplay(confusion_matrix=cm)

disp.plot(cmap='Blues', xticks_rotation=45)

```

```

plt.title(f'{model_name} - Confusion Matrix")
plt.grid(False)
plt.show()

# Classification Report

print(classification_report(y_test, y_pred))

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Kullandığımız modellerin listesi

feature_importance_models = {
    'Random Forest': rf_model,
    'Random Forest (Optimized)': best_rf_model,
    'Extra Trees': extra_trees_model,
    'Extra Trees (Optimized)': best_et_model,
    'Gradient Boosting': gb_model,
    'Decision Tree': decision_tree_model,
    'AdaBoost': ada_model
}

# Pipeline içinden kategorik + sayısal özellikleri al

num_features = numerical_cols # düz list

cat_encoder = pipeline.named_steps['preprocessor'].transformers_[1][1]
cat_feature_names = cat_encoder.get_feature_names_out(categorical_cols)
all_feature_names = num_features + list(cat_feature_names)

# Her model için özellik önemi

for model_name, model in feature_importance_models.items():
    if hasattr(model, 'feature_importances_'):
        print(f"\n {model_name} - Özellik Önem Dağılımı")

```

```

# 1. Feature importance değerlerini al

importances = model.feature_importances_

imp_df = pd.DataFrame({'Feature': all_feature_names, 'Importance': importances})

# 2. Kategorik olanları grupta:

imp_df['OriginalFeature'] = imp_df['Feature'].apply(lambda x: x.split('_')[0] if '_' in
x else x)

# 3. Gruplandır, yüzdeye çevir

grouped = imp_df.groupby('OriginalFeature').sum(numeric_only=True)

grouped['Importance (%)'] = (grouped['Importance'] / grouped['Importance'].sum())
* 100

# 4. Yüzdesel önem tablosu

display(grouped[['Importance (%)']].sort_values(by='Importance (%)',
ascending=False).round(2))

# 5. Grafik

plt.figure(figsize=(10, 7))

top_n = 20 if grouped.shape[0] > 20 else grouped.shape[0] # ilk 20 özelliği çiz

top_features = grouped.sort_values(by='Importance (%)',
ascending=False).head(top_n)

ax = sns.barplot(x=top_features['Importance (%)'], y=top_features.index,
palette='viridis')

plt.title(f'{model_name} - En Önemli {top_n} Özellik")
plt.xlabel("Önem (%)")
plt.ylabel("Özellik")

for container in ax.containers:

    ax.bar_label(container, fmt="%0.2f", label_type="edge")

plt.tight_layout()

plt.show()

else:

    print(f' {model_name} modelinde 'feature_importances_' özelliği bulunamadı.")

import shap

```

```

shap.initjs()

optimized_models = {
    "Random Forest (Optimized)": best_rf_model,
    "Extra Trees (Optimized)": best_et_model
}

sample_index = 5 # örnek gözlem

for model_name, model in optimized_models.items():
    print(f"\n SHAP Analizi Başlatılıyor: {model_name}")
    # X_test bir DataFrame değilse, isimlendirelim
    if not isinstance(X_test, pd.DataFrame):
        X_test_df = pd.DataFrame(X_test, columns=all_feature_names)
    else:
        X_test_df = X_test

    explainer = shap.TreeExplainer(model)
    shap_values = explainer.shap_values(X_test_df)

    # Sınıf çokluğuna göre ayarla
    if isinstance(shap_values, list):
        shap_vals = shap_values[0]
        expected_val = explainer.expected_value[0]
    else:
        shap_vals = shap_values
        expected_val = explainer.expected_value

    # Summary plot (dağılım)
    print(f" {model_name} – summary_plot (distribution):")
    shap.summary_plot(shap_vals, X_test_df)

    # Summary plot (bar)
    print(f" {model_name} – summary_plot (bar):")
    shap.summary_plot(shap_vals, X_test_df, plot_type="bar")

```

```
# Force plot

print(f" {model_name} – {sample_index}. gözlem için force_plot:")

shap.plots.force(expected_val, shap_vals[sample_index], matplotlib=True)


# Decision plot

print(f" {model_name} – {sample_index}. gözlem için decision_plot:")

shap.decision_plot(expected_val, shap_vals[sample_index],
feature_names=X_test_df.columns)
```