



**AMASRA TAŞKÖMÜRÜ İŞLETMESİNDE
İŞ KAZALARININ İNCELENMESİ**

YÜKSEK LİSANS TEZİ

Bilal ALTINDIŞ

Danışman

Doç. Dr. Fatih BAYRAM

MADEN MÜHENDİSLİĞİ ANABİLİM DALI

Ekim 2023

AFYON KOCATEPE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

AMASRA TAŞKÖMÜRÜ İŞLETMESİNDE İŞ KAZALARININ
İNCELENMESİ

Bilal ALTINDİŞ

Danışman

Doç. Dr. Fatih BAYRAM

MADEN MÜHENDİSLİĞİ ANABİLİM DALI

Ekim 2023

TEZ ONAY SAYFASI

Bilal ALTINDİŞ tarafından hazırlanan “Amasra Taşkömürü İşletmesinde İş Kazalarının İncelenmesi” adlı tez çalışması lisansüstü eğitim ve öğretim yönetmeliğinin ilgili maddeleri uyarınca 02 / 10 / 2023 tarihinde aşağıdaki jüri tarafından **oy birliği** ile Afyon Kocatepe Üniversitesi Fen Bilimleri Enstitüsü **Maden Mühendisliği Anabilim Dalı’nda YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Danışman : Doç. Dr. Fatih BAYRAM

Başkan : Prof. Dr. Yılmaz ÖZÇELİK
Hacettepe Üniversitesi, Mühendislik Fakültesi

Üye : Doç. Dr. Erkan ÖZKAN
Afyon Kocatepe Üniversite, Mühendislik Fakültesi

Üye : Doç. Dr. Fatih BAYRAM
Afyon Kocatepe Üniversite, Mühendislik Fakültesi

Afyon Kocatepe Üniversitesi
Fen Bilimleri Enstitüsü Yönetim Kurulu’nun
..... /..... /..... tarih ve
..... sayılı kararıyla onaylanmıştır.

.....
Prof. Dr. Bekir YALÇIN
Enstitü Müdürü

BİLİMSEL ETİK BİLDİRİM SAYFASI
Afyon Kocatepe Üniversitesi

Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- Görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- Başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- Atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- Kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- Ve bu tezin herhangi bir bölümünü bu üniversite veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

/ /

Bilal ALTINDİŞ

ÖZET

Yüksek Lisans Tezi

AMASRA TAŞKÖMÜRÜ İŞLETMESİNDE İŞ KAZALARININ İNCELENMESİ

Bilal ALTINDİŞ

Afyon Kocatepe Üniversitesi

Fen Bilimleri Enstitüsü

Maden Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Fatih BAYRAM

Günümüzde endüstrinin her kolunda olduğu gibi, madencilikte de gerek verimlilik gerekse iş güvenliği alanlarında çok sayıda veri toplanabilmektedir. Toplanan bu verilerin işletmeler için yararlı bilgiler haline getirilmesi gittikçe önem arz etmektedir. Verilerin işlenmesi ve işlenen verilerden faydalı bilgilerin çıkarılmasında veri madenciliği oldukça faydalıdır.

Bu tez çalışmasında, Türkiye Taşkömürü Kurumu Amasra Taşkömürü İşletmesinde 2010-2021 yılları arasında meydana gelen yaralanmalı iş kazaları verilerinin veri madenciliğiyle incelenmesi amaçlanmıştır. Bu amaç doğrultusunda, ilgili yıllara ait yaralanmalı kaza verileri işletme veri tabanından alınmış ve çalışmaya uygun biçimde düzenlenmiştir. Değerlendirme aşamasında ZeroR, J48 Karar Ağacı, k En Yakın Komşuluk, Yapay Sinir Ağları, Naive Bayes, Destekçi Vektör Makine ve Rassal Orman veri madenciliği algoritmaları kullanılmıştır. Bu algoritmaların uygulanması, açık kaynak kodlu bir uygulama olan WEKA programı ile yapılmıştır. Bu sonuçlara göre, işletme için kaza verilerini en iyi sınıflandıran ve tahmin eden algoritmalar tespit edilmeye çalışılmıştır. Farklı test yöntemlerine göre k En Yakın Komşuluk ve Destekçi Vektör Makine algoritmaları sınıflandırma ve tahmin çalışmalarında başarılı bulunmuştur. Çalışmanın önemli aşamalarından biri de birliktelik kuralları madenciliğiyle, kaza kökeninde etkili olan öz niteliklerin ve alt sınıfların belirlenmesi çalışmasıdır. Böylece, işletmede meydana gelen kazaların yerleri ve kök nedenleri konusunda daha detaylı bilgiler elde edilmiştir. Apriori ve Tahminci Apriori uygulamaları sonucunda kaza

yerlerine göre meydana gelen iş kazalarında kök nedenlerin işçilerin deneyimleri, vardiyadaki çalışma saati ve çalışan görevi olduğu ortaya konulmuştur. Bunun yanında, vardiya saati, kaza sebebi, aylık üretim ve aylık yevmiye değişkenlerinin de etkili oldukları tespit edilmiştir. Bu tespitler ışığında işletmeye yönelik öneriler sıralanmıştır.

2023, xi + 81 sayfa

Anahtar Kelimeler: Yeraltı kömür madenciliği, İş kazaları, Kömür üretimi, Veri madenciliği, WEKA, Birliktelik kuralları madenciliği.



ABSTRACT

M.Sc. Thesis

INVESTIGATION OF OCCUPATIONAL ACCIDENTS IN AMASRA HARD COAL ENTERPRISE

Bilal ALTINDİŞ

Afyon Kocatepe University

Graduate School of Natural and Applied Sciences

Department of Mining Engineering

Supervisor: Assoc. Prof. Dr. Fatih BAYRAM

Nowadays, as in every branch of industry, a large amount of data can be collected in mining, both in productivity and occupational safety. It is increasingly essential to transform this data into useful information for enterprises. Data mining is very useful in processing and extracting useful information from the processed data.

This thesis aims to analyze the data of occupational accidents with injuries between 2010 and 2021 in Amasra Hard Coal Enterprise of Turkish Hard Coal Corporation by data mining. For this purpose, the injured accident data for the relevant years were taken from the enterprise database and organized in a way suitable for the study. ZeroR, J48 Decision Tree, k Nearest Neighborhood, Artificial Neural Networks, Naive Bayes, Support Vector Machine, and Random Forest data mining algorithms were used in the evaluation phase. These algorithms were applied with the WEKA program, an open-source application. According to these results, it was tried to determine the algorithms that best classify and predict accident data for the enterprise. According to different test methods, k Nearest Neighborhood and Support Vector Machine algorithms succeeded in classification and prediction. One of the critical phases of the study is the determination of the attributes and subclasses that are effective in the origin of accidents by association rules mining. Thus, more detailed information was obtained about the locations and root causes of the accidents that occurred in the enterprise. As a result of Apriori and Predictive Apriori applications, it was revealed that the root causes of occupational accidents according to the accident locations are the experience of the workers, the working hours in the shift,

and the worker position. In addition, shift hours, accident cause, monthly production, and monthly wage variables were also influential. Given these results, recommendations for the enterprise are listed.

2023, xi + 81 pages

Keywords: Underground coal mining, Occupational accidents, Coal production, Data mining, WEKA, Association rules mining.



TEŞEKKÜR

Bu tez çalışmasının konusunun belirlenmesi, deneysel çalışmaların yönlendirilmesi, sonuçların değerlendirilmesi ve yazımı aşamasında yapmış olduğu büyük katkılarından dolayı tez danışmanım Sayın Doç. Dr. Fatih BAYRAM'a, her konuda öneri ve eleştirileriyle yardımlarını esirgemeyen hocalarım Prof. Dr. Yılmaz ÖZÇELİK'e, Doç. Dr. Erkan ÖZKAN'a ve arkadaşlarıma teşekkür ederim. Ayrıca bu çalışmanın gerçekleştirilmesinde, her türlü veriye erişimimde katkıları ve izni olan Amasra Taşkömürü İşletmesine şükranlarımı sunarım.

Bu araştırma boyunca maddi ve manevi desteklerinden dolayı aileme teşekkür ederim.

Bilal ALTINDİŞ
Afyonkarahisar 2023

İÇİNDEKİLER DİZİNİ

	Sayfa
ÖZET	i
ABSTRACT	iii
TEŞEKKÜR	v
İÇİNDEKİLER DİZİNİ.....	vi
SİMGELER ve KISALTMALAR DİZİNİ	viii
ŞEKİLLER DİZİNİ.....	ix
ÇİZELGELER DİZİNİ.....	xi
1. GİRİŞ.....	1
2. YERALTI KÖMÜR MADENCİLİĞİ	4
2.1 Yeraltı Kömür Üretim Yöntemleri	5
2.2 Amasra Taşkömürü İşletmesinde Uygulanan Yeraltı Kömür Üretim Yöntemleri	6
3. AMASRA TAŞKÖMÜRÜ İŞLETMESİ GENEL BİLGİLER.....	15
3.1 Amasra Taşkömürü İşletmesi Tarihçesi	16
3.2 Amasra Taşkömürü İşletmesi Sınırları	16
3.3 Amasra Taşkömürü İşletmesi Genel Jeolojisi	18
3.4 Amasra Taşkömürü İşletmesi Genel Durumu	22
3.5 Amasra Taşkömürü İşletmesi Ocak Kaynak Durumu	23
3.6 Amasra Taşkömürü İşletmesi İstihdam Durumu	23
3.7 Amasra Taşkömürü İşletmesi Kömür Üretim Miktarları	24
4. VERİ MADENCİLİĞİ	25
4.1 Veri Madenciliği Tarihçesi	27
4.2 Veri Madenciliği ve Bilgi Keşfi Süreci	28
4.3 Veri Madenciliği Yöntemleri.....	32
4.4 WEKA ile Veri Madenciliği	34
4.4.1 WEKA veri dosyasının hazırlanması	35
4.4.2 WEKA ile veri ön işleme	39
4.4.3 WEKA programında kullanılan algoritmalar.....	40
4.4.3.1 ZeroR.....	40
4.4.3.2 Naive Bayes.....	43
4.4.3.3 Karar ağacı (Decision tree).....	43
4.4.3.4 k en yakın komşuluk (k nearest neighbors – kNN).....	45
4.4.3.5 Yapay sinir ağları (Artificial neural networks)	45

4.4.3.6 Destekçi vektör makine (support vector machine – SVM)	47
4.4.3.7 Rassal orman (random forest)	47
4.5 Madencilikte Veri Madenciliği Uygulamaları	48
5. TEZ KAPSAMINDA YAPILAN ÇALIŞMALAR.....	52
5.1 Kaza Verilerinin Toplanması.....	53
5.2 Kaza verilerinin işlenmesi	54
5.3 Veri madenciliği çalışmaları	59
5.3.1 2010-2021 yılları kaza verilerinin veri madenciliği yöntemleriyle değerlendirilmesi	59
5.3.2 2010-2020 yılları kaza verilerinden 2021 yılı kaza verilerinin tahmini.....	65
5.3.3 2010-2021 yılları kaza verilerine yönelik birliktelik kuralları analizi	66
6. SONUÇLAR VE ÖNERİLER	75
7. KAYNAKLAR.....	78
ÖZGEÇMİŞ.....	81

SİMGELER ve KISALTMALAR DİZİNİ

Simgeler

%	yüzde
≤	küçük veya eşit
≥	büyük veya eşit
km ²	kilometrekare
m	metre
°	derece

Kısaltmalar

A.Ş.	Anonim şirket
ABD	Amerika Birleşik Devletleri
ANFIS	Uyarlamalı nöro-bulanık çıkarım sistemi
ANOVA	Analysis of variance
ARFF	Attribute-Relation File Format
ATİM	Amasra Taşkömürü İşletme Müessesesi
BKİ	Batı Karadeniz Kömür İşletmesi
EKİ	Ereğli Kömür İşletmesi
HAA	Hata ağacı analizi
kNN	k nearest neighbors
MAE	Mean absolute error
MLP	Multilayer Perceptron
MTA	Maden Tetkik ve Arama Genel Müdürlüğü
NIOSH	İş Sağlığı ve Güvenliği Ulusal Enstitüsü
RMSE	Root mean square error
SVM	Support vector machine
TMMOB	Türk Mühendis ve Mimar Odaları Birliği
TTK	Türkiye Taşkömürü Kurumu
WEKA	Waikato Environment for Knowledge Analysis

ŞEKİLLER DİZİNİ

	Sayfa
Şekil 2.1 -236/-300 tavan damar plan görünümü	7
Şekil 2.2 -236/-300 tavan damar güney kanatlar plan görünümü	8
Şekil 2.3 Klasik yöntem uzunayak perspektif görünüşü	8
Şekil 2.4 Klasik yöntem uzunayak içi görünümü.....	9
Şekil 2.5 Yarı mekanize ayak içi görünümüler	10
Şekil 2.6 -250/-40 6. damar basınçlı havalı patlatma plan görünümü.....	11
Şekil 2.7 -250/-300 kalın damar tam mekanize ayak plan görünümü.....	12
Şekil 2.8 -250/-300 mekanize ayak başı görünümü	13
Şekil 2.9 -250/-300 mekanize ayak içi görünümüler	14
Şekil 3.1 Amasra Taşkömürü İşletmesi yer bulduru haritası	15
Şekil 3.2 Amasra Taşkömürü İşletmesi işletme sınırı haritası	17
Şekil 3.3 Bartın-Amasra Havzası genelleştirilmiş stratigrafik kesiti	19
Şekil 3.4 Tavan damarın stampı ve kömür kalınlıkları	20
Şekil 3.5 Kalın damarın stampı ve kömür kalınlıkları	21
Şekil 3.6 Taşlı damarın stampı ve kömür kalınlıkları	21
Şekil 3.7 Alt taşlı damarın stampı ve kömür kalınlıkları	22
Şekil 4.1 Veri madenciliğinin etkileşim içinde olduğu bazı bilim dalları.....	26
Şekil 4.2 Bilgi keşfi sürecinde bir adım olarak veri madenciliği aşamaları.....	29
Şekil 4.3 Tipik bir veri madenciliği sisteminin mimarisi	32
Şekil 4.4 Veri madenciliği yöntemleri ve uygulanabilecek algoritmalar	33
Şekil 4.5 WEKA programı giriş arayüzü	35
Şekil 4.6 WEKA programında bir ARFF dosyası başlık bölümü görünümü.....	36
Şekil 4.7 WEKA programında bir ARFF dosyası veri bölümü görünümü	37
Şekil 4.8 WEKA Explorer arayüzü	37
Şekil 4.9 WEKA filtre özelliği görünümü.....	39
Şekil 4.10 WEKA programı örnek veri seti	40
Şekil 4.11 ZeroR algoritma sonucu.....	41
Şekil 4.12 ZeroR algoritma uygulaması sonucu elde edilen sınıflandırıcı çıktıları	42
Şekil 4.13 Karar ağacı algoritma sonucu elde edilen örnek bir karar ağacı	44
Şekil 4.14 Algoritma uygulamasında oluşan örnek bir yapay sinir ağı modeli	46

Şekil 5.1 Çalışmada izlenen yöntem	53
Şekil 5.2 Kaza yerlerine göre meydana gelen kaza sayıları	56
Şekil 5.3 Kaza yerlerine bağlı olarak diğer değişkenler açısından kaza sayıları	57
Şekil 5.3 Kaza yerlerine bağlı olarak diğer değişkenler açısından kaza sayıları (devam ediyor)	58
Şekil 5.4 Uygulanan çeyrekler açıklığı filtresi sonucu veri setindeki aykırı ve aşırı değerler.....	63
Şekil 5.5 Eğitim seti test yöntemine göre bazı algoritmaların hata matrisleri	64
Şekil 5.6 Çapraz doğrulama test yöntemine göre bazı algoritmaların hata matrisleri ...	64
Şekil 5.7 SVM algoritması hata matrisi	66



ÇİZELGELER DİZİNİ

	Sayfa
Çizelge 2.1 Yeraltı üretim yöntemleri sınıflamaları.....	6
Çizelge 3.1 ATİM kömür damarları ve damarlar arası kalınlık değerleri.....	20
Çizelge 3.2 ATİM ocak kaynak durumu	23
Çizelge 3.3 ATİM istihdam durumu	24
Çizelge 3.4 ATİM üretim miktarları	24
Çizelge 4.1 Veri madenciliği kronolojik gelişimi	27
Çizelge 5.1 Yıllara göre meydana gelen yaralanmalı iş kazaları sayıları.....	54
Çizelge 5.2 Kazaların sınıflandırılmasında kullanılan değişken ve bu değişkenlere ait alt sınıflar	55
Çizelge 5.3 2010-2021 yılları kaza verilerinin sınıflandırılmasında kullanılan algoritmalar ve başarı parametreleri	61
Çizelge 5.4 2021 yılı kaza verilerinin tahminine yönelik algoritmalar ve başarı parametreleri	66
Çizelge 5.5 Apriori yöntemiyle kaza yerleri için oluşturulan birliktelik kuralları	68
Çizelge 5.5 (Devam) Apriori yöntemiyle kaza yerleri için oluşturulan birliktelik kuralları.....	69
Çizelge 5.6 Apriori yöntemiyle oluşturulan birliktelik kurallarında öz nitelik ve alt sınıf dağılımları.....	70
Çizelge 5.7 Tahminci Apriori yöntemiyle kaza yerleri için oluşturulan birliktelik kuralları.....	71
Çizelge 5.7 (Devam) Tahminci Apriori yöntemiyle kaza yerleri için oluşturulan birliktelik kuralları	72
Çizelge 5.8 Tahminci Apriori yöntemiyle oluşturulan birliktelik kurallarında öz nitelik ve alt sınıf dağılımları	73

1. GİRİŞ

Yeraltı kaynakları, ülke ekonomilerinin gelişiminde önemli bir rol oynamaktadır. Bir ülkede yeraltı kaynakları olarak madenlerin üretilmesi ve işlenmesi, o ülkenin teknolojisinin gelişmesini, refah seviyesinin artmasını sağlar. Bu açıdan bakıldığında, hayatı kolaylaştırmak, gerekli araç-gereçlerin üretimini yapmak için madenlere ihtiyaç kaçınılmazdır. Dolayısıyla, endüstrinin tüm sektörleri (kimya, inşaat, makine, elektronik, bilişim, savunma sanayii vb.) madenlere ihtiyaç duymaktadır. Madencilik vazgeçilmez bir sektör olmasının yanında, sanayi kolları arasında iş sağlığı ve güvenliği açısından çok tehlikeli grupta yer almaktadır. Madencilik kapsamında yapılan faaliyetler oldukça tehlikeli durum barındıran, kaza ve ölüm riskinin yüksek olduğu çalışmalardır. Bu riskli faaliyetler sonucunda oluşan kazalarda yaralanmalar, uzuv kayıpları ve ölümler gerçekleşebilmektedir.

Madencilik faaliyetleri, genel anlamda açık ocak ve yeraltı işletmeciliği olarak yürütülmektedir. Her iki yöntem de iş sağlığı ve güvenliği açısından tehlikeli olmakla beraber yeraltı işletmeciliğine, yürütülen faaliyetler ve bunların barındırdığı tehlikeler göz önüne alındığında, daha önem verilmesi gerektiği ortaya çıkmaktadır. Hem açık ocak hem de yeraltı madenciliğinde yürütülen iş kalemleri düşünüldüğünde, dikkat edilmesi gereken hususların sayısının yeraltında fazla olduğu bir gerçektir. Yapılan madencilik faaliyetlerinde, iş sağlığı ve güvenliği yönünden sorunlar üretim aşamasında oluşmakta ve çalışanlar bu sorunlarla faaliyet süresince karşı karşıya kalmaktadırlar (Er 2018).

Yeraltı işletmeciliğinde tehlikeli kollardan biri de yeraltı kömür madenciliğidir. Özellikle göçükler, metan ve kömür tozu patlamaları sonucunda oluşan yangınlar önemli kazalara yol açmaktadır. Nitekim bu çalışmada inceleme alanı olarak seçilen Amasra Taşkömürü İşletmesinde 2022 yılı Ekim ayında bir patlama meydana gelmiş ve 43 maden işçisi hayatını kaybetmiştir. Yine Türkiye’de şu ana kadar yaşanmış en büyük maden kazalarından biri, 2014 yılında Manisa’nın Soma ilçesinde yeraltı kömür ocağında yaşanmıştır ve 301 işçinin ölümüyle sonuçlanmıştır. Bu kazalardan önce gerek sivil toplum kuruluşları (TMMOB Maden Mühendisleri Odası gibi) gerekse akademik çalışmalar neticesinde, işleyiş anlamında önlem alınmazsa bu tip kazaların oluşabileceği

öngörölmüştür. Bu tip kazalardan sonra hazırlanan raporlarda da alınmayan önlemlerden, yapılan ihmallerden söz edilmektedir.

Madencilik, iş güvenliği bakımından mevcut sanayi sektörleri arasında inşaat ve makine sektörleriyle beraber en tehlikeli gruptadır. Madencilik, çalışma ortamı ve şartları düşünüldüğünde en zor meslek grubudur. Dolayısıyla, tehlike ve zorluk seviyesi değerlendirildiğinde iş sağlığı ve güvenliğinin sağlanması mecburidir ve madencilik faaliyetlerinin planlanmalarının yapılmasından başlayarak düşünülmelidir. Çalışma ortamında iş sağlığı ve güvenliğinin sağlanması verimliliği artıracak önemli bir unsurdur.

Madencilik, mühendislik dalları arasında çok disiplinli, disiplinler arası faaliyetlerdir. Böyle olunca ilgili tüm disiplinlerin (inşaatla ilgili olarak yapı riskleri, makine kullanımına yönelik riskler gibi) içerdği riskleri kapsar. Bu risklerle mücadele anlamında tehlikeli durumların analizi ve risklerin sayısal bir şekilde ölçülmesi, önleme ve kabul edilebilir seviyelere indirme faaliyetleri için önemlidir. Bu tür ölçme faaliyetleri, maliyet ve verimlilik gibi alanlarda, özellikle iş sağlığı ve güvenliği alanında veri toplanmasını sağlamaktadır. Bu durumda da toplanan verilerin işlenmesi, yararlı bilgilerin çıkarılması ve iş güvenliğinde kullanımı ayrı bir iş kalemi haline gelmektedir (Erkayaoğlu 2020).

Madencilikle uğraşan mühendisler ve teknik personellerin birçoğu, iş sağlığı ve güvenliği konusunda veri toplama, elde ettikleri verileri işleme ve çeşitli yöntemlerle değerlendirerek işletmeler için yararlı bilgilerin çıkarımı konusunda uzman değillerdir. Aynı şekilde, maden işletmelerinde çalışan iş güvenliği uzmanları da bu konularda yetersizdir. Bu tip değerlendirmeler, veri işleme ve analizinde uzman kişiler tarafından yapılması gereken bir veri madenciliği çalışmasıdır. Veri madenciliği, özellikle karmaşık ve büyük verilerden anlamlı ve kullanılabilir çıkarımlar sağlayan bir bilimsel alandır. Bu alan sağlık, iktisat, ticaret, bilişim gibi sektörlerde yoğun bir şekilde kullanılmasına rağmen, madencilikte kullanımı oldukça kısıtlıdır. Madencilikte iş sağlığı ve güvenliği alanında da veri madenciliğiyle ilgili sınırlı sayıda çalışma bulunmaktadır. İş sağlığı ve güvenliği için oluşturulabilecek veri tabanları düşünüldüğünde, bu veri setlerinden gerekli bilgilerin çıkarımında veri madenciliği uygulamalarının avantajı büyüktür. Klasik istatistiksel yöntemlerle bulunamayacak ilişkiler veri madenciliği yöntemleriyle elde

edilebilir (Erkayaoğlu 2020).

Bu çalışmada da yukarıda belirtilen nedenlerden dolayı, çalışma alanı olarak belirlenen Amasra Taşkömürü İşletmesinde meydana gelen iş kazalarının veri madenciliği ile değerlendirilmesi amaçlanmıştır. Giriş bölümüyle başlayan çalışma, hem yeraltı kömür madenciliğinin hem de veri madenciliğinin anlatıldığı, literatürdeki ilişkili çalışmaların özetlendiği bölümlerle devam etmektedir. Daha sonraki bölümler tez kapsamında yapılan çalışmaların ve sonuçların verildiği bölümlerdir.



2. YERALTI KÖMÜR MADENCİLİĞİ

Ülkelerin gelişiminde etkili olan sanayilerin ihtiyaç duyduğu hammaddelerin büyük bir kısmı madencilik faaliyetlerinden elde edilmektedir. Madencilik faaliyetleri;

- a) Yerüstü (açık ocak) madenciliği,
- b) Yeraltı madenciliği,

olmak üzere iki şekilde gerçekleştirilir. Bu iki yöntem arasında yeraltı madenciliği, içerdiği çok sayıda risk sebebiyle daha tehlikelidir. Bu tehlike ve riskler de yeraltı madenciliğinin teknik ve jeolojik şartları gereğidir. İş kazalarının en sık yaşandığı endüstri kollarından biri olan madencilikte, mevcut risklerin üst üste gelip birbirini etkilemesi iş sağlığı ve güvenliğinin önemini ortaya koymaktadır. Bu çok çeşitli riskler nedeniyle madencilik sektöründe çalışanların iş kazası geçirme olasılığı fazladır. Bu yönüyle hem yerüstü hem de yeraltı madenciliğinde iş sağlığı ve güvenliği kapsamında önlemlerin alınması ve bunlara uyulması esastır.

Günümüzde uygulanmakta olan tüm iş sağlığı ve güvenliği tedbirlerinin ana hedefi, iş kazası ve meslek hastalığı seviyesini en aza düşürmektir. Bu hedefe ulaşmak için, özellikle bilimsel ve teknolojik araştırmalar yapılarak, risklerin belirlenmesi ve bu risklere dair tedbirlerin çalışma planında ve üretim sürecinde uygulanması ile mümkün olacaktır. Günümüzde gelişmiş ülkelerde alınan bu tedbirlerin çalışanlar için daha sağlıklı ve güvenli bir çalışma ortamı oluşturduğu ve bu ortamın daha verimli bir çalışma ortaya çıkardığı görülmektedir. Sonuç olarak, sağlanan bu ortam ve yapılan verimli çalışmalar toplumsal kalkınmaya da pozitif yönde fayda sağlayacaktır (Er 2018).

Madencilikte özellikle üretim miktarı ve iş güvenliği açısından yerüstü (açık ocak) madenciliği daha çok tercih edilmektedir. Yerüstü madenciliğinin uygulanamayacağı ekonomik maden yataklarında yeraltı madenciliğinin kullanımı kaçınılmazdır. Tüm dünyada olduğu gibi ülkemizde de kömür madenciliğinde yeraltı işletmeciliği kullanılabilmektedir. Kömür madenciliğinde de yeraltı üretim yöntemi olarak, uzunayak yöntemi öne çıkmaktadır.

2.1 Yeraltı Kömür Üretim Yöntemleri

Ekonomik bir cevher yatağı işletmeye alınmadan önce, yapılan detaylı incelemelerle yatakla ilgili yeterli bilgi toplanmalıdır. Bundan sonra en uygun madencilik yöntemini veya yöntemlerini seçmenin süreci başlar. Bu aşamada, seçim sadece bir proje ve fizibilite çalışması için temel olarak bir başlangıç niteliğindedir. Kullanılan temel prensipler açısından, günümüzde nispeten az sayıda madencilik yöntemi kullanılmaktadır. Her cevher yatağının farklılığı nedeniyle, bu yöntemlerin her birindeki varyasyonlar neredeyse sınırsızdır (Hustrulid ve Bullock, 2001).

Kömür oluşumu itibariyle turba, linyit ve taşkömürü olarak üç gruba ayrılabilir. Sedimanter kömür damarları farklı kalınlıklarda ve eğimlerde olabilir. Kömür yatakları derinlikleri itibariyle yerüstü ve yeraltı madencilik yöntemlerine elverişlidir. Yeraltı madenciliği genel hatları ile düşünüldüğünde farklı sınıflamalara tabi tutulabilir. Bu sınıflama iki temel üzerine oturtulabilir (Çizelge 2.1). Bunlar;

- a) Tavanın kontrol altına alınmasına göre
- b) Kazı arınının durumuna göre

Uygulanacak yeraltı kömür üretim yönteminin seçiminde cevherin tabakalaşması, damar kalınlığı, tabaka eğimi, mukavemeti ve cevher üstü tabakaların jeolojik özellikleri gibi birçok etken göz önünde bulundurulur. Tabakalaşmanın yatay ve az eğimli olduğu geniş cevher yataklarında, ağırlıklı olarak oda-topuk yöntemi uygulanmaktadır. Bu yöntem çoğunlukla Kuzey Amerika’da uygulanmaktadır. İngiltere ve Japonya gibi ülkelerde ise taşkömürü madenciliğinde kısmen kullanılmaktadır. Yüksek kazı randımanı sağlamanın yanında, bırakılan topuklardan dolayı üretim kaybına neden olması ve yüksek hazırlık maliyeti gibi dezavantajları da vardır (TBMM 2023).

Ülkemizde yeraltı kömür madenciliğinde ağırlıklı olarak uzunayak yöntemi kullanılmaktadır. Bu yöntem genelde emek yoğun bir çalışma gerektiren klasik uzunayak yöntemi olarak uzun zamandan beri uygulanmaktadır. Günümüzde teknolojinin de gelişmesiyle daha çok makine ve ekipmana dayalı, yarı mekanize ve tam mekanize ayak

uygulama yöntemleri daha çok kullanılır hale gelmiştir. Uzunayak üretim yöntemi, 0,4-4 m damar kalınlığında 0°-40° damar eğimine kadar olan kömür yataklarında uygulanan bir yöntemdir. 4 m'den daha kalın damarlarda damar dilimlere ayrılarak yukarıdan aşağıya göçertmeli ya da aşağıdan yukarıya dolgulu olarak üretim yapılmaktadır.

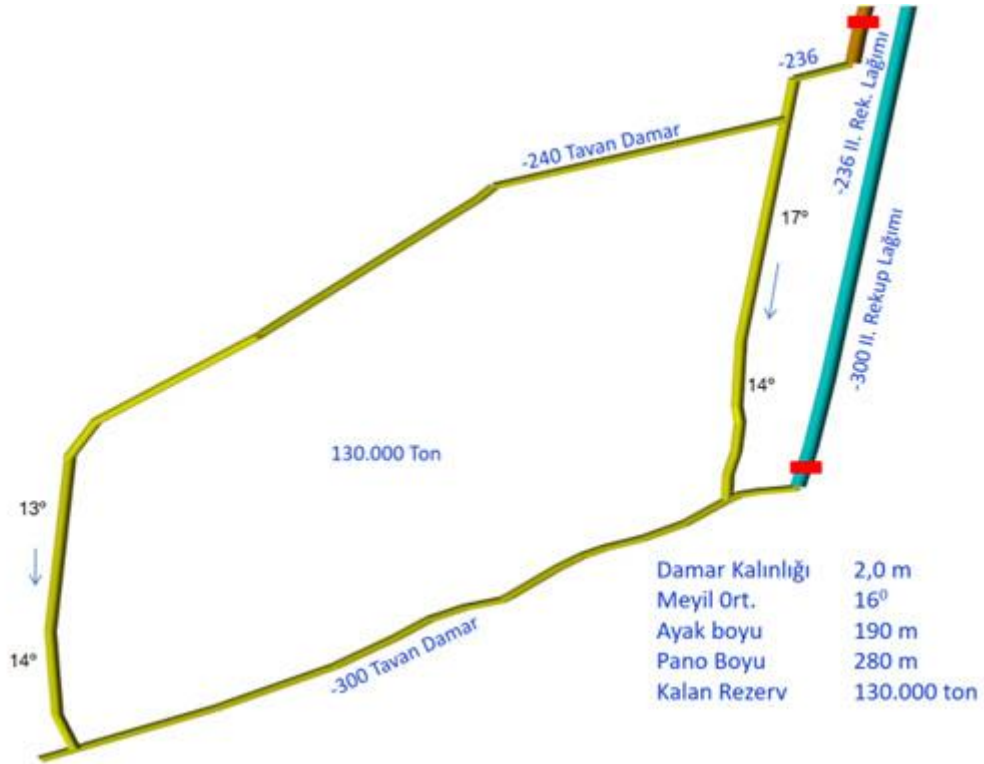
Çizelge 2.1 Yeraltı üretim yöntemleri sınıflamaları (Saltoğlu, 1988'den uyarlanmıştır)

Tavanın kontrol altına alınmasına göre	Alman sınıflaması	c) Tavanın, bırakılan cevher ve topuklarının desteği ile korunduğu yöntemler,
		d) Askıda kalan tavanın taş ve benzeri malzemelerle doldurulduğu yöntemler,
		e) Tavanın göçertildiği yöntemler
	Amerikan sınıflaması	f) Cevheri alınan yerin boş bırakılması (açık oda yöntemleri),
		g) Açılan boşluğun kazılmış cevher ile geçici olarak doldurulması (ambarlama yöntemleri),
		h) Açılan boşluk tavanının tahkimat ile tutulması (tahkimatlı yöntemler),
		i) Açılan boşluğun taş ile doldurulması (dolgulu yöntemler),
		j) Tavanın göçertilerek açılan boşluğun doldurulması (göçertmeli yöntemler),
		k) Kombine üretim yöntemleri
Kazı arınının durumuna göre	l) Uzun kazı arınlı yöntemler (Uzunayak, Diyagonal ayak)	
	m) Dar kazı arınlı yöntemler (Taban, Tavan, Yanal, Travers arınlı ayak)	
	n) Topuklu yöntemler (Topuklu, Travers, Arakatlı göçertmeli ayak)	
	o) Oda yöntemi (Oda, Arakatlı kazı, Oda-topuk)	
	p) Blok halinde üretim yöntemi (Blok göçertme)	

2.2 Amasra Taşkömürü İşletmesinde Uygulanan Yeraltı Kömür Üretim Yöntemleri

Amasra Taşkömürü İşletme Müessesesi (ATİM) sahasında bulunan damar yapıları göz önüne alındığında, ortalama 0-4 m damar kalınlığı ve $\leq 40^\circ$ eğimdeki tabakalaşma yoğunluktadır. Bu nedenle, ağırlıklı olarak geri dönüşlü göçertmeli uzunayak yöntemi

uygulandığı görülmektedir. Geri dönüşlü göçertmeli uzunayak yönteminde, pano içerisinde alt ve üst taban yolları sonuna kadar sürülür ve başyukarı ile birbirine bağlanır. Bu şekilde oluşturulan panoda ters yönde panonun başlangıcına doğru kazı işlemi yapılarak ilerleme sağlanır. Şekil 2.1’de görüldüğü gibi, -236/-300 tavan damar ayak panosu hazırlığı yapılmış, -240 kotunda üst taban yolu ve -300 kotunda alt taban yolu sürülmüştür. Sürülen bu taban yolları dip taraflarda 14° ve üst kısımlara doğru 13° eğime sahip bir başyukarı ile birbirine bağlanmış ve pano üretime hazır hale getirilmiştir.



Şekil 2.1 -236/-300 tavan damar plan görünümü

Şekil 2.2’de de bölgedeki farklı panoların durumları görülmektedir. Bölgede kömür damarı içerisinde süreksizlikler ve ondülasyonlar bulunmaktadır. Bu sebeple aynı bölgede bir panoda ortalama eğim 51° iken, diğer panoda 30° olabilmektedir. Dolayısıyla ATİM çalışma sahasında, aynı bölgede farklı panolarda farklı üretim yöntemleri uygulanabilmektedir.

Klasik yöntemle gerçekleştirilen uzunayaklarda kömür üretimi, kömürün patlayıcı madde ile gevşetilip sabit oluklar ile nakledilmesi esasına dayanmaktadır. Kömürde açılan

-236/-300 Tavan Damar Ayak

Kalınlık	2,20 m
Meyil	51°
Ayak boyu	84 m
Pano boyu	260 m
Rezerv	82.000 ton
Kalan Rezerv	72.500 ton

Damar Kalınlığı	2,20 m
Meyil Ort.	30°
Ayak boyu	8 m
Pano Boyu	33 m
Rezerv	1000 ton

8



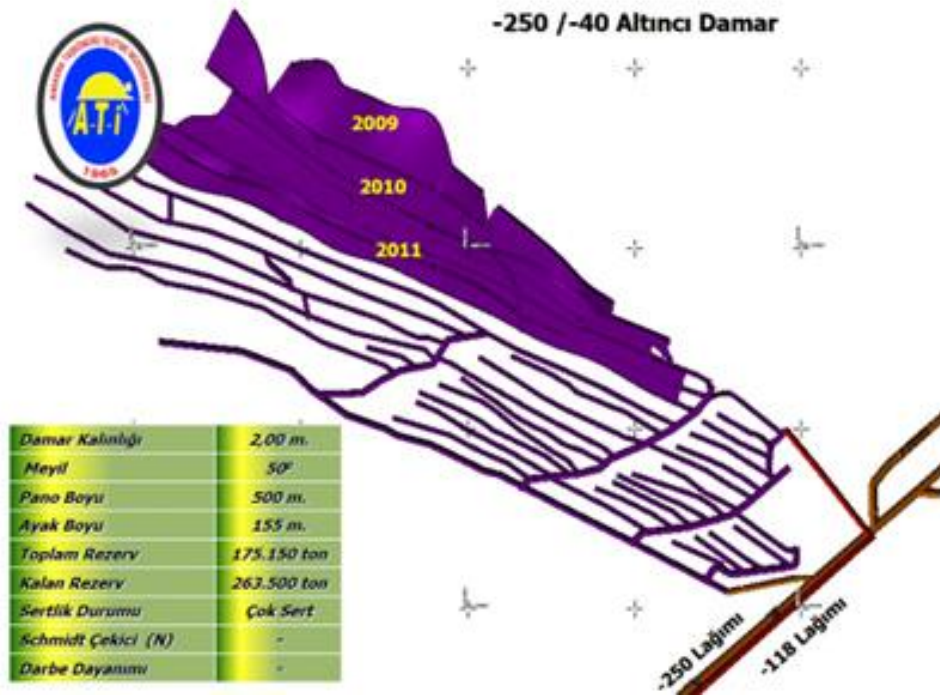
Şekil 2.4 Klasik yöntem uzunayak içi görünümü

Teknolojinin gelişmesi ile birlikte, klasik yönteme göre daha ekonomik ve daha yüksek üretim hızına sahip yarı mekanize ayak yöntemine geçilmiştir. Şekil 2.5'te yarı mekanize ayak yönteminin uygulanmasına yönelik ayak içi görünüm verilmektedir. Bu yöntemde de yine kömür kazı işlemi delme-patlatma yöntemi ile yapılmaktadır. Kazılan kömür, ayak içerisinden konveyörler vasıtasıyla nakledilmekte, fakat tahkimat sistemi olarak kalkan tahkimatlar kullanılmaktadır. Bu yöntemde ayak ilerleme hızı, tahkimat öteleme hızı klasik yönteme göre daha yüksektir.



Şekil 2.5 Yarı mekanize ayak içi görünüm

İşletme jeolojik yapısı itibarıyla kazılması zor, sert kömür damarlarına sahiptir. 45° üzerindeki eğimli kömür damarları 15 metre aralıklarla, kılavuz sürülerek basınçlı havalı (pnömatik) patlatma sistemiyle kazılmaktadır. Basınçlı havalı patlatma yöntemiyle yapılan kazıda, kömür çok sert olması nedeniyle taban ve tavan taşından ayrılmamakta ve askıda kalmaktadır. Dolayısıyla, rezerv kaybı gerçekleşmektedir. Şekil 2.6’da görüldüğü gibi, -40/-250 kotları arasında 6. damar içerisinde 15 m kot farkı ile sürülen kılavuzlarda yukarıdan aşağı kotlara doğru her kılavuzda arından geriye doğru basınçlı havalı patlatma ile kazılan kömür yine konveyörler vasıtasıyla nakledilerek üretim gerçekleştirilmektedir.



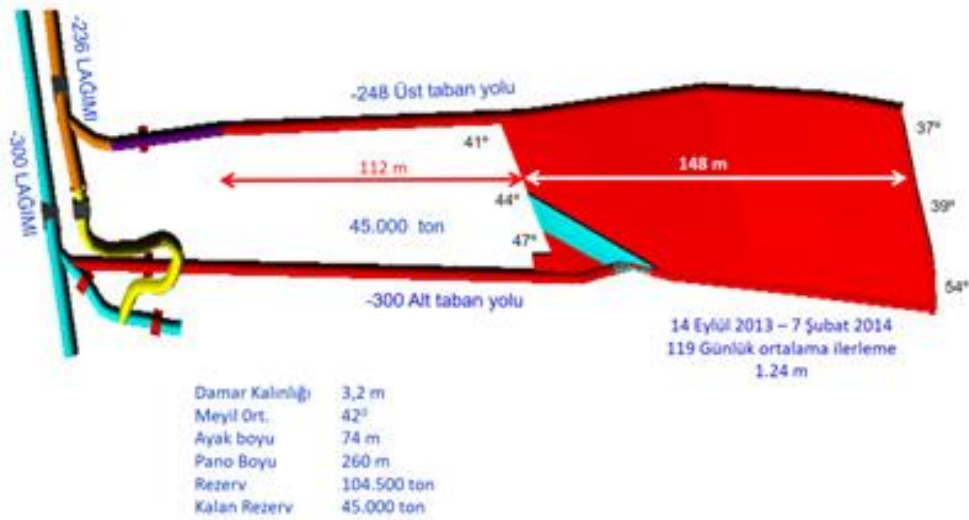
Şekil 2.6 -250/-40 6. damar basınçlı havalı patlatma plan görünümü

Bu yöntemde, kontrol edilemeyen tavan taşı düşerek kömüre karışmakta ve temiz kömür üretimini olumsuz yönde etkilemekte, üretim yerinden kömür yıkama tesisine kadar olan bölgede işçilik ve enerji maliyeti artmakta ve lavvar tesislerinde ise fazla enerji sarfıyatı olmaktadır.

Amasra kömürlerinin tutuşma sıcaklığı 140 – 144 °C arasında değişmektedir. Ayrıca damarların nem içeriğinin %2,8 – 3,5 gibi düşük ve uçucu madde miktarının % 31 – 35 gibi yüksek oluşundan dolayı yangına müsait olması nedeniyle, alınamayan kömürler

oksidasyona uğrayarak ocak yangınlarına neden olmaktadır. Bu durum çalışanların can güvenliğini, üretim sürecini, bir alt kat hazırlık dönemini olumsuz yönde etkilemektedir. Yangından sonra açılan panolardaki taban, kılavuz ve başyukarılarda yapılan bakım onarım gibi işler, işçilik ve malzeme giderlerinin, dolayısıyla üretim giderlerinin artmasına neden olmaktadır.

İşletme rezerv kayıplarının önlenmesi, üretim miktarının artırılması, birim maliyetlerinin düşürülmesi, dönem zararlarının azaltılması ve güvenli çalışma ortamında üretim yapılması için $\geq 45^\circ$ kömür damarlarında da tam mekanize üretim yöntemi kullanılmaya başlanmıştır. Şekil 2.7’de tam mekanize sistem ile üretim yapılacak -248 kotunda üst taban yolu ve -300 kotunda alt taban yolu sürülerek hazırlanan kalın damar panosu şematik görünümü verilmiştir.



Şekil 2.7 -250/-300 kalın damar tam mekanize ayak plan görünümü

Bu sistemde, işletmede 8 saatlik 3 vardiya usulüne göre çalışılmaktadır. Bu şekilde günlük ortalama 2,4 m ilerleme yapılmaktadır. Günlük ortalama 3 adet kesim, 3 adet tahkimat ötelemesi ve tahkimat ve kesici bakımları gerçekleştirilmektedir. Ayak içerisinde (ayak dibinde) kesicinin girebileceği 6 m uzunluğunda bir cep bulunmaktadır. Kesim aşağıdan yukarı doğru yapılmaktadır. Kesici aşağıdan yukarı doğru çıkarken normal kesimini yaparak çıkar ve aşağı inerken eğer tavan taşından ayrılmayan bölgelerde kömür var ise bunları tıraşlayarak aşağı iner ve cebe girer. Bu şekilde kesim

işlemi tamamlanmış olur ve tahkimat öteleme işlemine geçilir. Şekil 2.8’de ayak başından bir görünüm verilmiştir.



Şekil 2.8 -250/-300 mekanize ayak başı görünümü

Tahkimat öteleme işlemi de aşağıdan yukarı doğru yapılır. En dipteki kalkan tahkimata 1 no’lu kalkan tahkimat diyecek olursak 1’den yukarı doğru ayak başına kadar bütün kalkan tahkimatlar arına doğru ötelenir. Şayet bu esnada tavanda herhangi bir çekme olayı olduysa ve tavanda bir açıklık oluştuysa, bu boşluk ayak içerisine indirilen ağaç malzeme ile tahkim edilerek kalkan tahkimatın üzeri beslenir ve tahkimat tavanının olabildiğince düz olmasına özen gösterilir. Bu şekilde ayak içerisindeki bütün kalkan tahkimatlar ötelenir ve sıra ayak başındaki saptırma makarasının ötelenmesine gelir. Üzerinde elektrik kablosu, su hortumu, çekme halatı ve emniyet halatının da bulunduğu bu makaranın da arına doğru ötelenmesi işlemi bittikten sonra öteleme işlemi sona erer ve tekrar kesim işlemine başlanır.

Şekil 2.9’da tam mekanize ayak içinden görünüm verilmektedir. Yukarıda anlatılan bu döngü günde ortalama 3 kez tekrarlanır ve sistem bu şekilde çalışmaya devam eder.



Şekil 2.9 -250/-300 mekanize ayak içi görünüm

3. AMASRA TAŞKÖMÜRÜ İŞLETMESİ GENEL BİLGİLER

Türkiye Taşkömürü Kurumu (TTK) Genel Müdürlüğü'ne bağlı olarak faaliyette bulunan ATİM, kurumun merkezinin bulunduğu Zonguldak ili dışındaki tek müessesesidir. İmtiyaz sahaları, Bartın ve Kastamonu illeri sınırları içinde olup merkezi Bartın ili Amasra ilçesindedir. İşletmenin kuruluş amacı; devletin enerji politikasına uygun olarak, taşkömürü rezervlerinin en iyi şekilde değerlendirilmesi, ekonomiye kazandırılmasıdır. Yenilenemeyen enerji kaynaklarından olan kömürün gelecekte de önemini koruyacağı görülmektedir. Dünya ile beraber ülkemizin kömüre olan ihtiyacı önümüzdeki yıllarda daha da artacaktır (Tural 2019). ATİM için yer bulduru haritası Şekil 3.1'de verilmiştir.



Şekil 3.1 Amasra Taşkömürü İşletmesi yer bulduru haritası

Türkiye'nin toplam taşkömürü rezervi 1,5 milyar ton civarındadır. Bu rezervin yaklaşık 620 milyon tonu Amasra Havzası'nda bulunmaktadır. Koklaşmayan ve zayıf koklaşan damarlardan oluşan kömürler, koklaşır kömürlere karıştırılarak demir çelik sanayisinde kullanılabilir. Ayrıca sanayinin diğer kollarında ve ısıtma amaçlı olarak da

kullanılmaktadır.

3.1 Amasra Taşkömürü İşletmesi Tarihçesi

Amasra Havzasındaki ilk üretim faaliyetleri 1848 yılında başlamıştır. Özel kişilerce yapılan bu faaliyetler, havzanın devlete geçtiği 1940 yılına kadar devam etmiştir. Saha 1953 yılında Ereğli Kömür İşletmesi'ne (EKİ) devredilmiştir ve Tarlaağzı köyünde devlet sermayesi ile kurulan ilk ocak faaliyete başlamıştır.

Halen üretim çalışmalarının yapıldığı Demirci deresindeki ocaklarda ilk galerinin sürülmesine 1965 yılında başlanmış ve Ocak 1968 yılına kadar Kozlu Bölgesi Hazırlık Ocağı; 1968–1973 yılları arasında da “Müstakil Başmühendislik” olarak devam etmiştir. 18.02.1973 tarihinde EKİ Müessesesi'ne bağlı “Amasra Bölge Müdürlüğü”; 01.07.1979 tarihinde TTK'ye bağlı “Batı Karadeniz Kömür İşletmesi” (BKİ) haline dönüştürülmüş, 07.04.1981 tarihinde ise yeniden EKİ'ye bağlı Bölge Müdürlüğü olmuştur. 1984 yılbaşından itibaren TTK'ye bağlanarak ATİM tüzel kimliği alınmıştır (Tural 2019).

05.04.1994 kararları ile Müessesenin 1994 yılı sonuna kadar özelleştirilmesi veya kapatılması kararlaştırılmıştır. Ancak bu kararlar uygulanmamış, Müessese statüsü İşletme Müdürlüğüne dönüştürülmüş ve Eylül 1994'den itibaren “Amasra İşletme Müdürlüğü” statüsünde TTK Üzülmüş Taşkömürü İşletme Müessesesi'ne bağlanmıştır. Üç yılı aşkın süren işletme döneminden sonra TTK Yönetim Kurulunun 24.07.1997 tarih ve 16/192 sayılı kararı ile 01.01.1998 tarihinden itibaren “Amasra Taşkömürü İşletmesi Müessesesi” olarak yeniden tüzel kişilik kazanılmıştır (İnt. Kyn. 1).

3.2 Amasra Taşkömürü İşletmesi Sınırları

ATİM sahası, yaklaşık 49 km²'lik bir alana sahip olup batıda Tarlaağzı köyü; doğuda Abbasköy, Saraydüzü, Karainler köyleri; kuzeyde Karadeniz ve güneyde Bartın Merkez ilçe ile sınırlanmıştır. Bu sahanın Amasra–A Sahası olarak anılan kuzeydeki bölümü 7,6 km²'lik bir alan olup; bu sahada istihsal ve hazırlık çalışmaları Müessese tarafından yürütülmektedir. İşletme sınırları Şekil 3.2'de verilmiştir.



Şekil 3.2 Amasra Taşkömürü İşletmesi işletme sınırı haritası

Amasra–A Sahası dışında kalan bölümü, Amasra–B Sahasını oluşturmaktadır. Amasra–B Sahasının rödovans karşılığı işletmeciliği, Amasra–A Sahasının –400 kotunun altında kalan kısmı ile beraber yapılan ihale sonucu 15.04.2005 tarihinde imzalanan bir sözleşme ile HEMA Endüstri A.Ş. firmasına verilmiştir (İnt. Kyn. 2).

49,5 km²'den oluşan 72582 sicil nolu ruhsat sahası 7,6 km²'si Amasra–A Sahası ve 41,8 km²'si Amasra–B Sahası şeklinde bölünerek 86970 sicil no ile ruhsatlandırılmıştır. 72582 sicil nolu alanından bölünmek suretiyle oluşturulan 86970 sicil sayılı ruhsatın sahada rödovans usulü faaliyetlerini sürdüren HEMA Endüstri A.Ş.'ye mevcut rödovans sözleşmesi hükümleri saklı kalmak kaydıyla ruhsatın devrinin yapılması 01.10.2019'da kabul edilmiştir. Sürecin devamında 86970 sicil sayılı Amasra–B Sahası ruhsatı, HEMA Endüstri A.Ş.'ye, 11.02.2020 tarihli “Devir Zaptı” ile MAPEG nezdinde imzalanarak devredilmiştir.

Müesseseye bağlı Bartın-Kurucaşile’de 1 adet saha ile Kastamonu sınırları içinde bulunan

3 adet saha ile birlikte Bartın–Yeşilkaya’da 1 adet kil sahası ve Amasra–Bozköy’de 1 adet kuvars kumu sahası rödovans karşılığı ihale edilmiş olup sahada yüklenici firmaların çalışmaları kapsamında üretim ve sevkiyat devam etmektedir.

3.3 Amasra Taşkömürü İşletmesi Genel Jeolojisi

ATİM çalışma sahasında, havzanın en önemli maden yatağı olan taşkömürü içeren formasyonları bulunduran Yılanlı Formasyonu bulunmaktadır. Kireçtaşı, dolomitik kireçtaşı ve dolomit içeren bu formasyon, Paleozoyik yaşlı kayaçların en eskisidir ve Devoniyen-Vizeen (Karbonifer) aralığında oluşmuştur. Yılanlı Formasyonu üzerinde ince kömür damarları içeren Alacağzı Formasyonu gelmektedir. Alacağzı Formasyonu üzerine bol kömürlü iki formasyon gelmektedir. Bunlar Kozlu ve Karadon Formasyonlarıdır. Sahada Paleozoyik döneme ait Çakraz Formasyonu, Karadon Formasyonu üzerindedir. Formasyonlar genel olarak konglomera, kumtaşı, kiltası, silttaşı ve kömür içermektedir.

Bu formasyonlardan sonra Mesozoyik yaşlı formasyonlar başlamaktadır. Sırasıyla Zonguldak, Kilimli, Cemaller, Başköy, Dinlence, İkse, Kazpınar ve Alaplı Formasyonları yer almaktadır. Bu formasyonlar kumlu ve marnlı kireçtaşlarından, andezit, tuf ve marndan oluşmaktadır.

Mezozoyik yaşlı kayaçlar üzerinde Senozoyik yaşlı birimler bulunmaktadır. Kuaterner yaşlı bu birimler, yamaç molozu ve alüvyonlar içermektedir. ATİM çalışma sahası, Hersiniyen ve Alp orojenezleri etkisinde kıvrılarak ve kırılmalar geçirerek şimdiki yapısına ulaşmıştır (Haliloğlu 2000).

ATİM çalışma sahasının da içinde bulunduğu Bartın-Amasra Havzasının genelleştirilmiş stratigrafik kesiti Şekil 3.3’te verilmiştir.

Üst Sistem	Sistem	Seri	Kat	Formasyon	Kalınlık (m)	Sütun Kesiti	Litoloji ve Açıklama	Orojenik Fazlar
Senozoik	Kuvaterner						Alüvyon yamaç molozu	Alp Orojenezi
Mesozoik	Kretase	Üst Kretase	Meastrihtiyen	Alaplı	15-200		Marnlı kireçtaşı	
			Kampaniyen	Kazpınar	60-200		Andezit	
			İkse	İkse	15-130		Marnlı kireçtaşı	
			Santoniyen	Dinlence	30-350		Aglomera , andezitik tüfit	
			Koniasiyen	Dinlence	30-350		Aglomera , andezitik tüfit	
			Turoniyen	Başköy	5-25		Marnlı kireçtaşı	
			Senomaniyen	Cemaller	5-350		Kumlu kireçtaşı	
		Alt Kretase	Apsiyen	Kilimli	0-300		Fliş	
			Barremiyen	Zonguldak	35-500		Kireçtaşı	
Paleozoik	Permian			Çakraz	35-350		Konglomera kumtaşı silttaşı	Hersiniyen Orojenezi
	Karbonifer	Üst Karbonifer	Westfaliye BCD	Karadon	200-780		Konglomera kumtaşı silttaşı kömür	
			Westfaliye A	Kozlu	100-1000		Konglomera kumtaşı kilttaşı kömür	
			Namuriyen	Alacağzı	100-570		Kumtaşı kilttaşı kömür	
	Alt Karbonifer		Vizeen	Yılanlı			Dolomitik kireçtaşı	

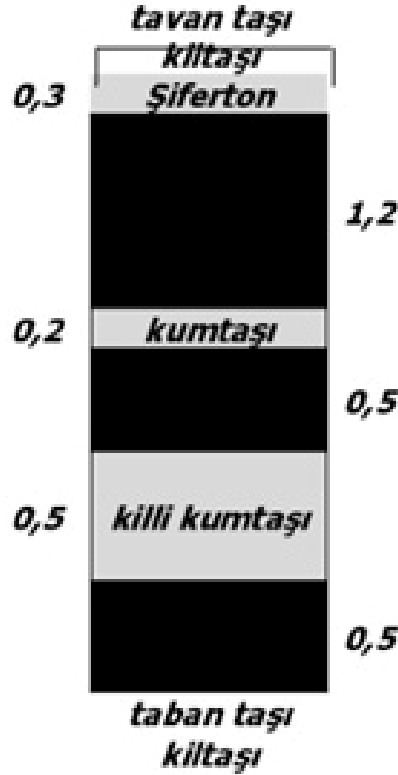
Şekil 3.3 Bartın-Amasra Havzası genelleştirilmiş stratigrafik kesiti (Haliloğlu 2000)

ATİM sahasında dört ekonomik kömür damarı işletilmektedir. Kömür damarlarının üstten alta doğru sıralaması ve kalınlıkları Çizelge 3.1’de verilmiştir.

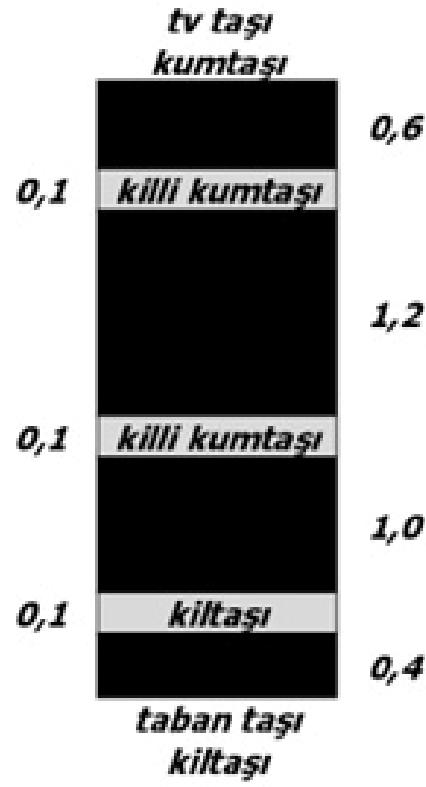
Çizelge 3.1 ATİM kömür damarları ve damarlar arası kalınlık değerleri

Kömür Damarı	Damar ve Ara Kalınlık
Tavan	1,65 – 3,35 m (Damar Kalınlığı) 30 m (İki Damar Arası Kalınlık)
Kalın	1,50 – 3,70 m (Damar Kalınlığı) 15 m (İki Damar Arası Kalınlık)
Taşlı	1,50 – 3,20 m (Damar Kalınlığı) 25 m (İki Damar Arası Kalınlık)
Alt taşlı	1,60 – 3,50 m (Damar Kalınlığı)

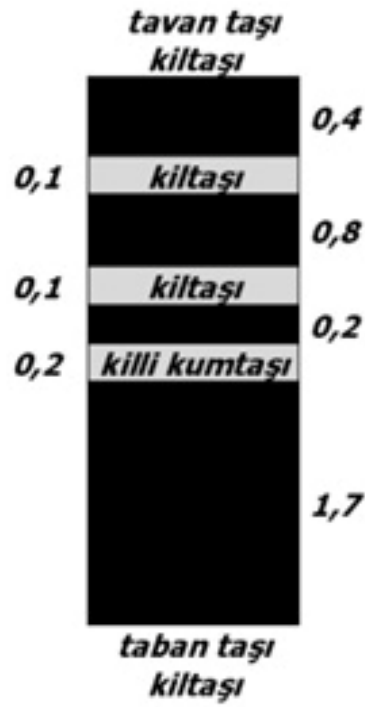
Çizelge 3.1’de bulunan kömür damarlarının stamları ve damar içi kalınlık değerleri tavan damar için Şekil 3.4, kalın damar için Şekil 3.5, taşlı damar için Şekil 3.6 ve alt taşlı damar için Şekil 3.7’de verilmiştir.



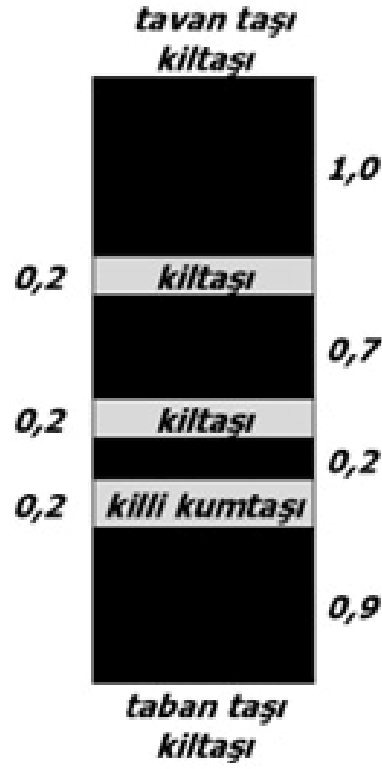
Şekil 3.4 Tavan damarın stampları ve kömür kalınlıkları



Şekil 3.5 Kalın damarın stampı ve kömür kalınlıkları



Şekil 3.6 Taşlı damarın stampı ve kömür kalınlıkları



Şekil 3.7 Alt taşlı damarın stampı ve kömür kalınlıkları

3.4 Amasra Taşkömürü İşletmesi Genel Durumu

Çalışmaların devlet tarafından plan ve projelere bağlı başlatıldığı tarihten itibaren, bugüne kadar +167/–350 kotları arasında 10 ana kat teşkil ettirilmiştir. +167, +100, +40, ±0, –30, –100 ve –175 katlarında üretim faaliyetleri bitirilmiş; –250, –300 ve –350 katları arasında üretim ve hazırlık çalışmaları devam etmektedir. –350 katı ve –325 kuyu dibi dik damarlar lağımı “–250/–350 Kat Hazırlığı” projesine göre açılmakta olup, proje kapsamında –250, –300, –325 ve –350 katları arasında üretim ve hazırlık çalışmaları devam etmektedir.

Amasra ocaklarında 2022 yılı Aralık ayı sonuna kadar yatırımlardan 53.260 m, cari hesaplardan 12.796 m olmak üzere, toplam 66.056 m galeri sürülmüştür. Amasra Müessesesinde 14/10/2022 tarihinde P III vardiyasında meydana gelen grizu patlaması nedeniyle üretim ve hazırlık faaliyetlerine ara verilmiştir.

3.5 Amasra Taşkömürü İşletmesi Ocak Kaynak Durumu

Amasra–Bartın Sahasındaki ilk etüt ve arama çalışmaları 1936 yılında başlamıştır. 1941 yılında Tarlaağzı köyü civarında ilk sığ sondajlar yapılmıştır. 1956 yılında bölgede MTA’nın sondaj çalışmaları başlamış olup 1999 yılına kadar toplam 152 adet sondaj yapılmıştır. Amasra–Bartın sahasındaki rezervler için daha kapsamlı çalışmalara 1984 yılından itibaren başlanmıştır. Bu tarihten sonra, çeşitli kamu ve özel kuruluşlar tarafından sırasıyla 1986, 1992, 2003, 2017 ve 2020 yıllarında kaynak ve rezerv belirleme çalışmaları yapılmıştır (İnt. Kyn. 2). Tüm yapılan çalışmalar sonucunda 2020 yılı itibariyle ATİM sahasındaki son kaynak durumu Çizelge 3.2’de verilmiştir.

Çizelge 3.2 ATİM ocak kaynak durumu

Kaynak	Amasra A-Sahası (ton)	Amasra B-Sahası (ton)	Amasra Toplam (ton)
Hazır	385.000	-	385.000
Görünür	5.605.000	395.954.757	401.559.757
Muhtemel	7.690.000	152.661.950	160.351.950
Mümkün	56.619.859	2.192.919	58.812.778
TOPLAM	70.299.859	550.809.626	621.109.485

Amasra havzasının kömür damarları,

’nın diğer bölgelerindeki damarların aksine koklaşma özelliği olmayan ya da çok zayıf koklaşır kömür damarlarından oluşmaktadır. Bununla beraber koklaşabilen diğer Zonguldak kömürleri ile karıştırılarak demir çelik sanayisinde kullanılabilir. Amasra kömürleri, siyah renkli, mat ve parlak zonlar gösteren sert, yer yer killi, düzensiz kırıklı bir yapı gösterir. Yangına müsait olan bu damarlar, TTK Havzasının gerek parça gerekse yerinde dayanım açısından en sert damarlarıdır.

3.6 Amasra Taşkömürü İşletmesi İstihdam Durumu

İşletmede, 2015 yılından bu yana işçi statüsünde yeraltı ve yerüstü kadrolarındaki

alışanların sayıları ile yerüstünde alışan bir işiye karşılık yeraltında alışan işi sayıları izelge 3.3'te verilmiştir.

izelge 3.3 ATİM istihdam durumu

	2015	2016	2017	2018	2019	2020	2021	2022
Pano Ayak Üretim İşisi	213	218	179	180	344	276	254	170
Yeraltı	499	424	395	384	526	514	508	462
Yerüstü	136	125	116	110	96	100	86	80
Toplam	635	549	511	494	622	614	594	542
Yeraltı / Yerüstü	3,67	3,39	3,41	3,49	5,48	5,14	5,91	5,78

3.7 Amasra Taşkömürü İşletmesi Kömür Üretim Miktarları

Yıllar itibariyle tüvenan ve satılabilir üretimleri, program ve fiili durumları izelge 3.4'te gösterilmektedir. ATİM'de, Ekim 2022 tarihinde meydana gelen grizu patlaması nedeniyle üretim ve hazırlık faaliyetlerine ara verilmiştir.

izelge 3.4 ATİM üretim miktarları

YILLAR	Tüvenan Üretim(ton)		Gerçekleşme (%)	Satılabilir Üretim (ton)		Gerçekleşme (%)
	Program	Fiili		Program	Fiili	
2014	252,840	221,243	88	130,935	100,785	77
2015	256,704	194,746	76	132,356	90,527	68
2016	204,180	170,819	84	106,074	83,578	79
2017	197,846	116,437	59	103,730	53,980	52
2018	169,425	100,953	60	84,336	46,841	56
2019	148,694	130,640	88	74,347	51,763	70
2020	166,400	116,080	70	99,632	46,516	47
2021	181,545	125,406	69	108,780	55,981	51
2022	158,500	104,508	66	91,000	41,380	45

4. VERİ MADENCİLİĞİ

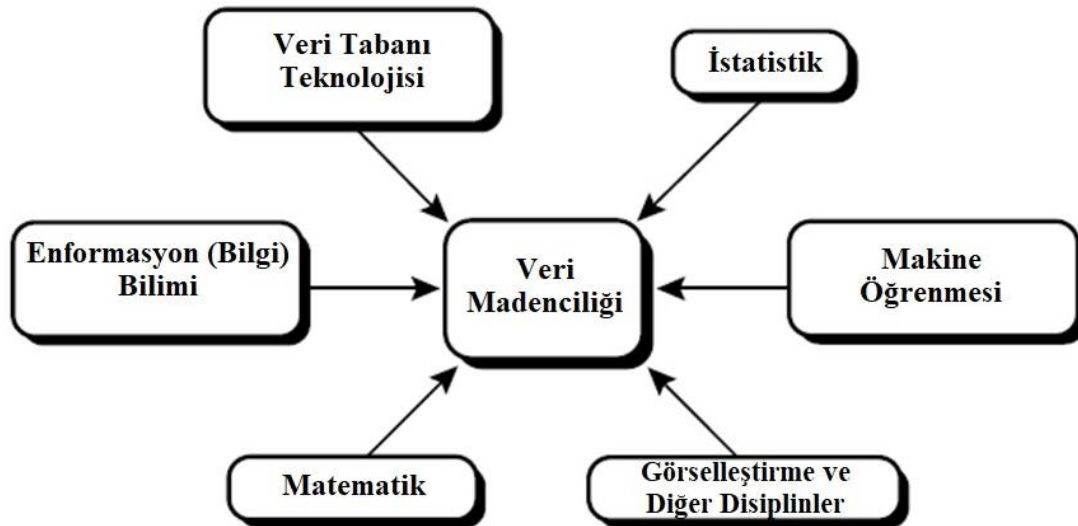
Günümüz dünyasında bilginin kullanımı ve dolayısıyla teknolojik gelişim yadsınamaz durumdadır. İnsan, yüksek standartlarda bir yaşam oluşturmak için daima bilgi ve teknolojiyi kullanmaktadır. Zaten bu bilgi ve teknolojiyi kullanabilen toplumlar daha gelişmiş, daha müreffeh standartlara ulaşabilmektedirler. Teknolojinin gelişimiyle hayatımızın her alanında gerçekleştirdiğimiz işlemlerin ve yaşadığımız olayların kayıt altına alınması, depolanması ve istenildiğinde ulaşılması sağlanmıştır. Bu çeşit bir arşivleme, elde edilen bu verilerin daha sonra tekil olarak kullanımını oldukça kolaylaştırmıştır. Buna karşılık, bu verilerin ilişkisel anlamda değerlendirilerek geçmiş dönem için sonuçların çıkarılması ve ileriye dönük kararların alınabilmesi ayrı bir uzmanlıktır. Günümüzde yapılan oldukça fazla veri içeren analizler, veri biliminde uzmanlaşma ve farklı araçların kullanımını gerektirmektedir. Büyük miktarda veriyi elde etmek ve bu veri topluluğundan bilgi çıkararak kullanabilmek, hayatımızın her alanında değerli hale gelmektedir. Veri madenciliği de bu anlamda fayda sağlayan bir mekanizma görevi üstlenmektedir. Veri madenciliği, büyük miktarda verinin erişilebilir olması ve bu verilerin yararlı bilgi ve enformasyona dönüştürülmesine duyulan ihtiyaç nedeniyle, son yıllarda bilgi endüstrisinde ve bir bütün olarak toplumda büyük ilgi görmüştür.

Basit bir ifadeyle veri madenciliği, büyük miktarda veriden bilgi madenciliği yapmaktır. Aslında bu ifade, klasik madencilikte yapılan faaliyetlerin üretilen madene göre isimlendirildiği düşünüldüğünde (örneğin; altın madenciliği, kömür madenciliği, doğal taş madenciliği gibi) eksik bir ifade olmuştur. Bu nedenle veri madenciliği, daha uygun bir şekilde “verilerden bilgi madenciliği” olarak adlandırılabilirdi. Bu uzun ifade yerine kullanılabilecek “bilgi madenciliği” de bilginin büyük veriden çıkarıldığını ifade etmekte yetersiz kalmaktadır. Dolayısıyla, “veri madenciliği” ifadesi çok daha beğenilen, rağbet gören, kullanılan bir seçim olmuştur (Han ve Kamber 2006).

Veri madenciliği (verilerden bilgi keşfi) büyük miktarda veriden ilginç (önemsiz olmayan, örtülü, önceden bilinmeyen ve potansiyel olarak yararlı) sonuçların, örüntülerin ve bilginin çıkarılmasıdır. Veri tabanlarında bilgi keşfi (madenciliği) [knowledge discovery (mining) in databases – KDD], bilgi çıkarımı (knowledge extraction),

veri/örüntü analizi (data/pattern analysis), veri arkeolojisi (data archeology), veri tarama (data dredging), bilgi hasadı (information harvesting) gibi terimler de veri madenciliğine benzer anlamlara sahiptir.

Veri madenciliği, büyük ölçekli verilerden açıklayıcı, anlaşılır ve tahmine dayalı modellerin yanı sıra anlayışlı, ilginç ve yeni modelleri keşfetme sürecidir. Mevcut problemleri çözmek, kritik kararları almak veya geleceğe yönelik tahminleri yapmak için gerekli olan bilgileri elde etmeye yarayan bir araçtır. Ortaya çıkarılması hedeflenen bilgiler; üstü kapalı, çok net olmayan, önceden bilinmeyen, daha önce keşfedilmemiş; ancak potansiyel olarak kullanışlı, anlamlı ve kritik bilgilerdir. Diğer bir deyişle, veri madenciliği verilerin, yeni ve potansiyel olarak yararlı bilgi bulma amaçlı analiz sürecidir. Bu süreç, bulunması zor örüntülerin ortaya çıkarılması için matematiksel araçların kullanımını içerir. Veri madenciliğinin elde edilen verilerin içinde bulunan bilgileri insana gereksinim duymadan, kendi başına çıkarması mümkün değildir. Kullanıcı tarafından hiçbir işlem yapılmadan, veriler üzerinden problemlere otomatik yanıt veren herhangi bir veri madenciliği aracı bulunmamaktadır. Bu anlamda birçok bilim dalı ile ilişkili şekilde işlemler yürütülmektedir (Han ve Kamber 2006). Şekil 4.1’de veri madenciliğinin etkileşim içinde olduğu bazı bilim dalları verilmiştir.



Şekil 4.1 Veri madenciliğinin etkileşim içinde olduğu bazı bilim dalları (Han ve Kamber 2006)

4.1 Veri Madenciliği Tarihçesi

Önceki yıllarda, verileri depolayabilen ve analiz edebilen mevcut bilgi sistemlerinin eksikliği nedeniyle çeşitli sıkıntılar yaşanmaktaydı. Ancak, günümüzde ve çağımızda, verilere inceleme ve örüntüleri çıkarma konusunda yeni bir model ortaya çıkmış ve birçok alanda ilgililerin kullanımına sunulmuştur. 2000’li yıllardan önce, giderek artan miktarda veri, elektronik olarak depolanmaya başlamıştır. Bu hacim giderek artmış ve gelecekte de artmaya devam edecektir. Son yirmi yıla kadar toplanan veri hacminin artmasına rağmen, bu verilerden çok azından istenilen şekilde ve değerde yararlanılabilmektedir. Bunun nedeni, verileri tam olarak analiz etme ve ortaya çıkabilecek temel örüntüleri ayırt etme işleminin zorluğuydu (Aldana 2000). Veri madenciliğinin gelişimine genel bir bakış Çizelge 4.1’de verilmiştir.

Çizelge 4.1 Veri madenciliği kronolojik gelişimi (Aldana 2000’den uyarlanmıştır)

Gelişim Aşaması	Cevaplanan Soru Örneği	Kullanılabilen Teknolojiler	Özellikler
Veri Toplama (1960lar)	Geçen 5 yılda kar ne kadardı?	Bilgisayarlar, teypler, diskler	Geriye dönük, statik veri sunumu
Veri Erişimi (1980ler)	Geçen ay birim satışlar ne kadardı?	İlişkisel veri tabanları, Yapılandırılmış sorgulama dili (SQL)	Geriye dönük, dinamik veri sunumu
Veri Ambarlama ve Karar Destek Sistemleri (1990lar)	Geçen ay birim satışlar ne kadardı?	Çevrimiçi analitik işleme (OLAP), çok boyutlu veri tabanları, veri ambarları	Çoklu seviyelerde geriye dönük, dinamik veri sunumu
Veri Madenciliği (Günümüzde Gelişmekte Olan)	Gelecek ay birim satışlar muhtemelen ne olur? Niçin?	İleri düzeyde algoritmalar, çok işlemcili bilgisayarlar, büyük veri tabanları	İleriye dönük, önleyici bilgi sunumu

Veri madenciliği teknikleri, uzun bir araştırma ve ürün geliştirme sürecinin sonucudur. Bu gelişim, verilerin ilk kez bilgisayarlarda depolanmasıyla başlar ve veriye erişimdeki gelişmelerle devam eder. Veri madenciliği algoritmaları en az 40 yıldır var olan teknikleri bünyesinde barındırmaktadır; ancak son zamanlarda eski istatistiksel yöntemlerden sürekli olarak daha iyi performans gösteren olgun, güvenilir, anlaşılabilir araçlar olarak uygulanmaktadır. Veriden bilgiye doğru gelişimde, her yeni adım bir öncekinin üzerine

inşa edilmiştir. Çizelge 4.1’de verilen dört adım devrim niteliğindedir; çünkü yeni problemlerin doğru ve hızlı bir şekilde yanıtlanmasına olanak tanımaktadır. Veri madenciliği teknolojisinin temel bileşenleri, istatistik, yapay zeka ve makine öğrenimi gibi araştırma alanlarında onlarca yıldır geliştirilmektedir. Günümüzde bu tekniklerin olgunluğu, yüksek performanslı ilişkisel veri tabanı motorları ve geniş veri entegrasyonu çabalarıyla birleştiğinde, bu teknolojileri mevcut veri ambarı ortamları için pratik hale getirmektedir (Aldana 2000).

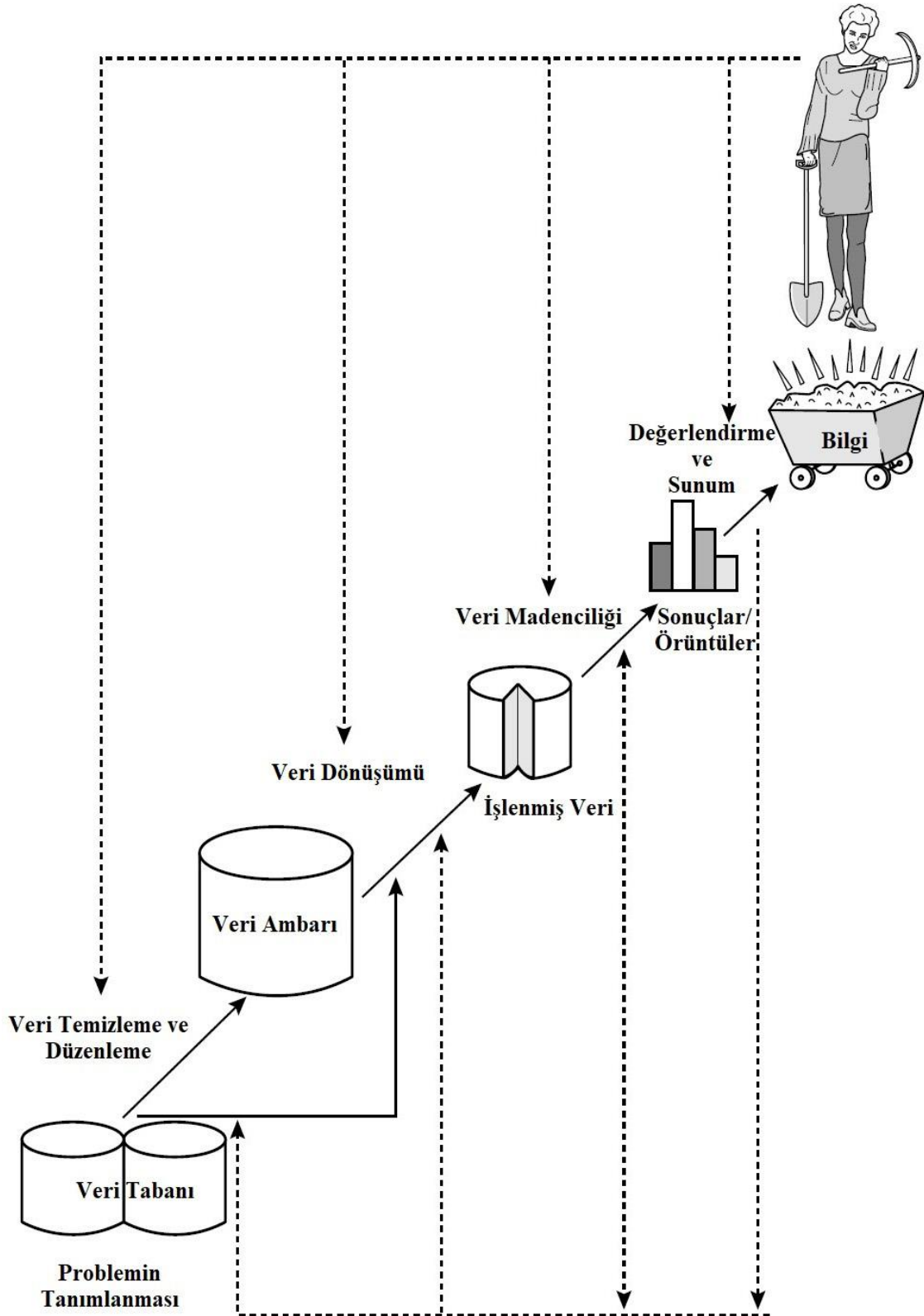
4.2 Veri Madenciliği ve Bilgi Keşfi Süreci

Veri madenciliği gizli kalmış bilgilerin ortaya çıkarılma, keşfedilme sürecidir. Bilgi keşfi sürecinde veri madenciliği önemli bir rol oynamaktadır. Bu sürecin her aşaması önemlidir. Çünkü, birbirini takip eden aşamalardan oluşur ve her aşamasının çıktısı bir sonraki aşamanın girdi verisini oluşturmaktadır.

Veri madenciliğiyle bilgi keşfinde standart bir süreç vardır. Bu süreç, The Cross-Industry Standard Process for Data Mining (CRISP-DM) oluşumu tarafından ortaya konmuş ve uluslararası düzeyde standardı kabul edilmiştir. CRISP-DM’ye göre bu süreç, aşağıdaki altı aşamadan oluşmaktadır (İnt. Kyn. 3):

- a) Problemi ve problem ortamını anlama
- b) Veriyi anlama
- c) Veriyi hazırlama
- d) Modelleme
- e) Değerlendirme
- f) Çıktıları sunma

Yukarıdaki standart süreç düşünüldüğünde, bilgi keşfinin bir aşaması olarak veri madenciliği aşamaları Şekil 4.2’de verilmiştir.



Şekil 4.2 Bilgi keşfi sürecinde bir adım olarak veri madenciliği aşamaları (Han ve Kamber 2006)

a. Problemin Tanımlanması: Bilginin keşfi ve veri madenciliğinin başarılı olabilmesinin en önemli adımı, var olan problemin doğru bir şekilde tanımlanmasıdır. Ulaşılmak istenen amaç, problemle tam anlamıyla örtüşmeli ve problem açık bir şekilde tanımlanmış olması gerekir. Problemin bu şekilde belirlenip çözüme ulaştırılabilmesi için, öncelikle problem ilgili bilgi ve verilerle eşleştirilmelidir. Veri madenciliğiyle uğraşan analist, karşısına çıkan ve çözülmeyi bekleyen her probleme ve sektöre uzman gibi yaklaşamaz. Burada ilgili sektörlerden uzmanlar tarafından desteklenmelidir. İlgili sektörlerdeki uzmanlarla iş birliği içinde yapılarak problemin ve veri tabanının belirlenmesi bilgi keşfi işleminin önemli bir parçasıdır.

b. Veri Temizleme ve Düzenleme: Veri temizleme ve düzenleme olarak adlandırabileceğimiz verinin hazırlanma aşaması, veri madenciliğinde önemli aşamalardan biridir. Bu aşama ne kadar doğru yapılırsa, ileriki aşamalarda kurulacak modellerde sorunların çıkma olasılığı azalmış olur. Yapılacak uygulamalar için seçilen veriler istenilen özelliklere uygun olmayabilir. Eksik veya aykırı, gürültülü, tutarsız veriler veri tabanında yer alabilir. Verilerin analiz aşamasından önce değerlendirilerek, istatistiksel anlamda sağlıklı bir hale getirilmesi gerekmektedir. Bu amaç için, veriler üzerinde aşağıdaki işlemler uygulanmaktadır.

i. Veri Seçme: Yapılacak olan analiz ile ilgili olan verilerin belirlenmesi işlemidir. Tanımlanan problemle ilgili verilerin toplanması ve bu verilerin elde edileceği kaynakların tespit edilmesidir.

ii. Veri Temizleme: Veri tabanında yer alan tutarsız ve hatalı verilere gürültü; var olan uç verilere de aykırı veriler denir. Veriler toplanma biçimine ve elde edildikleri kaynaklara bağlı olarak gürültülü, tutarsız, eksik özelliklere sahip olabilirler. Sonuçlar üzerinde bu özelliklerin etkilerinin azaltılması gerekmektedir. Veri temizleme işlemi, gürültülü ve tutarsız verilerin çıkarılması işlemidir.

iii. Veri Bütünleştirme: Birçok veri kaynağının birleştirilmesi işlemidir. Farklı kaynaklardan gelen veriler tek bir veri ambarında toplanabilir.

iv. Veri Azaltma: Veri hacmi çok büyük olduğunda, uygulanacak veri

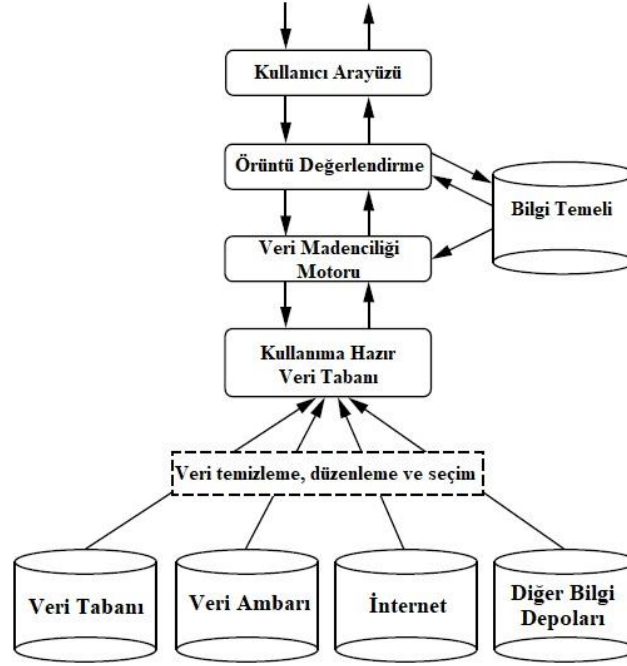
madenciliği algoritmalarından sonuç elde etmek oldukça fazla zaman alabilmektedir. Veri azaltma işlemi, aynı veri sonuçları elde edebilecek şekilde veri miktarının azaltılması, bir bakıma örneklem hazırlanmasıdır.

- c. **Veri Dönüşümü:** Verinin, veri madenciliği yöntemine uygun hale dönüşümünün gerçekleştirilmesidir. Veri madenciliği uygulamasında kullanılacak teknik ve modeller her veri grubu ile çalışmayabilir. Bazı algoritmalar yalnızca sayısal verilerle çalışırken, bazıları da yalnızca sözel verilerle çalışabilir.
- d. **Veri Madenciliği:** Verilerdeki örüntülerin belirlenmesi için veri madenciliği yöntemlerinin uygulanmasıdır. Hazır hale gelen verilere veri madenciliği algoritmaları uygulanır. Uygulanacak algoritma, problemin tanımına uygun olarak seçilir ve çözümlenecek veri kümesine uygulanır.
- e. **Değerlendirme ve Sunum:** Bazı ölçütlere göre elde edilmiş ilginç örüntülerin bulunması ve değerlendirilmesi işlemidir. Algoritmaların uygulanması sonucu, elde edilen sonuçların değerlendirildiği ve anlaşıldığı bölümdür. Bu aşama veri madenciliğinin, dolayısıyla bilgi keşfinin en önemli aşamalarından biridir. Elde edilen sonuçların doğru bir şekilde yorumlanıp, grafiksel metinlere aktarılması gerekir. Böylece kullanılabilir bilginin, son kullanıcılara doğru ve yararlı bir şekilde iletilmesi sağlanır. Ayrıca, elde edilen verinin ne kadar ilginç olduğu ve bilginin etki gücü önemlidir.

Görüldüğü gibi, veri madenciliği tüm süreçte yalnızca bir adımdır; ancak değerlendirme için gizli örüntüleri ortaya çıkarması, verilerden bilgi keşfi sürecinin en önemli aşaması olmasını sağlamaktadır. Dolayısıyla tüm bu süreç aslında veri madenciliği olarak anılmaktadır.

Yukarıda belirtilen aşamalar çerçevesinde tipik bir veri madenciliği sisteminin mimarisi aşağıdaki ana bileşenlere sahip olabilir (Şekil 4.3). İlk etapta veri, veri tabanı, veri ambarı, internet ve diğer bilgi depolarında toplanır. Bu aşamada, elde edilen veriler üzerinde gerekli ise veri temizleme, düzenleme, bütünleştirme ve seçme gibi işlemler yapılabilir. Bu işlemler neticesinde kullanılabilir veri tabanı oluşur. Veriler, veri madenciliği motoru ile çeşitli ilişkisel analizlere tabi tutulur. Örüntü değerlendirme aşamasında, ilginç örüntülerin ortaya çıkarılması söz konusudur. Bilgi temeli, araştırmayı yönlendirmek

veya ortaya çıkan örüntülerin ilginçliğini değerlendirmek için kullanılan alan bilgisidir. Kullanıcı arayüzü ise, veri madenciliği sistemi ile kullanıcı arasındaki bağlantıyı oluşturmaktadır.



Şekil 4.3 Tipik bir veri madenciliği sisteminin mimarisi (Han ve Kamber 2006)

4.3 Veri Madenciliği Yöntemleri

Genel olarak veri madenciliği yöntemleri iki sınıfa ayrılabilir:

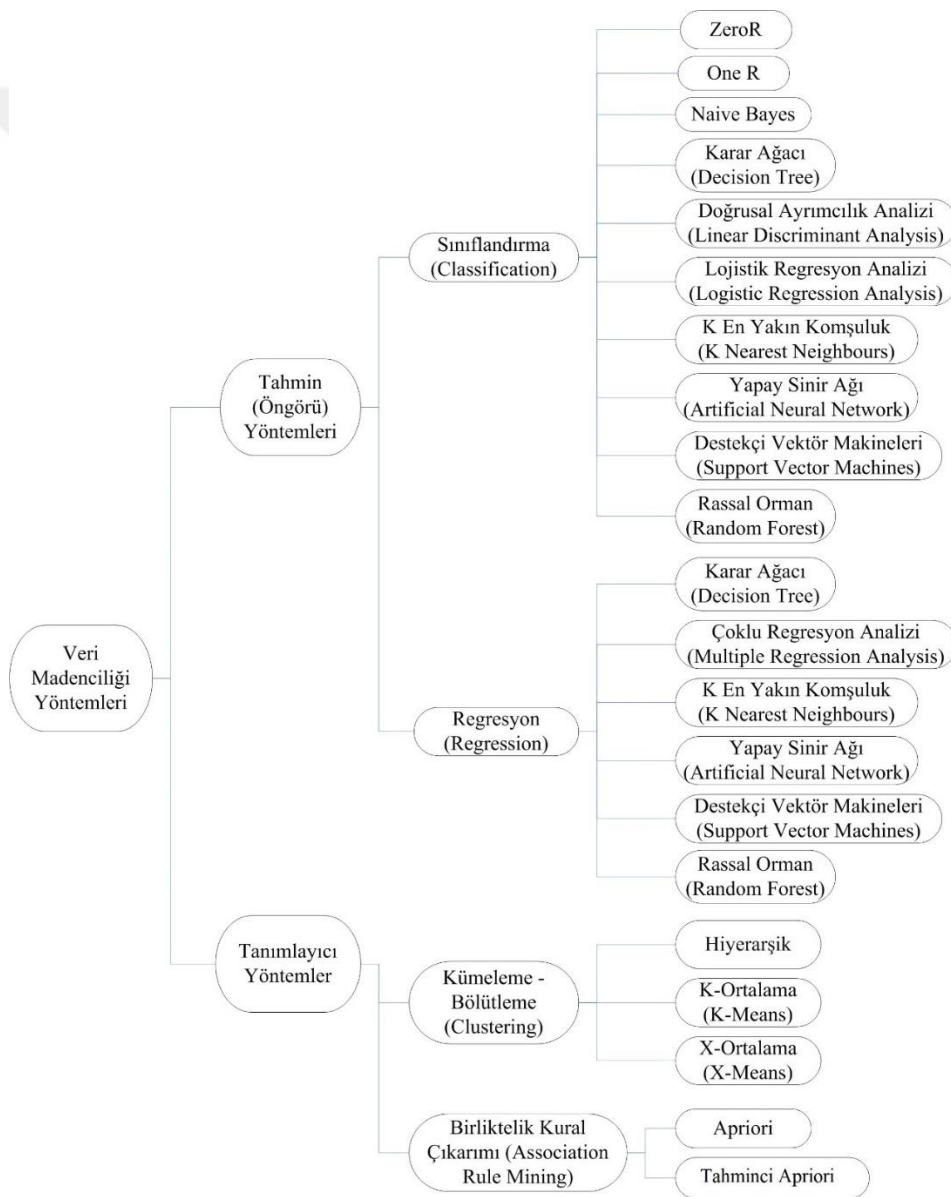
- Tahmin (Öngörü) Yöntemleri (Prediction Methods)
- Tanımlayıcı Yöntemler (Description Methods)

Tahmin (öngörü) yöntemleri, sonuçları bilinen verileri temsil edecek en uygun modeli kurar ve bu modeli kullanarak sonuçları bilinmeyen veri setlerini tahmin etmeye çalışır. Öngörü amacı ile var olan verilerden yorum çıkarılmasını hedefler. Tanımlayıcı yöntemlerde ise karar vermede kullanılacak eldeki verilerin tanımlanması gerçekleştirilir. Amaç, mevcut veriler arasındaki ilişkileri, bağıntıları ve davranışları bulmaktır. Veriyi tanımlayan yorumlanabilir örüntülerin bulunması hedeflenir (Argüden ve Erşahin 2008).

Veri madenciliği yöntemleri, yaptıkları işlemlere göre de temel olarak üç gruba ayrılabilir. Bunlar:

- i. Sınıflandırma ve Regresyon
- ii. Kümeleme
- iii. Birliktelik Kuralları ve Ardışık Zamanlı Örüntüler

İşlevlerine göre yöntemlerden sınıflandırma ve regresyon yöntemleri tahmin (öngörü) yöntemleri arasındadır. Kümeleme, birliktelik kuralları ve ardışık zamanlı örüntüler tanımlayıcı yöntemlerdir. Bu çerçevede veri madenciliği yöntemlerinin ve uygulanabilecek algoritmaların sınıflaması Şekil 4.4'te verilmiştir.



Şekil 4.4 Veri madenciliği yöntemleri ve uygulanabilecek algoritmalar (İnt. Kyn. 4)

Tahmin (öngörü) yöntemlerinden sınıflandırma yönteminde, sınıfı belirlenmiş olan veriler kullanılarak sınıfı belli olmayan veriler sınıflandırılır. Veri madenciliğinde en yaygın kullanılan yöntemler arasındadır. Öncelikle kullanılan veri setine yönelik model oluşturulur, daha sonra bu model temel alınarak hedef verilerin sınıfları tahmin edilir. Regresyon da aynı şekilde, değerleri bilinen değişkenler (bağımsız değişkenler) kullanılarak hedef edilecek değişkenler (bağımlı değişkenler) tahmin edilir. Kümeleme yönteminde temel amaç; karmaşık, heterojen olan veri setinin homojen kümelere ayrılmasıdır. Kullanılan çeşitli algoritmalarla veriler, benzerliklerine göre bir grupta toplanır. Birliktelik kural çıkarımı yönteminde ise veri seti ilişkisel olarak incelenir ve muhtemel ilişkiler belirlenir. Güvenilirlik, ilgi, kanaat gibi ölçütlere göre hedef veriye ilişkin kurallar oluşturulur.

Şekil 4.4'te verildiği gibi, veri madenciliği uygulamalarında tahmine ve tanımlamaya dayalı yapılan işlemlerde kullanılan temel algoritmalar mevcuttur. Bu yöntemler ve algoritmalar kullanıcı dostu çeşitli programlarla (RapidMiner, WEKA, R gibi) uygulanabilmektedir. Bu çalışma kapsamında, veri madenciliği uygulamaları WEKA programı ile gerçekleştirilmiştir.

4.4 WEKA ile Veri Madenciliği

WEKA, son teknoloji makine öğrenimi algoritmaları ve veri ön işleme araçlarından oluşan açık kaynak kodlu bir programdır. Mevcut yöntemlerin, yeni veri kümeleri üzerinde hızlı bir şekilde denenebilmesi için tasarlanmıştır. Girdi verilerinin hazırlanması, öğrenme tablolarının istatistiksel olarak değerlendirilmesi ve girdi verilerinin ve öğrenme sonuçlarının görselleştirilmesi de dahil olmak üzere tüm deneysel veri madenciliği süreci için kapsamlı destek sağlar. Çeşitli öğrenme algoritmalarının yanı sıra, çok çeşitli ön işleme araçları içerir. Bu çeşitli ve kapsamlı araç setine ortak bir arayüz aracılığıyla erişilir, böylece kullanıcılar farklı yöntemleri karşılaştırabilir ve eldeki sorun için en uygun olanları belirleyebilir (Witten vd. 2011).

WEKA, Yeni Zelanda'daki Waikato Üniversitesi'nde geliştirilmiştir; adı **Waikato Environment for Knowledge Analysis (WEKA)** kelimelerinin baş harflerinden

oluşmaktadır (ayrıca weka, sadece Yeni Zelanda adalarında bulunan meraklı bir yapıya sahip uçamayan bir kuş türünün ismidir). Sistem Java dilinde yazılmış ve GNU Genel Kamu Lisansı koşulları altında dağıtılmıştır. Neredeyse her platformda çalışır (Linux, Windows ve Macintosh işletim sistemleri). Birçok farklı öğrenme algoritması için tek tip bir arayüzün yanı sıra, ön ve son işleme ve herhangi bir veri kümesindeki öğrenme şemalarının sonucunu değerlendirme yöntemleri sağlar. WEKA programı 3.8.6 sürümünün giriş arayüzü Şekil 4.5’te verilmiştir.



Şekil 4.5 WEKA programı giriş arayüzü

WEKA programında temel veri ön işleme ve sınıflandırma (classify), kümeleme-bölütme (clustering), birliktelik analizi (association rules), eğri uydurma (regression) gibi veri madenciliği işlemleri yapılabilmektedir. Verileri tanımlamak çalışmanın ayrılmaz bir parçasıdır. WEKA birçok veri görselleştirme olanağı ve veri ön işleme aracı sağlamaktadır.

4.4.1 WEKA veri dosyasının hazırlanması

WEKA programında tüm algoritmalar, girdi verilerini genel olarak ARFF (Attribute-Relation File Format) adı verilen bir dosya şeklinde almaktadır. ARFF dosyaları, WEKA programında kullanılmak üzere oluşturulmuş bir dosya tipidir. WEKA programı, Comma-Separated Values (csv) dosyalarını da kullanabilmektedir. Son sürümlerinde “WEKA Package Manager” vasıtasıyla yapılacak eklentiyle Microsoft Excel dosyaları

(xls, xlsx) da doğrudan kullanılabilmekte ve ARFF dosya biçiminde kaydedilebilmektedir. ARFF dosyaları iki farklı bölümden oluşmaktadır. Birinci bölüm başlık bölümü, ikinci bölüm ise veri bölümüdür. Başlık bölümünde % işareti ile başlayan ve WEKA tarafından dikkate alınmayan satırlar bulunmaktadır. Bu satırlar dosya hakkında genel bilgileri, veriler hakkında bilgileri ve özet istatistikleri içermektedir (Şekil 4.6). @ işareti ile başlayan veri bölümü ise veri kümesinin adını (@RELATION), öz niteliklerin listesini ve türlerini (@ATTRIBUTE) ve verileri (@DATA) içermektedir. Şekil 4.7’de WEKA programı içerisindeki örnek bir ARFF dosyasının veri bölümünün görünümü verilmiştir.

```
% 1. Title: Iris Plants Database
%
% 2. Sources:
%   (a) Creator: R.A. Fisher
%   (b) Donor: Michael Marshall (MARSHALL%PLU@io.arc.nasa.gov)
%   (c) Date: July, 1988
%
% 3. Past Usage:
%   - Publications: too many to mention!!! Here are a few.
%   1. Fisher, R.A. "The use of multiple measurements in taxonomic problems"
%     Annual Eugenics, 7, Part II, 179-188 (1936); also in "Contributions
%     to Mathematical Statistics" (John Wiley, NY, 1950).
%   2. Duda, R.O., & Hart, P.E. (1973) Pattern Classification and Scene Analysis.
%     (Q327.D83) John Wiley & Sons. ISBN 0-471-22361-1. See page 218.
%   3. Dasarthy, B.V. (1980) "Nosing Around the Neighborhood: A New System
%     Structure and Classification Rule for Recognition in Partially Exposed
%     Environments". IEEE Transactions on Pattern Analysis and Machine
%     Intelligence, Vol. PAMI-2, No. 1, 67-71.
%   -- Results:
%   -- very low misclassification rates (0% for the setosa class)
%   4. Gates, G.W. (1972) "The Reduced Nearest Neighbor Rule". IEEE
%     Transactions on Information Theory, May 1972, 431-433.
%   -- Results:
%   -- very low misclassification rates again
%   5. See also: 1988 MLC Proceedings, 54-64. Cheeseman et al's AUTOCLASS II
%     conceptual clustering system finds 3 classes in the data.
%
% 4. Relevant Information:
%   --- This is perhaps the best known database to be found in the pattern
%       recognition literature. Fisher's paper is a classic in the field
%       and is referenced frequently to this day. (See Duda & Hart, for
%       example.) The data set contains 3 classes of 50 instances each,
%       where each class refers to a type of iris plant. One class is
%       linearly separable from the other 2; the latter are NOT linearly
%       separable from each other.
%   --- Predicted attribute: class of iris plant.
%   --- This is an exceedingly simple domain.
%
% 5. Number of Instances: 150 (50 in each of three classes)
%
% 6. Number of Attributes: 4 numeric, predictive attributes and the class
%
% 7. Attribute Information:
%   1. sepal length in cm
%   2. sepal width in cm
%   3. petal length in cm
%   4. petal width in cm
%   5. class:
%       -- Iris Setosa
%       -- Iris Versicolour
%       -- Iris Virginica
%
% 8. Missing Attribute Values: None
%
% Summary Statistics:
%
%   Min   Max   Mean   SD   Class Correlation
% sepal length: 4.3 7.9  5.84  0.83  0.7826
% sepal width:  2.0 4.4  3.05  0.43  -0.4194
% petal length:  1.0 6.9  3.76  1.76  0.9490 (high!)
% petal width:  0.1 2.5  1.20  0.76  0.9565 (high!)
%
% 9. Class Distribution: 33.3% for each of 3 classes.
```

Şekil 4.6 WEKA programında bir ARFF dosyası başlık bölümü görünümü

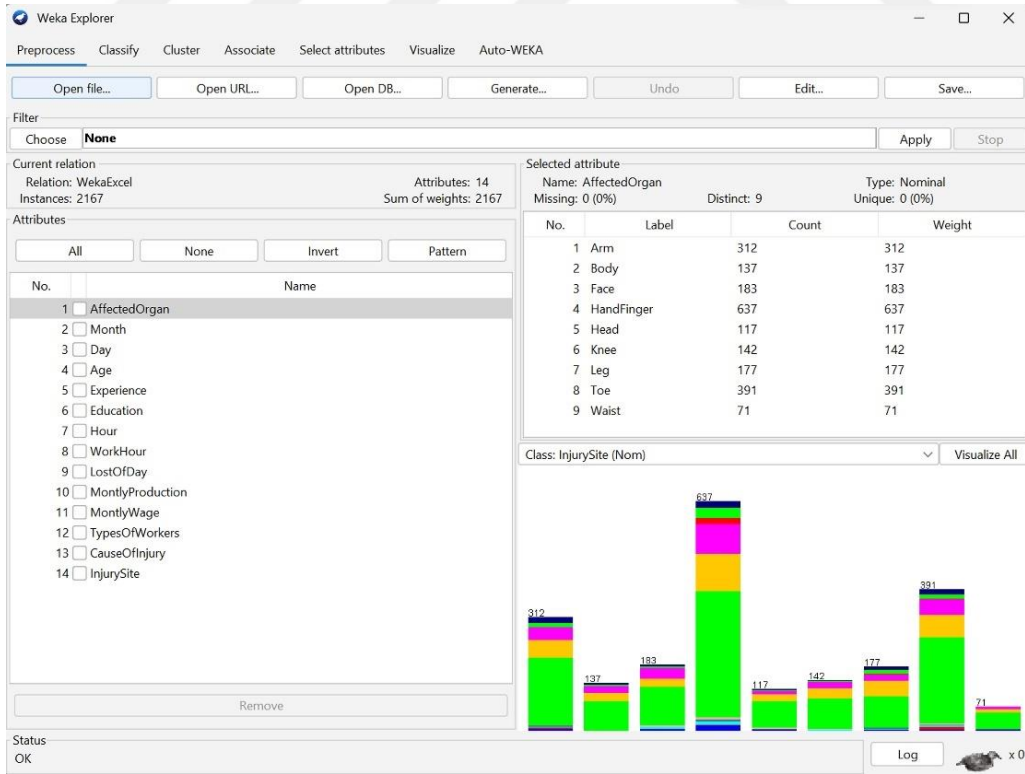
```
@relation weather.symbolic

@attribute outlook {sunny, overcast, rainy}
@attribute temperature {hot, mild, cool}
@attribute humidity {high, normal}
@attribute windy {TRUE, FALSE}
@attribute play {yes, no}

@data
sunny,hot,high,FALSE,no
sunny,hot,high,TRUE,no
overcast,hot,high,FALSE,yes
rainy,mild,high,FALSE,yes
rainy,cool,normal,FALSE,yes
rainy,cool,normal,TRUE,no
overcast,cool,normal,TRUE,yes
sunny,mild,high,FALSE,no
sunny,cool,normal,FALSE,yes
rainy,mild,normal,FALSE,yes
sunny,mild,normal,TRUE,yes
overcast,mild,high,TRUE,yes
overcast,hot,normal,FALSE,yes
rainy,mild,high,TRUE,no
```

Şekil 4.7 WEKA programında bir ARFF dosyası veri bölümü görünümü

ARFF dosyasının büyük bir kısmı örneklerin (satırlar) listesinden oluşur ve her örnek için öz nitelik değerleri virgülle ayrılır. Hazır haldeki ARFF biçimindeki veri dosyası, WEKA programında başlangıç arayüzündeki “Explorer” sekmesinden açılmaktadır (Şekil 4.8).



Şekil 4.8 WEKA Explorer arayüzü

Şekil 4.8, ARFF dosyası yüklendikten sonraki ekranı göstermektedir. Burada, örnek (instances) ve öz nitelik (attributes) sayıları görülmektedir. Bunun yanında, öz nitelik listesi ve veri seti içindeki öz niteliklerin hedef veriye göre dağılımları da yer almaktadır. “Edit” sekmesiyle veri seti görülebilir ve gerekli düzenlemeler yapılabilir. “Explorer” arayüzü “Preprocess” sekmesiyle açılmaktadır. Bu sekme, veri seti üzerinde gerekli ön işleme işlemlerinin yapıldığı alandır. Bunu takip eden Classify, Cluster, Associate, Select Attributes gibi sekmelerle veri madenciliğinin diğer uygulamaları kullanılabilir.

ARFF dosyalarında öz niteliklerin veri tipleri aşağıdaki şekilde sınıflandırılabilir;

- i. Numeric
 - Real (Ondalıklı sayılar, reel sayılar kullanılacaksa)
 - Integer (Tam sayılar kullanılacaksa)
- ii. Date (Tarih tanımlamak için)
- iii. String (Metinsel verileriniz varsa)
- iv. Nominal (Bir kümenin elamanı şeklinde ise, o kümeye aitlik belirtiliyorsa)

ARFF dosyalarında dikkat edilmesi gerekli bazı hususlar bulunmaktadır. Bunlar;

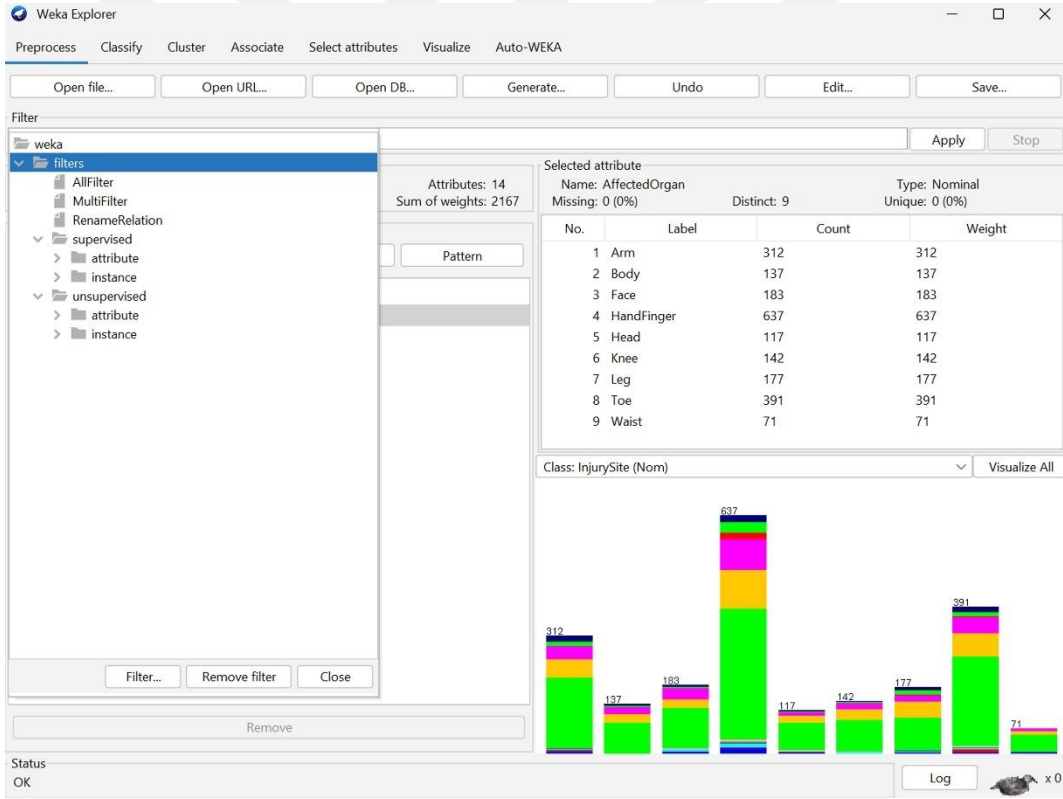
- i. Türkçe karakterleri kullanmamakta fayda vardır.
- ii. İki kelime arasında boşluk bırakılmaz veya bu kelimeler tırnak (‘.....’) içerisine alınmalıdır.
- iii. Veri bilminde değeri olmayan özellikler ‘null’ şeklinde ifade edilir. ARFF dosyalarında herhangi bir öz niteliğin değeri yoksa bu ‘?’ işareti ile gösterilir. Örneğin;

```
@data
Marble,65,7,55
Limestone,75,8,60
Granite,100,11,68
Travertine,55,?,47
```

4.4.2 WEKA ile veri ön işleme

WEKA “Preprocess” ekranında “Filter” özelliğini kullanarak veriyi işlememiz mümkündür (Şekil 4.9). Filtre özelliği ikiye ayrılır.

- a. Denetimli (Supervised)
 - i. Öz nitelik (Attribute)
 - ii. Örnek (Instance)
- b. Denetimsiz (Unsupervised)
 - i. Öz nitelik (Attribute)
 - ii. Örnek (Instance)



Şekil 4.9 WEKA filtre özelliği görünümü

Denetimli filtreler sınıf değerini dikkate alırken, denetimsiz filtreler dikkate almaz. Kullanılacak çeşitli filtre özellikleriyle veri üzerindeki tekrarlanan, çıkarılması gereken satır ve sütunlar temizlenebilir. Ayrıca, bazı durumlarda sayısal (numeric) verilerin sözel (nominal) verilere çevrilmesi gibi işlemler yapılabilir. Bunun yanında, boyut azaltma

yöntemi olarak adlandırılan temel bileşen analizi (principal component analysis) de gerçekleştirilir. Temel bileşen analizi, büyük bir değişken kümesini, büyük kümedeki bilgilerin çoğunu içeren daha küçük bir değişken kümesine dönüştürerek büyük veri kümelerinin boyutluluğunu azaltmak için sıklıkla kullanılan bir boyut azaltma yöntemidir.

4.4.3 WEKA programında kullanılan algoritmalar

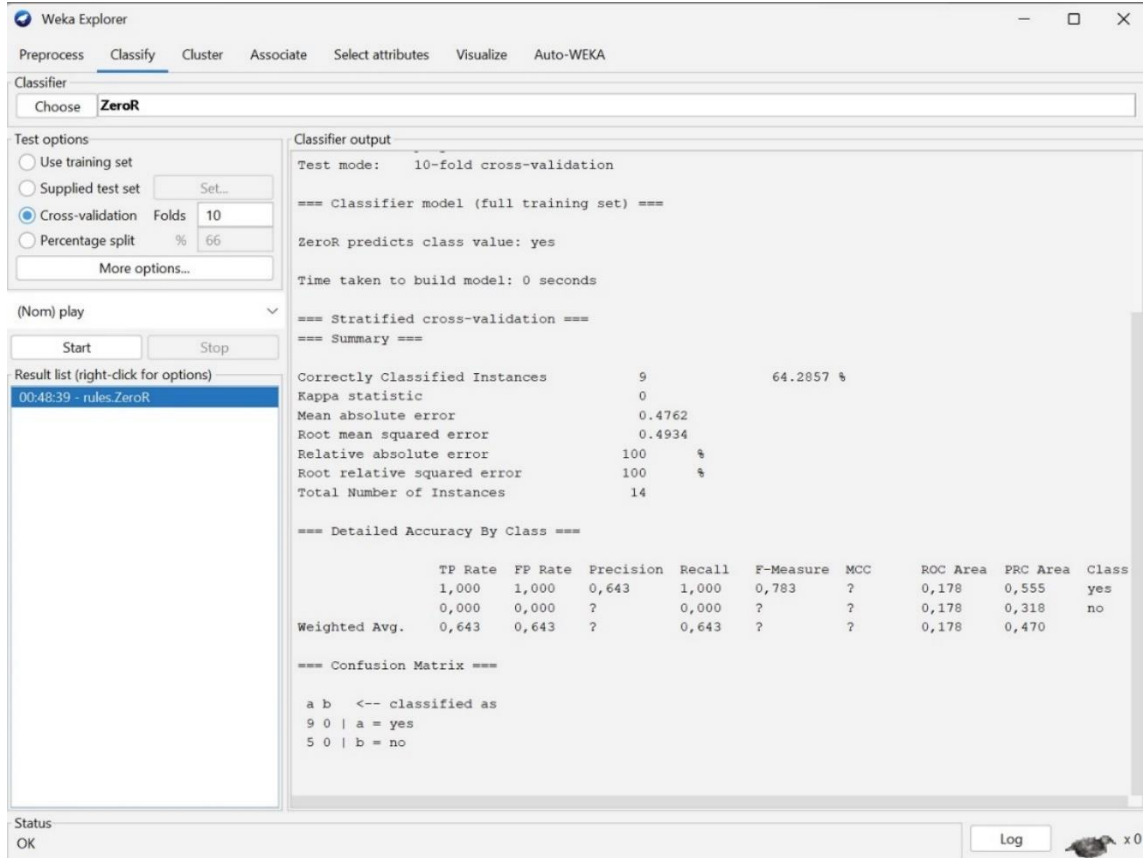
Şekil 4.4'te verildiği gibi, veri madenciliği algoritmalarının WEKA programında uygulamaları aşağıda sıralanmıştır.

4.4.3.1 ZeroR

ZeroR, hedefe odaklanan ve tüm tahmin edicileri göz ardı eden en basit sınıflandırma yöntemidir. ZeroR sınıflandırıcı, basit şekilde çoğunluk kategorisini (sınıfını) tahmin eder. ZeroR'de tahmin yeteneği olmamasına rağmen, diğer sınıflandırma yöntemleri için bir ölçüt olarak temel bir performans belirlemek adına kullanışlıdır (İnt. Kyn. 5). ZeroR algoritmasının WEKA programında kullanımına yönelik olarak örnek bir veri seti Şekil 4.10'da ve uygulama sonucu Şekil 4.11'de verilmiştir.

Viewer					
Relation: weather.symbolic					
No.	1: outlook Nominal	2: temperature Nominal	3: humidity Nominal	4: windy Nominal	5: play Nominal
1	sunny	hot	high	FALSE	no
2	sunny	hot	high	TRUE	no
3	overcast	hot	high	FALSE	yes
4	rainy	mild	high	FALSE	yes
5	rainy	cool	normal	FALSE	yes
6	rainy	cool	normal	TRUE	no
7	overcast	cool	normal	TRUE	yes
8	sunny	mild	high	FALSE	no
9	sunny	cool	normal	FALSE	yes
10	rainy	mild	normal	FALSE	yes
11	sunny	mild	normal	TRUE	yes
12	overcast	mild	high	TRUE	yes
13	overcast	hot	normal	FALSE	yes
14	rainy	mild	high	TRUE	no

Şekil 4.10 WEKA programı örnek veri seti



Şekil 4.11 ZeroR algoritma sonucu

Şekil 4.11 incelendiğinde, iki önemli bölüm öne çıkmaktadır. Bunlar, kullanılan test yöntemi ve sınıflandırıcı çıktısıdır.

Algoritma uygulamasında önemli noktalardan biri, test yöntemidir. WEKA programında algoritmaların uygulanmasında kullanılan test yöntemleri şunlardır (Bouckaert vd. 2018):

- Eğitim setini kullan (Use training set):** Bu test yönteminde sınıflandırıcı, üzerinde eğitildiği örneklerin sınıfını ne kadar iyi tahmin ettiğine göre değerlendirilir.
- Test setini kullan (Supplied test set):** Burada sınıflandırıcı, bir dosyadan yüklenen aynı öz niteliklere sahip bir dizi örneğin sınıfını ne kadar iyi tahmin ettiğine göre değerlendirilir.
- Çapraz doğrulama (Cross-validation):** Sınıflandırıcı, belirlenen katsayıyı kullanılarak çapraz doğrulama ile değerlendirilir. Bu katsayıya (n) göre, algoritma n defa çalışacak ve veri seti n parçaya ayrılacak demektir. Her seferinde n-1 parça

eğitim verisi, kalan tek parça ise test verisi olur. Her döngüde eğitim ve test verileri değişir.

- d. Yüzdesel bölme (Percentage split):** Sınıflandırıcı, test için tutulan verilerin belirli bir yüzdesini ne kadar iyi tahmin ettiğine göre değerlendirilir. Dışarıda tutulan veri miktarı % alanına girilen değere bağlıdır.

Hangi test yöntemi kullanılırsa kullanılsın, çıktı olarak alınan model her zaman tüm eğitim verilerinden oluşturulan modeldir. Genel olarak, çapraz doğrulama yöntemi en çok tercih edilen test yöntemidir.

Algoritma sonucu elde edilen sınıflandırıcı çıktısı bölümünde algoritma çalıştırma bilgileri, sınıflandırıcı modeli, kullanılan test yöntemi, çıktı özet istatistikleri, sınıfların doğruluk tablosu ve hata matrisi bulunmaktadır (Şekil 4.12).

```
Classifier output

=== Run information ===

Scheme:      weka.classifiers.rules.ZeroR
Relation:    weather.symbolic
Instances:   14
Attributes:  5
             outlook
             temperature
             humidity
             windy
             play
Test mode:   10-fold cross-validation

=== Classifier model (full training set) ===

ZeroR predicts class value: yes

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      9           64.2857 %
Kappa statistic                    0
Mean absolute error                 0.4762
Root mean squared error             0.4934
Relative absolute error             100 %
Root relative squared error         100 %
Total Number of Instances          14

=== Detailed Accuracy By Class ===

              TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
              1,000    1,000    0,643     1,000    0,783      ?       0,178    0,555    yes
              0,000    0,000    ?         0,000    ?         ?       0,178    0,318    no
Weighted Avg.   0,643    0,643    ?         0,643    ?         ?       0,178    0,470

=== Confusion Matrix ===

a b  <-- classified as
9 0 | a = yes
5 0 | b = no
```

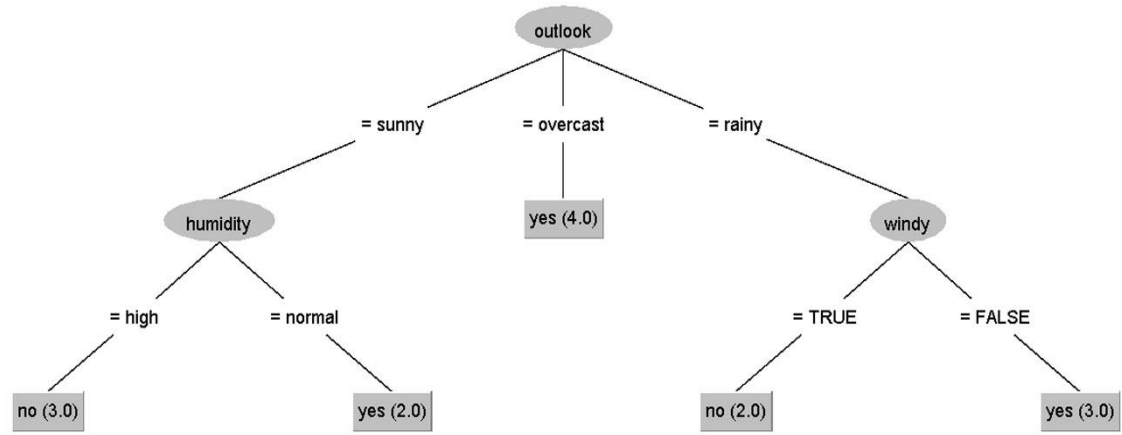
Şekil 4.12 ZeroR algoritma uygulaması sonucu elde edilen sınıflandırıcı çıktıları

4.4.3.2 Naive Bayes

Naive Bayes algoritması, sonuç ihtimaline göre çalışan bir algoritmadır. Her öz niteliğin sonuca olan etki etme olasılığının hesaplanmasına dayanmaktadır. Birden fazla sınıfın alt sınıflarının oluşma olasılığıyla sınıf olasılığının çarpımı şeklinde bir hesaplamadan bahsedilebilir. Temel olarak Bayes teoremine dayanmaktadır. Bu teorem; belirsizlik taşıyan herhangi bir durumun modelinin oluşturularak, bu durumla ilgili evrensel doğrular ve gerçekçi gözlemler doğrultusunda belli sonuçlar elde edilmesine olanak sağlar. Naive Bayes modelinin oluşturulması kolaydır ve karmaşık yinelemeli parametre tahmini gerektirmez. Bu da onu özellikle çok büyük veri kümeleri için kullanışlı hale getirir. Basitliğine rağmen, Naive Bayes sınıflandırıcı genellikle şaşırtıcı derecede iyi sonuç verir ve genellikle daha karmaşık sınıflandırma yöntemlerinden daha iyi performans gösterdiği için yaygın olarak kullanılır (Argüden ve Erşahin 2008, İnt. Kyn. 6).

4.4.3.3 Karar ağacı (Decision tree)

Karar Ağacı algoritması, veriyi sınıflara bölerek çalışan bir algoritmadır. Orijinal hali C4.5 olan algoritma, daha sonra açık kaynak kodlu kullanım için J48 ismiyle tekrar yazılmıştır. Bir karar ağacı, bir kök düğümden yukarıdan aşağıya doğru inşa edilir ve verilerin benzer değerlere sahip (homojen) örnekler içeren alt kümelere bölünmesini içerir. Karar ağacı, bir ağaç yapısı şeklinde sınıflandırma veya regresyon modelleri oluşturur. Bir veri kümesini daha küçük alt kümelere ayırırken, aynı zamanda ilişkili bir karar ağacı aşamalı olarak geliştirilir. Nihai sonuç, karar düğümleri ve yaprak düğümleri olan bir ağaçtır. Bir karar düğümünün iki veya daha fazla dalı vardır. Yaprak düğüm bir sınıflandırmayı veya kararı temsil eder. Ağaçtaki en üstteki karar düğümü, kök düğüm olarak adlandırılan en iyi tahmin ediciye karşılık gelir. Karar ağaçları hem kategorik hem de sayısal verileri işleyebilir. ZeroR modelinde hiçbir tahmin edici yoktur. Naive Bayes, Bayes teoremini ve tahmin ediciler arasındaki bağımsızlık varsayımlarını kullanarak tüm tahmin edicileri içerir; ancak karar ağacı tahmin ediciler arasındaki bağımlılık varsayımlarıyla tüm tahmin edicileri içerir (İnt. Kyn. 7). Algoritma sonucunda elde edilen örnek bir karar ağacı Şekil 4.13'te verilmiştir.



Şekil 4.13 Karar ağacı algoritma sonucu elde edilen örnek bir karar ağacı

Bir karar ağacı oluşturmak için entropi ve bilgi kazanımını değerleri kullanır. Eğer örnek tamamen homojen ise entropi sıfırdır ve eğer örnek eşit olarak bölünmüş ise entropi birdir. Entropinin küçük değer içermesi durumunda özelliğin önemi karar ağacı algoritması için artmaktadır. Diğer taraftan 1'e yaklaştıkça özelliğinin önemi azalır. Bilgi kazanımı ise, tüm sistem entropisi (ulaşılacak istenen amaca yönelik) ile özelliğin bu amaca ulaşılması için hesaplanan entropisi farkıdır. Bu yönüyle entropinin tam tersi bir durum söz konusudur. Karar ağacı inşa edilirken en yüksek bilgi kazanım değerine sahip özellik seçilir. Entropisi sıfır olan bir dal yaprak düğümdür. Entropisi sıfırdan büyük olan bir dalın daha fazla bölünmesi gerekir. Algoritma tüm veriler sınıflandırılana kadar yaprak olmayan dallar üzerinde yinelemeli olarak çalıştırılır (İnt. Kyn. 7).

Şekil 4.13'te verilen karar ağacı, WEKA programında örnek olarak verilen, farklı hava koşullarında golf oynanabilirliğine dair bir veri setine aittir. Bir karar ağacı, karar düğümden yaprak düğümlere tek tek eşlenerek kolayca bir kurallar kümesine dönüştürülebilir. Bu karar ağacı için oluşturulacak kurallar şunlardır:

- 1- Eğer hava yağmurlu ve rüzgar yoksa golf oynanır.
- 2- Eğer hava yağmurlu ve rüzgar varsa golf oynanamaz.
- 3- Eğer hava bulutlu, kapalı ise golf oynanır.
- 4- Eğer hava güneşli ve nem yüksekse golf oynanamaz.
- 5- Eğer hava güneşli ve nem normalse golf oynanır.

4.4.3.4 k en yakın komşuluk (k nearest neighbors – kNN)

Algoritma benzerlik fonksiyonlarının kullanıldığı bir algoritmadır. Sınıflandırılmak istenen noktanın en yakın k komşuya yakınlığıyla ilişkilidir. Genelde yakınlık Öklid mesafesine göre belirlenir. $k = 1$ ise, örnek basitçe en yakın komşusunun sınıfına atanır. $k = 3$ için yeni bir elemanın sınıflandırılması, daha önce sınıflandırılmış elemanlardan en yakın 3 tanesi alınarak yapılır. k için en uygun değerin seçilmesi, en iyi şekilde önce verilerin incelenmesiyle gerçekleştirilir.. Genel olarak, büyük bir k değeri genel gürültüyü azalttığı için daha kesindir; ancak bunun bir garantisi yoktur. Çapraz doğrulama, k değerini doğrulamak için bağımsız bir veri kümesi kullanarak iyi bir k değerini geriye dönük olarak belirlemenin başka bir yoludur. Literatüre göre, çoğu veri kümesi için en uygun k değeri 3-10 arasında olmuştur. Bu, $k = 1$ 'den çok daha iyi sonuçlar üretir (İnt. Kyn. 8).

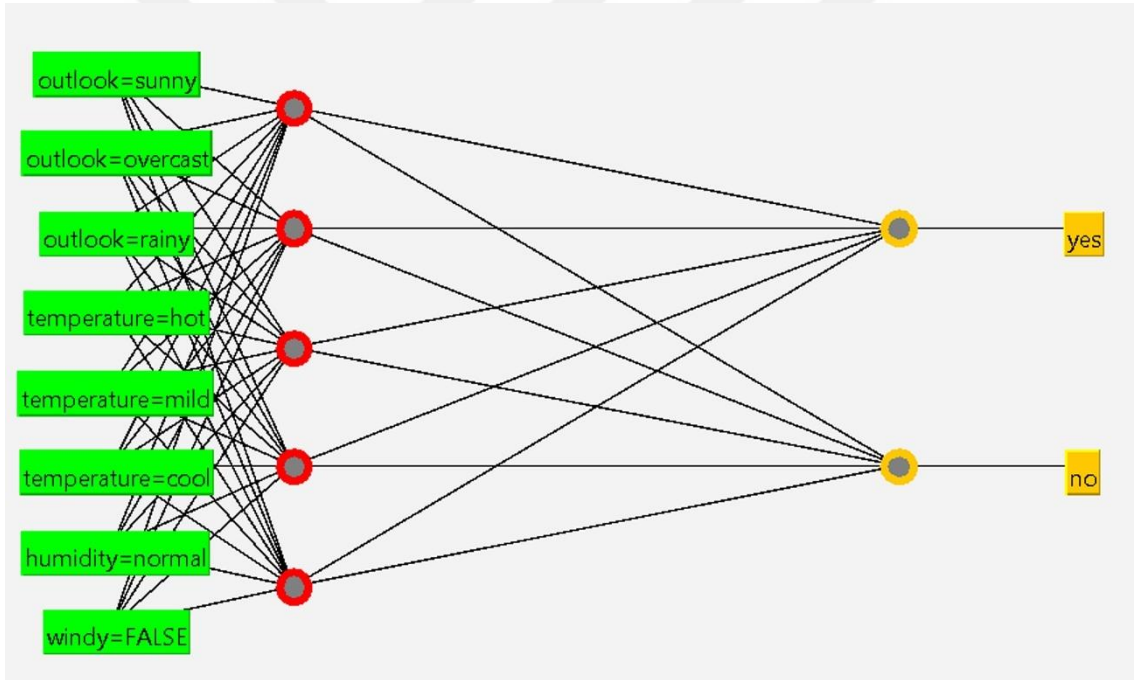
4.4.3.5 Yapay sinir ağları (Artificial neural networks)

Yapay sinir ağı, beyin gibi biyolojik sinir ağını temel alan bir sistemdir. Beyin, elektro-kimyasal sinyaller aracılığıyla iletişim kuran yaklaşık 100 milyar nörona sahiptir. Nöronlar sinaps adı verilen bağlantı noktaları aracılığıyla birbirine bağlanır. Her nöron diğer nöronlarla binlerce bağlantı kurar ve hücre gövdesine ulaşmak için sürekli olarak gelen sinyalleri alır. Sinyallerin toplamı belirli bir eşiği aşarsa, akson aracılığıyla bir yanıt gönderilir. Yapay sinir ağıyla da bilgisayarlarda matematiksel modellemelerle çözülemeyen problemler için sinir ağlarına benzer bir yapı oluşturulmuştur.

Yapay sinir ağı, yapay nöronlardan (düğümler) oluşan bir ağdan meydana gelir. Bu düğümler birbirlerine bağlıdır ve birbirleriyle olan bağlantılarının gücüne göre bir değer atanır: engelleme (inhibition - maksimum -1.0) veya uyarma (excitation - maksimum +1.0). Bağlantının değeri yüksekse, bu güçlü bir bağlantı olduğunu gösterir. Her düğümün tasarımında bir transfer fonksiyonu yerleşiktir. Bir yapay sinir ağında üç tür nöron vardır: giriş düğümleri, gizli düğümler ve çıkış düğümleri. Her bölümdeki düğümler de katmanları oluşturur. Bu katmanlar giriş katmanı, gizli katman ve çıkış katmanı olarak adlandırılır. Giriş düğümleri, sayısal olarak bilgi alır. Bilgi, her düğüme

bir sayı verildiği aktivasyon değerleri olarak sunulur ve sayı ne kadar yüksekse aktivasyon o kadar büyük olur. Bu bilgi daha sonra ağ boyunca aktarılır. Bağlantı güçleri (ağırlıklar), engelleme veya uyarma ve transfer fonksiyonlarına bağlı olarak aktivasyon değeri düğümden düğüme aktarılır. Düğümlerin her biri aldığı aktivasyon değerlerini toplar; daha sonra değeri kendi transfer fonksiyonuna göre değiştirir. Aktivasyon, çıkış düğümlerine ulaşana kadar gizli katmanlar aracılığıyla ağ boyunca akar. Çıktı düğümleri daha sonra girdiyi anlamlı bir şekilde tanımlar (İnt. Kyn. 9).

WEKA programında yapay sinir ağı algoritması, çok katmanlı algılayıcı (Multilayer Perceptron – MLP) olarak isimlendirilmiştir. Algoritma uygulamasında oluşan örnek bir yapay sinir ağı modeli Şekil 4.14’te verilmiştir.



Şekil 4.14 Algoritma uygulamasında oluşan örnek bir yapay sinir ağı modeli

Farklı sinir ağı türleri vardır. Bunlar genellikle ileri beslemeli ve geri beslemeli ağlar olarak sınıflandırılır. İleri beslemeli ağ, girişler, çıkışlar ve gizli katmanlar içeren tekrarlamayan bir ağıdır. Girdi verileri, hesaplamaları gerçekleştirdiği bir işlem elemanları katmanına aktarılır. Her işlem elemanı, girdilerinin ağırlıklı bir toplamına dayalı olarak hesaplamasını yapar. Hesaplanan yeni değerler, daha sonra bir sonraki katmanı besleyen yeni giriş değerleri haline gelir. Bu süreç tüm katmanlardan geçip çıktıyı belirleyene

kadar devam eder. Geri beslemeli ağ döngüler kullanarak her iki yönde hareket eden sinyallere sahip olabilirler. Bu döngüler bir denge durumuna ulaşana kadar sürekli değişen doğrusal olmayan dinamik bir sistem haline gelir. WEKA programında kullanılan çok katmanlı algılayıcı, geri beslemeli bir sınıflandırıcı olarak çalışmaktadır.

4.4.3.6 Destekçi vektör makine (support vector machine – SVM)

Destekçi vektör makine algoritması, başlangıçta verileri sınıflandırmak için kullanılmıştır ve sınıflandırma konusunda kullanılan oldukça etkili ve basit yöntemlerden birisidir. Bu sınıflandırmanın temeli, bazı fonksiyonlar kullanarak nesneleri iki sınıfa ayırmaktır. Bu işlemin yapılması için iki sınıfa da yakın ve birbirine paralel iki sınır çizgisi çizilir. Burada aralarında en büyük aralık bulunan sınır çizgileri kullanılır ve bu sınır çizgileri birbirine yaklaştırılarak ortak sınır çizgisi üretilir. Sınıfları ayıran düzlemler sınır çizgileri olarak bilinir ve amaç tekrarlı optimizasyon kullanarak en büyük ayrımı oluşturan ortak sınır çizgisini belirlemektir. Fonksiyonlar çekirdek (kernel) olarak adlandırılır ve işlem haritalama olarak adlandırılır. Ortak sınır çizgisinin her iki tarafında en yakın mesafede yer alan sınır çizgileri, destek vektörleri olarak adlandırılır. SVM yönteminde hem sınıflandırma hem de regresyonda bazı kısıtların etkilerini ortadan kaldırmak için bir hata fonksiyonu bulunmaktadır. Sınıflandırma yaparken kullanılacak çekirdek (kernel), işlemin başarısını etkileyebilir. Verileri sınıflamak için kullanılacak sınır çizgilerinin yapısını (doğrusal, polinom, sigmoid vb.) belirtir (Martin ve Miranda 2012).

4.4.3.7 Rassal orman (random forest)

Denetimli sınıflandırma algoritmalarından biridir. Hem regresyon hem de sınıflandırma problemlerinde kullanılmaktadır. Algoritma birden fazla karar ağacı üreterek sınıflandırma işlemi esnasında, sınıflandırma değerini yükseltmeyi hedefler. Rassal oluşturulmuş karar ağaçlarından elde edilen çıktıları birleştirerek sınıflandırma ve regresyon işlemlerinde kullanır (Breiman 2001). Rassal orman oluşturmak için kaç karar ağacı kullanılacağı önemlidir. Genel olarak, kullanıcı ağaç sayısını deneme yanılma yöntemiyle belirler. Bazen ağaç sayısını artırdığında, aslında neredeyse hiçbir performans kazancı olmadan yalnızca daha fazla hesaplama gücü harcanmaktadır.

4.5 Madencilikte Veri Madenciliği Uygulamaları

Madenlerde iş sağlığı ve güvenliği çalışmaları, sektörde gerçekleştirilen teknolojik ilerlemeler gibi araştırma konusu olmuştur. Birçok araştırmacı madencilikte iş sağlığı ve güvenliği üzerinde çalışmalar gerçekleştirmiş, genel olarak yapılan çalışmalarda riskler belirlenmiş, alınabilecek önlemler tartışılmıştır. Bu çalışmalar çoğunlukla istatistiksel yöntemlerle ele alınmış, daha derin çıkarımlar ve sonuçlar sağlayan veri madenciliği yöntemlerinin kullanımı kısıtlı kalmıştır. Konuyla ilgili literatüre bakıldığında aşağıda verilen çalışmalara ulaşılabilmektedir.

Dessureault vd. (2007) yapmış oldukları çalışmada, madenlerdeki iş sağlığı ve güvenliği verilerini kullanarak veri madenciliği uygulama potansiyelini araştırmayı amaç edinmişlerdir. Bu çalışmalarında ABD’de faaliyet gösteren İş Sağlığı ve Güvenliği Ulusal Enstitüsü (NIOSH) veri setini kullanmışlardır. Kullanacakları veri tabanını üzerinde gelişmiş analitik işlemler yapabilmeleri için veri ambarı metodolojisini kullanarak veri tabanını modernize etmişlerdir. Günümüzde kullanılmakta olan veri madenciliği yazılımları iş kazalarının nedenlerini belirlemek amacıyla da kullanılmaktadır.

Shuangyue ve Li (2013), kömür madeni gizli tehlike verileri üzerinde Apriori algoritması kullanmışlar ve birliktelik kuralları üretmişlerdir. Gizli tehlike yönetimi ve önlenmesinde işletmeler için iyileştirme önerileri sağlamanın ve sonuçta kazaları önlemede ve kaza kayıplarını azaltmada önemli bir pratik öneme sahip olduğu belirtilmiştir.

Mevsim (2016) yapmış olduğu çalışmada, yeraltı kömür işletmelerindeki grizu patlamalarının ana nedenlerini belirlemeyi hedeflemiştir. Bu çalışmasında, TTK’de 1875-2014 yılları arasında meydana gelen grizu patlaması olaylarını baz almıştır. Bu dönem içerisinde TTK’de 815 çalışanın yaşamını yitirmesine neden olan 67 grizu patlaması meydana gelmiştir. Çalışmada patlamaların kök nedenlerini belirlemede bir tümden gelim risk analiz tekniği olan hata ağacı analizi (HAA) tekniği kullanılmıştır. Meydana gelen grizu patlamalarının ana ve yan sebeplerinin belirlenmesinde, ReliaSoft BlockSim-7 yazılımı kullanılmıştır. Çalışma bulgularında, mekanik tutuşma kaynaklarının grizu patlamalarında en önemli etkiye sahip olduğu görülmüştür. Bunun yanında havalandırma

sistemindeki yanlılıklar, kıvılcım, elektrik tutuşma kaynakları ve metan degajının da büyük öneme sahip olduğu görülmüştür. Çalışma sonucunda bir grizu patlamasının meydana gelme olasılığının 108 ayda %100 olduğu görülmüştür; ancak grizu patlamasının meydana gelmesinde en büyük öneme sahip olan mekanik tutuşma kaynağının ortadan kaldırılması durumunda %100 olasılık ihtimalinin 255 aya yükseldiği görülmüştür. Bu çalışmada kantitatif risk değerlendirmesinin yanında kalitatif HAA analizi de uygulanmıştır. Kalitatif ağaçta 27 orta dereceli ve 65 basit olay (kök neden) tespit edilmiş ve diğer grizu patlamaları çalışmaları ile karşılaştırıldığında kapsamlı bir HAA sunulmuştur. Türkiye Taş Kömürü Kurumunda meydana gelen grizu patlamalarının bu yöntemle irdelenmesi ülkemiz madencilik sektörü için büyük önem arz etmektedir. Ayrıca, bu çalışma literatüre de büyük katkı sağlamıştır. Yapılacak çalışmalar ile kapsam daha da genişletilerek hem kamu hem de özel sektör iş yerleri de dahil edilerek diğer maden kazaları da incelenebilir. Meydana gelen grizu patlamalarının kök nedenleri belirlendikten sonra, bu nedenleri önlemek için alınması gereken tedbirler konusunda iş sağlığı ve güvenliği yasal mevzuatına da katkı sağlanabilir.

Erdoğan (2016), yapmış olduğu çalışmada TTK'de 2000-2014 yılları arasında meydana gelen iş kazalarını baz almıştır. Çalışmasında varyans analizi (ANOVA) tekniğini kullanarak iş kazası istatistiklerini değerlendirmiştir. Bu yöntem ile iş kazalarının sayısı, çeşidi ve yeri ile kazazedelerin eğitim durumu, yaşı ve tecrübesi ve kaza kaynaklı yaralanma ve iş gücü kaybına ilişkin verileri irdelemiştir. Tehlikeler kişisel, operasyonel ve çalışma ortamı şeklinde sınıflandırılmıştır. İş kazası sonucunda meydana gelen yaralanmalar ve iş gücü kayıpları da kaza şiddeti olarak kabul edilmiştir. Her bir tehlike sınıfı için ayrı ayrı olay ağaçları oluşturulmuş, yaralanmalar ve iş gücü kaybı için alt ve üst risk seviyeleri belirlenmiştir. Belirlenen bu risk sınırlarının düşürülmesi için alınması gereken önlemler konusunda önerilerde bulunulmuştur.

Sanmiquel vd. (2018), İspanya Çalışma ve Sosyal Güvenlik Bakanlığı'nın 2005-2015 yılları arasındaki verilerini kullanarak madencilik sektöründeki iş kazalarının veri madenciliği teknikleri ile analizini yapmışlardır. İş kazalarında yer alan en önemli değişkenler ve bunların birliktelik kuralları belirlenmiştir. Veri madenciliği teknikleri, kazaların temel nedenini bulmak için yararlı bir araç olarak seçilmiştir.

Ak (2019), nöro-bulanık mantık yaklaşımı kullanarak bir yeraltı metal madeninde risk değerlendirme çalışması yapmış ve uyarlamalı nöro-bulanık çıkarım sistemi (ANFIS) tabanlı bir model uygulamıştır. Çalışma yapılan madenler için kullanılan geleneksel yöntemlerden daha iyi sonuçlar aldığını ifade etmiştir.

Gerassis vd. (2020) yapmış oldukları çalışmada, İspanya'daki maden işletmelerinde meydana gelen iş kazası istatistiklerini irdelemişlerdir. Bu çalışmada, yapay zeka uygulamaları ile farklı iş kazaları için bilinmeyen ilişkilerin araştırılabilmesi veri madenciliği yöntemleriyle uygulanabilirliği ifade edilmiştir. Aynı zamanda madencilik işletmelerindeki bakım, havalandırma, ekipman performansı gibi operasyonel verilerde veri madenciliği teknikleri kullanılarak irdelenmiştir. İş sağlığı ve güvenliğini hedeflenen düzeye çıkarabilmek için işletmelerdeki tüm istatistiki bilgiler veri tabanlarında toplanmalı ve veri madenciliği teknikleriyle irdelenmelidir. Veri tabanlarında çalışan, ekipman performansı, saha durumları ile ilgili veriler toplanabilmektedir. Madencilik sektöründe iş sağlığı ve güvenliğine öncelik veren ülkelerde bu veri tabanlarının oluşturulmasına ve değerlendirilmesine yönelik izlenecek yol haritası hakkında resmi kuruluşlar tarafından da öneriler de bulunmaktadır. Uygulanan veri entegrasyonu veya veri madenciliği tekniğinin en iyi şekilde yorumlanabilmesi için öncelikle kullanılan verilerin özelliklerinin iyi anlaşılması gerekmektedir.

İphar ve Çukurluöz (2020), klasik karar matrisi risk değerlendirme yönteminde karşılaşılan belirsizliklerin üstesinden gelmek için, risk değerlendirme sürecini geliştirmeye yönelik bulanık mantık tabanlı bir güvenlik değerlendirme yöntemi tanıtmışlardır. Önerilen yöntem, Türkiye'deki mekanize kömür madenlerinde mevcut olabilecek risklerin daha gerçekçi bir değerlendirmesi olarak önerilmiştir. Bu sayede mekanize yeraltı kömür madenlerindeki riskli durumlar ve operasyonlar uzman bilgisi ve mühendislik muhakemesi ile dilsel formlarda belirlenmiştir. Dolayısıyla, böyle bir değerlendirmenin yakın gelecekte tam mekanize kömür üretiminin yapılacağı kömür madenleri için değerli bir rehber olacağını belirtmişlerdir.

Opokua ve Opoku (2021) bir yeraltı maden ocağında meydana gelecek göçükte işçilerin hayatta kalma oranını tahmin etmek için bir uyarlamalı nöro-bulanık çıkarım sistemi

(ANFIS) geliřtirmiřtir. Hayatta kalma oranını dūřuk ve yūkksek řekilde tahmin edebilmek iin alıřma vardiyaları, kaya tipleri, kayacın zellikleri, tavan kalınlıęı ve alıřan ayaęın zellikleri gibi parametreler kullanılarak yapay sinir aęı ile bulanık kurallar tūretilmiř ve bazı durumlarda %90'a yakın doęru tahminlerin yapıldıęı ifade edilmiřtir.

Tūrkiye'de yapılan lisansūstū tez alıřmaları Haziran 2023 itibariyle incelendięinde, veri madencilięi ile ilgili toplam 695 adet alıřma bulunduęu gze arpmaktadır. Bu alıřmaların būyūk bir blūmū iřletme, ekonomi, istatistik, bilgisayar bilimleri anabilim dallarında ilgili olduęu grūlmektedir. Maden Mūhendislięi disiplininde yapılan tez alıřmaları kısıtlıdır. Aynı řekilde, TTK'de meydana gelen iř kazaları ile ilgili veriler kullanılarak yapılmıř veri madencilięi yntemlerini ieren alıřmalar azdır. Bunun yanında, veri madencilięi zellikle iř kazaları verileri gibi karmařık ve ok parametrelili verilerin aıklanmasında olduka etkilidir. Literatūrde yapılan alıřmalara katkı konulması amacıyla, TTK Amasra Tařkmūrū iřletmesinde 2010-2021 yılları arasında meydana gelen yaralanmalı iř kazaları, veri madencilięi yntemleriyle irdelenerek iřletme iin de yararlı bilgilerin ıkarımı hedeflenmiřtir.

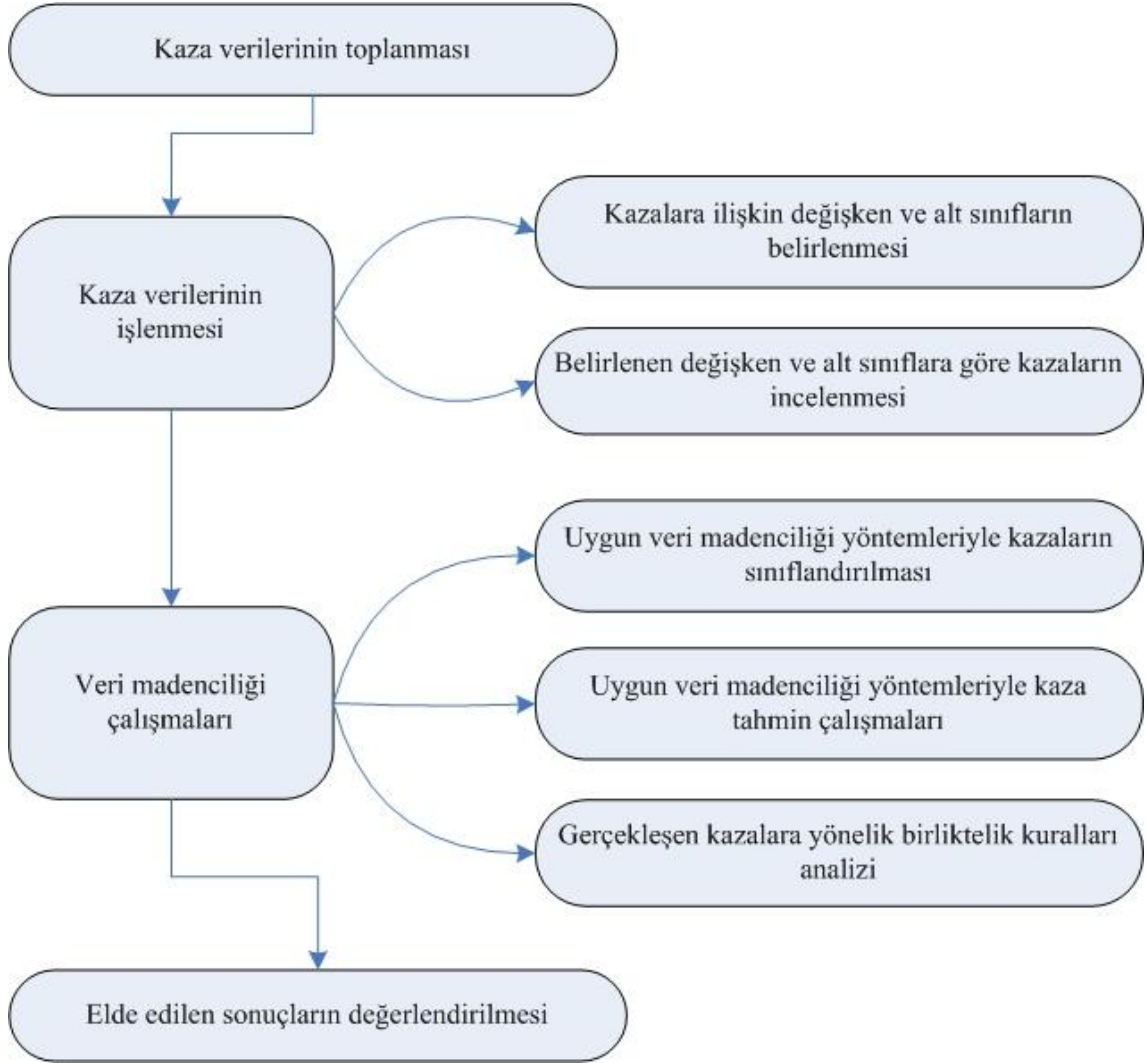
5. TEZ KAPSAMINDA YAPILAN ÇALIŞMALAR

Bu tez kapsamında, TTK ATİM çalışma alanı olarak seçilmiştir. ATİM’de 2010-2021 yılları arasında meydana gelen yaralanmalı iş kazaları incelenmiştir. Meydana gelen yaralanmalı iş kazalarına yönelik toplanan verilerle iş kazalarının meydana geldiği alanlar veri madenciliği yöntemleri kullanılarak ilişkilendirilmiştir. Çalışmada analizler için, WEKA 3.8.6 versiyonu açık kaynak kodlu veri madenciliği programı kullanılmıştır.

Bu kapsamda, veri işleme çalışmalarından sonra üç aşamada uygulama gerçekleştirilmiştir:

- a) Çalışmanın birinci aşamasında, 2010-2021 yıllarını kapsayan veriler kazaların meydana geldiği yerler temel alınarak sınıflandırılmıştır. Bu bölümde kazalarla ilgili olarak belirlenen değişkenler (14 adet) kullanılmıştır. Çalışmada uygulanan analizlerde yanıt olarak nitelendirilen kaza yerlerinin tespitinde, diğer 13 değişken tahmin edici olarak kullanılmıştır. Uygun veri madenciliği algoritmalarıyla kaza yerlerinin tahminine yönelik bir sınıflama gerçekleştirilmiştir.
- b) Çalışmanın ikinci aşamasında, işletmede 2010-2020 yılları arasında meydana gelen yaralanmalı kaza verileri incelenerek 2021 yılında meydana gelen kazalar, meydana geliş yerleri açısından tahmin edilmeye çalışılmıştır. Bu aşamada da veri madenciliğinde kullanılan uygun algoritmalar ve yöntemler seçilmiştir.
- c) Çalışmanın son aşamasında ise, 2010-2021 yılları arasındaki kaza verileri birliktelik kuralları analizine tabi tutulmuş ve oluşan kazalarda parametrelerin birlikte görülme durumları da tespit edilmiştir.

2010-2021 yılları arasında işletmede meydana gelen yaralanmalı iş kazaları dikkate alınarak yapılan çalışmada izlenen yöntem Şekil 5.1’de verilmiştir.



Şekil 5.1 Çalışmada izlenen yöntem

5.1 Kaza Verilerinin Toplanması

Çalışmada kullanılan veri seti, ATİM’de 2010-2021 yılları arasında gerçekleşen yaralanmalı kazaları içermektedir. Bu yıllar içerisinde meydana gelen kaza sayılarıyla ilgili bilgiler Çizelge 5.1’de verilmiştir. İnceleme yapılan yıllarda toplam 2169 iş kazası meydana gelmiştir. Bu kazaların 2167 adedi yaralanmalı iş kazasıdır ki çalışmanın temel incelenecek verileri bu kazalardır. Bu kazalarla ilgili detaylı bilgiler, TTK’nin kaza istatistiklerini tuttuğu elektronik veri tabanından, gerekli izinler alınarak çekilmiştir.

Çizelge 5.1 Yıllara göre meydana gelen yaralanmalı iş kazaları sayıları (İnt. Kyn. 10)

Yıllar	Yaralanmalı iş kazaları sayıları
2010	267
2011	251
2012	194
2013	224
2014	136
2015	151
2016	137
2017	133
2018	112
2019	190
2020	164
2021	208
TOPLAM	2167

5.2 Kaza verilerinin işlenmesi

Veri tabanından alınan ham iş kazaları verilerinin yapılacak olan istatistiksel çalışmalara girdi parametreleri haline getirilebilmesi için, öncelikle işlenmesi ve sınıflandırılması gerekmektedir. 2010-2021 yıllarını kapsayan 12 yıllık kaza verilerinin tamamı veri tabanından çekilmiş ve öncelikle ölümlü kazalar tespit edilmiştir. Araştırma kapsamındaki yıllar itibariyle işletmede sadece 2012 ve 2013 yıllarında birer olmak üzere toplam 2 ölümlü kaza meydana gelmiştir. Ölümlü kaza verilerinin çıkarılmasıyla toplam 2167 yaralanmalı kaza incelemeye alınmıştır.

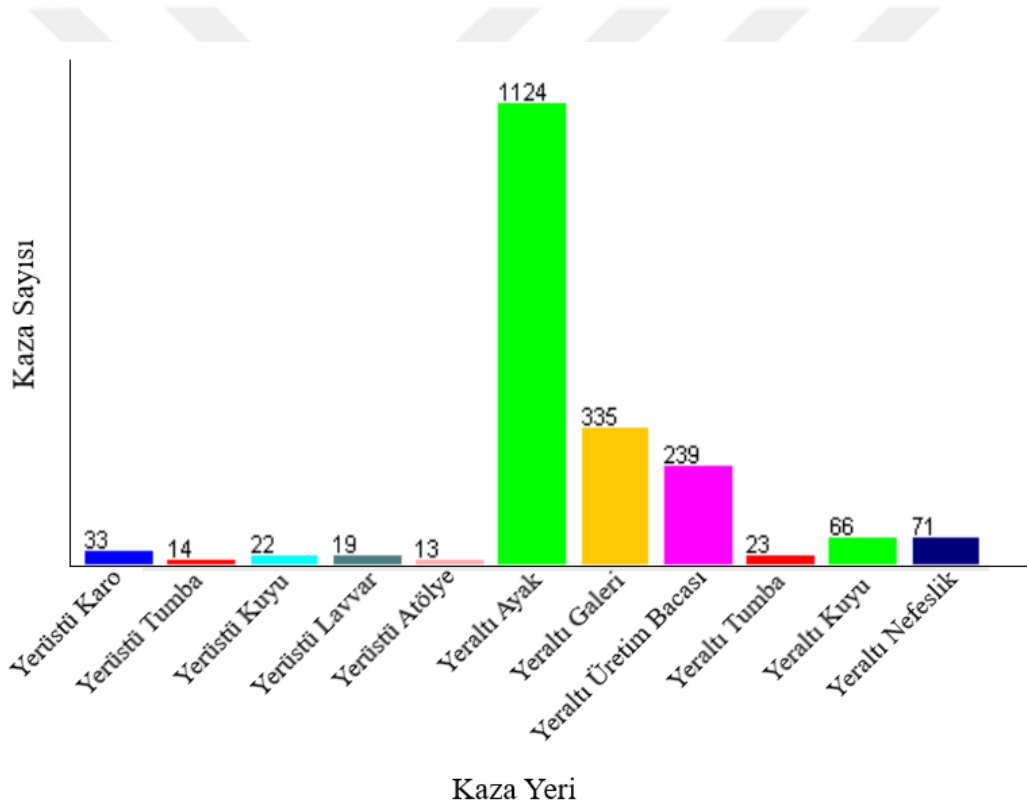
Yaralanmalı kazalar, belirlenen 14 temel değişken çerçevesinde gruplandırılmıştır. Bu 14 temel değişken arasında oluşan kazalarla doğrudan ilgili değişkenler (etkilenen organ, kazanın gerçekleştiği ay, kazanın gerçekleştiği hafta günü, kazalananın yaşı, kazalananın deneyimi, kazalananın eğitim durumu, kazanın saat olarak hangi vardiyada gerçekleştiği, vardiyanın hangi saatinde meydana geldiği, kaza sonucu oluşan kayıp gün sayısı, çalışan tipi, kaza sebebi ve kaza yeri) yanında kazanın meydana geldiği ayda yapılan aylık üretim miktarı ve aylık yevmiye sayıları da bulunmaktadır. Kaza incelemesi için seçilen 14 değişken, 108 alt sınıfa ayrılmıştır. Kazalara ilişkin oluşturulan 14 değişken ve 108 alt sınıf Çizelge 5.2’de verilmiştir.

Çizelge 5.2 Kazaların sınıflandırılmasında kullanılan değişken ve bu değişkenlere ait alt sınıflar

Değişkenler	Değişken Tanımı	Alt sınıflar
Etkilenen Organ	Kaza sonucunda çalışanın zarar gören organını belirtmektedir.	Kol, Vücut, Yüz, El parmağı, Baş, Diz, Bacak, Ayak parmağı, Bel
Ay	Kazanın meydana geldiği ayı tanımlamaktadır.	Ocak, Şubat, Mart, Nisan, Mayıs, Haziran, Temmuz, Ağustos, Eylül, Ekim, Kasım, Aralık
Gün	Kazanın meydana geldiği günü belirtmektedir.	Pazartesi, Salı, Çarşamba, Perşembe, Cuma, Cumartesi, Pazar
Yaş	Kaza tarihinde kazalananın yaşı tanımlanmaktadır.	18-25, 26-30, 31-35, 36-40, 41-45, ≥ 45
Deneyim	Kaza tarihinde kazalananın kaç yıllık deneyiminin olduğu verilmektedir.	0-5, 6-10, 11-15, 16-20, 21-25, ≥ 26
Eğitim Durumu	Kazalananın eğitim durumunu vermektedir.	İlkokul, Ortaokul, Lise, Önlisans, Lisans
Vardiya	Kazanın meydana geldiği vardiya saatini tanımlamaktadır.	0-8, 8-16, 16-24
Çalışma Saati	Kazanın vardiyada kaçınıcı saatte meydana geldiğini göstermektedir.	0-1, 1-4, 4-8
Kayıp İş Günü Sayısı	Çalışanın kazadan dolayı çalışamadığı gün sayısıdır.	0, 1-3, 4-28, 29-180, >180
Aylık Üretim	Kazanın meydana geldiği ayda ton olarak yapılan tüvenan kömür üretimini göstermektedir.	≤ 7999 , 8000-9999, 10000-11999, 12000-13999, 14000-15999, 16000-17999, 18000-19999, 20000-21999, 22000-23999, 24000-25999, ≥ 26000
Aylık Yevmiye	Kazanın meydana geldiği ayda yapılan üretime karşılık harcanan toplam yevmiye sayısıdır.	≤ 3999 , 4000-4999, 5000-5999, 6000-6999, ≥ 7000
Çalışan Görevi	Kazalananın asli görevini ifade etmektedir.	Barutçu, Başmadenci, Sondajcı, Tarama-Söküm İşçisi, Meydan-Çevre İşçisi, Nakliyat İşçisi, Makine Operatörü, Bakım-Onarım İşçisi, Mekanize Üretim İşçisi, Hazırlık İşçisi, Pano-Ayak Üretim İşçisi, Vardiya Mühendisi, Pano-Ayak Nezaretçisi, Teknisyen, Lavvar İşçisi
Kaza Sebebi	Kazaya sebep olan ana nedeni ifade etmektedir.	Yerüstü Ekipman Teknik Kullanımı, Yerüstü Vurma-Çarpma, Yerüstü Malzeme Taşıma-Kullanma, Yerüstü Kayma-Düşme-Burkulma, Yerüstü Diğer, Yeraltı Ekipman Teknik Kullanımı, Yeraltı Göçük, Yeraltı Vurma-Çarpma, Yeraltı Malzeme Taşıma-Kullanma, Yeraltı Kayma-Düşme-Burkulma, Yeraltı Diğer
Kaza Yeri	Kazanın meydana geldiği yeri belirtmektedir.	Yerüstü Karo, Yerüstü Tumba, Yerüstü Kuyu, Yerüstü Lavvar, Yerüstü Atölye, Yeraltı Ayak, Yeraltı Galeri, Yeraltı Üretim Bacası, Yeraltı Tumba, Yeraltı Kuyu, Yeraltı Nefeslik

Çizelge 5.2’de verilen kaza değişkenleri ve bunların alt sınıfları belirlenirken, belli standartlarda bir değerlendirme yapılabilmesi açısından kazaların tamamı oluş yerleri, oluş biçimleri açısından incelenmiştir. Ayrıca, işletme koşulları ve çalışma şekilleri de dikkate alınmıştır. Bu inceleme neticesinde kazaların alt sınıflarına ilişkin bir birliktelik sağlanmıştır.

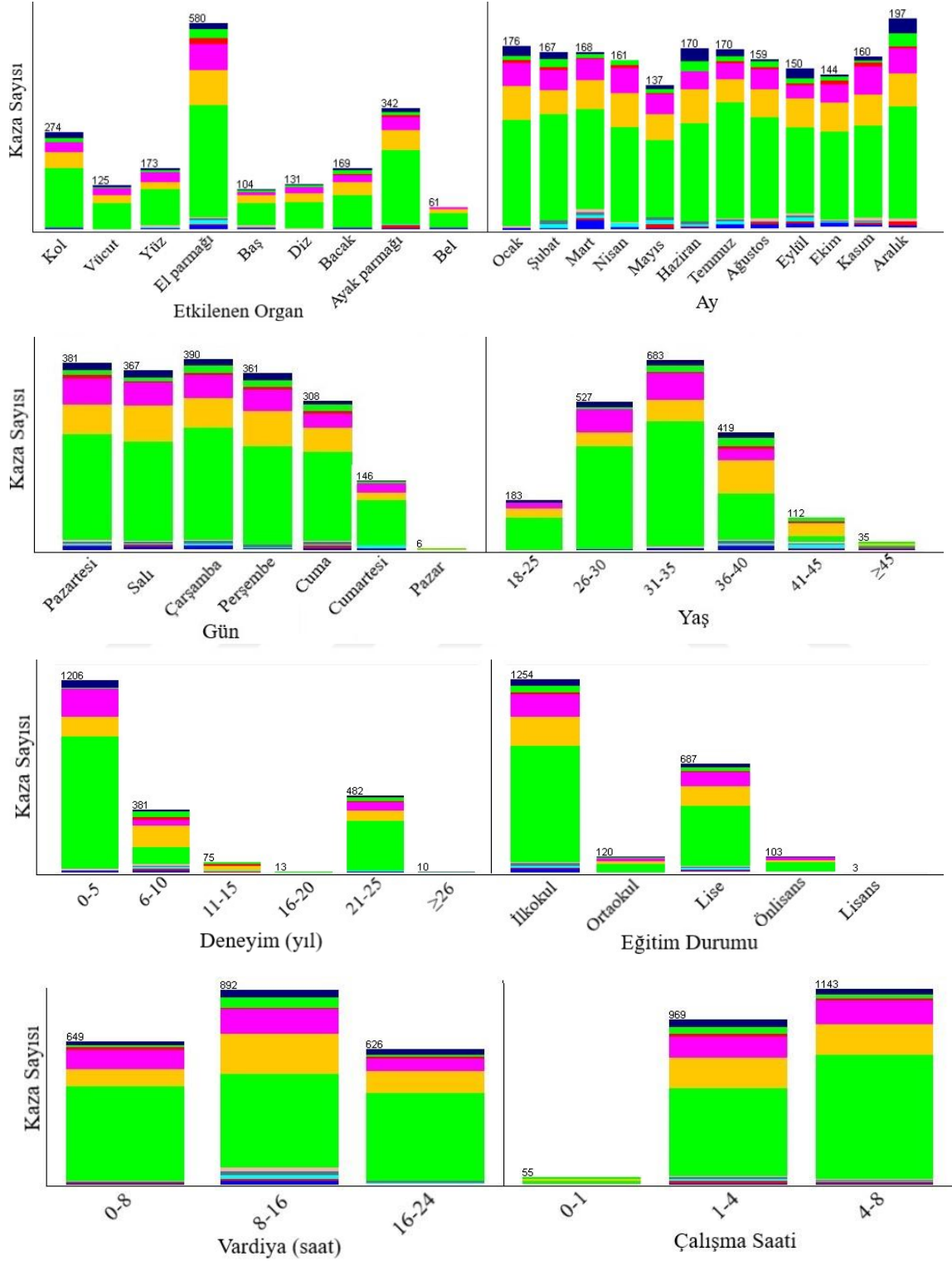
Değişkenler ve alt sınıflar belirlendikten sonra, çalışmada yapılan analizlerde yanıt olarak tanımlanan kaza yeri değişkenine ve bu değişkene bağlı olarak tahmin edici değişkenlere göre kaza sayıları dağılım grafiklerinde incelenmiştir. Kaza yeri değişkenine göre meydana gelen kaza sayıları Şekil 5.2’de verilmiştir.



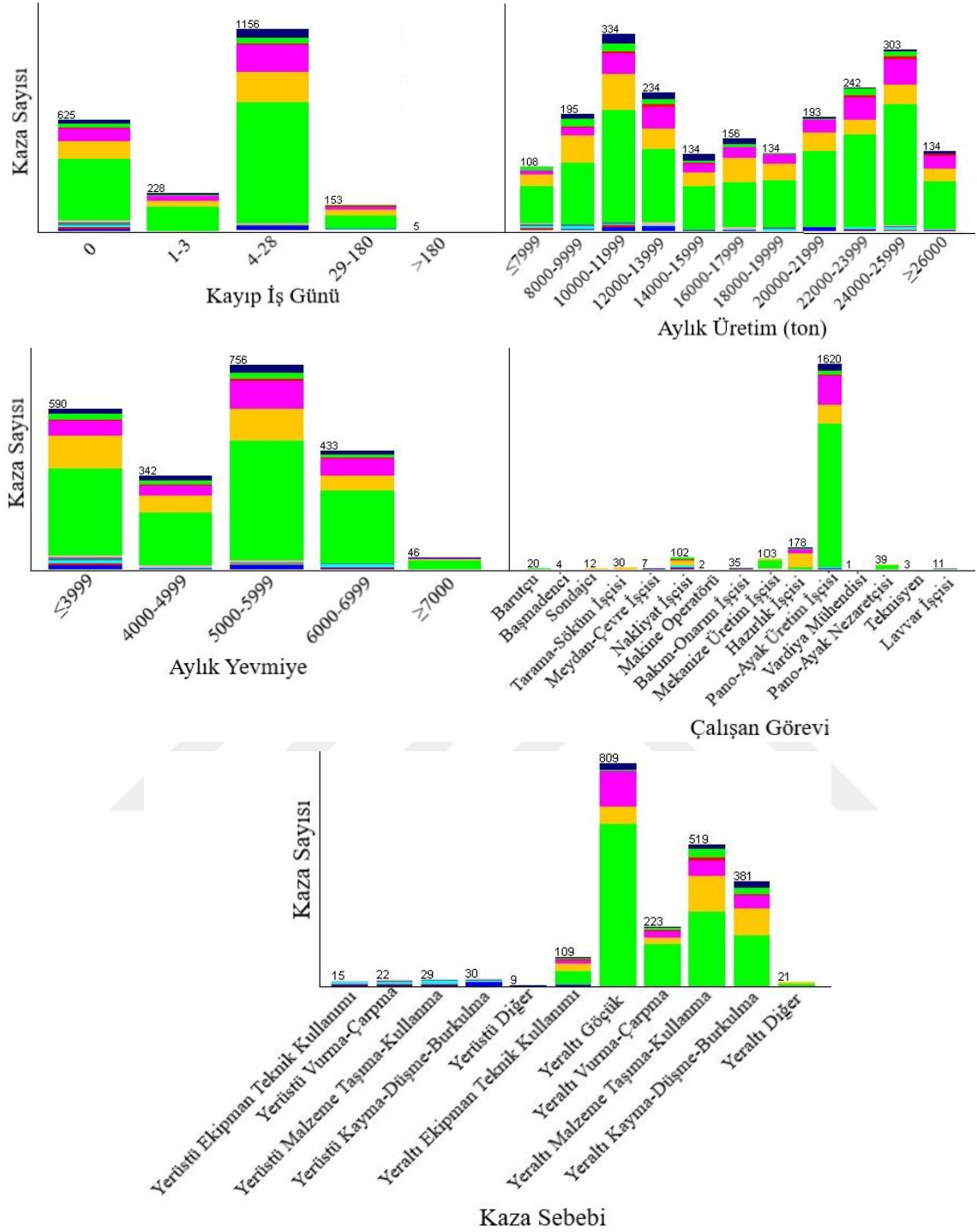
Şekil 5.2 Kaza yerlerine göre meydana gelen kaza sayıları

Şekil 5.2’ye göre işletmede en fazla kaza yeraltında ayaklarda meydana gelmiştir. Daha sonra yeraltı galerilerinde ve yeraltı üretim bacalarında kazaların meydana geldiği görülmektedir. Yeraltı ayak ve yeraltı üretim bacalarında meydana gelen kazalar doğrudan kömür üretimi sırasında oluşan kazalardır. Bu şekilde toplam 1363 kaza meydana gelmiş ve bu toplam yaralanmalı kazaların %62,9’unu oluşturmaktadır.

Kaza yerlerine bağlı olarak diğer değişkenler açısından kaza sayıları da incelenmiştir (Şekil 5.3).



Şekil 5.3 Kaza yerlerine bağlı olarak diğer değişkenler açısından kaza sayıları



Şekil 5.3 Kaza yerlerine bağlı olarak diğer değişkenler açısından kaza sayıları (devam ediyor)

Şekil 5.3'te çalışmada kullanılan tahmin edici değişkenlerinin kaza sayıları açısından dağılımları verilmiştir. Bu şekillerde kaza yerlerinin dağılımlar içerisindeki ağırlıkları da görülmektedir. Şekil 5.2'de verilen kaza yerleri renkleri, Şekil 5.3'te verilen dağılımlarda aynı renklerle yer almaktadır.

Etkilenen organ açısından en fazla kazalanan organların el ve ayak parmağı olduğu görülmektedir. Kazalar gerçekleştiği ay olarak değerlendirildiğinde, ortalama aynı sayılarda gerçekleştiği söylenebilir. Kaza sayıları, mesai günlerinde (Pazartesi, Salı, Çarşamba, Perşembe ve Cuma) yaklaşık aynı seviyededir. En çok kazalanan yaş grubu 31-35 yaş aralığıdır. Kazalanan çalışanların deneyimlerine bakıldığında 0-5 yıl aralığındaki çalışanlar en fazla kazaya maruz kalan gruptur. Eğitim durumları açısından ilkökul mezunu çalışanlar en fazla kazalanmıştır. Kazaların meydana geldiği vardiya ve vardiya içerisindeki çalışma saati dikkate alındığında en fazla kaza sırasıyla 8-16 vardiyasında ve vardiyanın 4-8 saat aralığında meydana gelmektedir. Vardiyaların ilk saatlerinde (0-1 saat) kaza sayıları oldukça düşüktür. Meydana gelen kazalarda kayıp iş günü en fazla 4-28 gün aralığında gerçekleşmiştir. Aylık üretim ve aylık yevmiye sayılarına göre kaza sayıları dalgalanma göstermektedir. Çalışan görevine bakıldığında pano-ayak üretim işçileri en fazla kazalanan gruptur. Aynı şekilde kaza sebeplerine göre yeraltında oluşan göçükler en fazla kaza sebebidir.

5.3 Veri madenciliği çalışmaları

ATİM’de 2010-2021 yılları arasında meydana gelen kazalar üç aşamada incelenmiştir. Öncelikle tüm yıllara ait kazalar uygun veri madenciliği algoritmalarıyla sınıflandırılmış, bir sonraki aşamada 2021 yılı kaza yerlerinin, 2010-2020 yılları verilerinden tahminine yönelik bir analiz gerçekleştirilmiştir. Son aşamada ise tüm yıllara ait kaza verilerine birliktelik kuralları analizi uygulanmıştır.

5.3.1 2010-2021 yılları kaza verilerinin veri madenciliği yöntemleriyle değerlendirilmesi

Çalışmanın bu aşamasında, ele alınan yıllar itibariyle yaralanmalı kaza verilerinden oluşturulan veri tabanında uygun veri madenciliği algoritmaları kullanarak sınıflandırma işlemleri yapılmıştır. Bu sınıflandırma işleminde iki temel amaç söz konusudur. Bunlardan birincisi, tüm veriyi en iyi şekilde temsil eden algoritmanın belirlenmesidir. Diğeri ise, kullanılacak olan veri madenciliği algoritmalarının istenilen parametreyi tahmin etme kabiliyetinin tespit edilmesidir. WEKA programıyla yapılan

değerlendirmelerde, bu amaçlar doğrultusunda çeşitli test yöntemleri kullanılmaktadır. Çalışmanın bu aşamasında, algoritmalarla yapılan modelleme çalışmalarında eğitim seti (training set) ve çapraz doğrulama (cross validation) test yöntemi olarak seçilmiştir. Bu test yöntemleri yanında algoritmaların başarı oranlarının belirlenmesinde bazı istatistikler kullanılmıştır. Kappa istatistiği, ortalama mutlak hata (MAE) ve ortalama kare kök hata (RMSE) değerleriyle algoritma başarısı değerlendirilmiştir.

Veri seti üzerinde çeşitli algoritmaların uygulanmasında tahmin yeteneğinin başarı oranının artırılması amaçlanmıştır. Başarı oranı temel olarak, doğru şekilde sınıflandırılan veri sayısının veri seti içindeki yüzdesiyle belirlenmektedir. Veri madenciliği uygulamalarında algoritmaların ana ayarlarında herhangi bir değişiklik yapılmamıştır. Yalnızca başarı oranının durumu göz önünde bulundurularak, daha iyi sonuç alınmasına yönelik bazı parametre değerleri değiştirilerek denemeler yapılmıştır. Modelleme çalışmalarında ilk uygulanan algoritma ZeroR algoritmasıdır. Bu algortmada en basit veri madenciliği algoritmasıdır ve herhangi bir ayarlama söz konusu değildir. Bu algoritma, yapılan modelleme çalışmalarında temel bir sonuç verir. Yapılan diğer algoritmaların başarısını görebilmek, ilişki düzeyinin artışının gözlenmesi açısından önemli bir algortmadır. Veri madenciliğinde modelleme çalışmalarının başarısı, ZeroR uygulamasının başarısının, yani bu algoritma sonucunda elde edilen korelasyon katsayısının ne kadar ilerletildiğiyle ölçülebilir.

Veri seti üzerinde uygulanan diğer algoritmalar sırasıyla J48 Karar Ağacı (J48), k En Yakın Komşuluk (kNN), Yapay Sinir Ağları (MLP – Multilayer Perceptron), Naive Bayes (NB), Destekçi Vektör Makine (SVM – Support Vector Machine) ve Rassal Orman (RF) algoritmalarıdır. Veri seti üzerinde uygulanan algoritmalar ve elde edilen sonuçlar Çizelge 5.3'te verilmiştir.

Çizelge 5.3'te ZeroR algoritmasının başarısı temel başarı olarak varsayılmıştır. Dolayısıyla, ZeroR başarısı altındaki başarı oranının algoritma için bir başarısızlık olarak kabul edilmelidir. Yapılan algoritma uygulamaları için her iki test yöntemi için böyle bir durum söz konusu değildir. ZeroR uygulamasının başarısı her iki test yöntemi için % 57,86'dır. Diğer algoritmaların başarısı bu değeri ne kadar yükselttikleri ile

Çizelge 5.3 2010-2021 yılları kaza verilerinin sınıflandırılmasında kullanılan algoritmalar ve başarı parametreleri

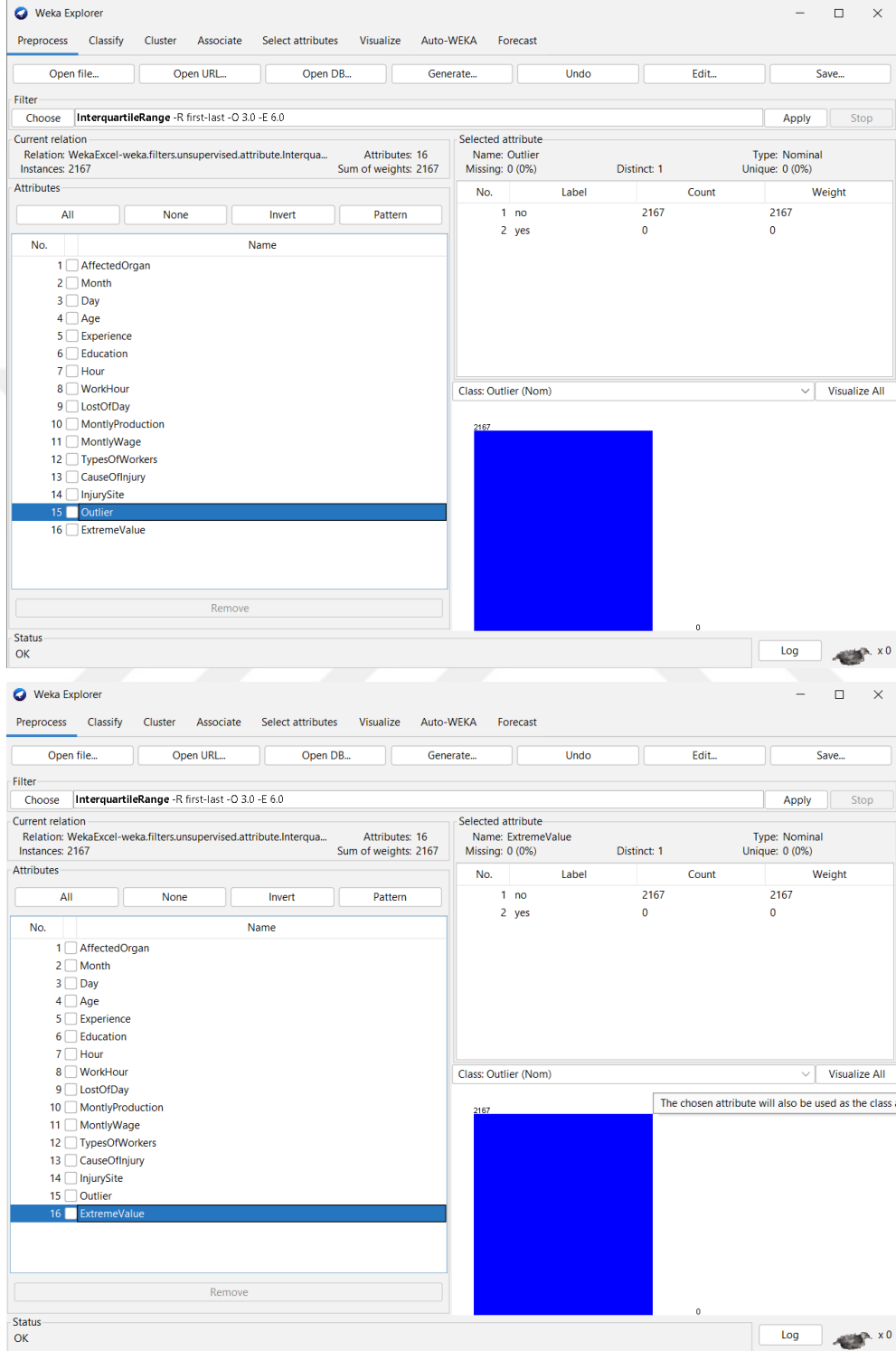
Algoritma Test Yöntemi										
Algoritma	Eğitim Seti					Çapraz Doğrulama				
	Değiştirilen Ayarlar	Doğru Sınıflandırılan (Toplam kaza sayısı: 2167)	Kappa	MAE	RMSE	Değiştirilen Ayarlar	Doğru Sınıflandırılan (Toplam kaza sayısı: 2167)	Kappa	MAE	RMSE
ZeroR	-	1254 (% 57,86)	0	0,113	0,2374	-	1254 (% 57,86)	0	0,113	0,2374
NB	-	1489 (% 68,71)	0,4207	0,0796	0,2048	-	1394 (% 64,32)	0,3368	0,0861	0,2189
J48	Budama yapıldı	1867 (% 86,16)	0,7653	0,0313	0,1251	-	1425 (% 65,76)	0,3266	0,093	0,2217
kNN	-	2167 (% 100)	1	0,0008	0,0015	Komşu sayısı: 9	1332 (% 61,46)	0,163	0,0927	0,2226
MLP	2 gizli katman (34, 30)	2074 (% 95,70)	0,9299	0,0099	0,0804	2 gizli katman (34, 30)	1236 (% 57,03)	0,2791	0,08	0,2665
SVM	-	1553 (% 71,66)	0,4522	0,1501	0,2658	-	1432 (% 66,08)	0,3424	0,1512	0,2677
RF	-	2167 (% 100)	1	0,034	0,0797	-	1388 (% 64,05)	0,2866	0,0931	0,215

ölçülebilir. Eğitim seti test yönteminde kNN ve RF algoritmaları oldukça iyi sonuçlar vermiş ve %100 başarı sağlamışlardır. Veri madenciliğinde böyle durumlarda (% 100 başarı) ilk akla gelen husus, algoritmalar açısından aşırı öğrenme ve veriyi ezberleme (overfitting) problemidir. Ancak, her %100 başarı bu şekilde değerlendirilmemelidir. Eğer veri setinde aşırı öğrenmeye ve veriyi ezberlemeye sebep olacak gürültülü veri (noisy data) yoksa, bu başarı kuşku ile karşılanacak bir durum olmayabilir. Bunun için WEKA programında veri setinin gürültülü veri içerip içermediği kontrol edilebilir. WEKA programında çeyrekler açıklığı (interquartile range) filtresi kullanılarak veri setindeki aykırı değerler (outliers) ve aşırı değerler (extreme values) belirlenmeye çalışılır. Yapılan filtreleme işleminde, bu değerlerin varlığına ilişkin olumlu sonuç alınır, veri setinden bu değerler atılarak gürültülü veri temizliği yapılmış olur. Çalışmanın bu aşamasında kullanılan veri setine çeyrekler açıklığı filtresi uygulandığında her iki değer adına olumlu sonuç çıkmamıştır (aykırı ve aşırı değer sayısı 0'dır). Bu da veri setinin içinde anlamsız veri olmadığını göstermektedir. Bu analiz sonucuna ilişkin WEKA programı arayüzleri Şekil 5.4'te verilmiştir.

Eğitim seti test yönteminde, diğer başarı parametreleri de incelenmiştir. WEKA programında standart olarak verilen başarı ölçütlerinden Kappa istatistiği, MAE ve RMSE değerlerine bakıldığında en başarılı algoritmanın kNN algoritması olduğu görülmektedir. Başarı ölçütlerinden Kappa istatistiği algoritma ile sınıflandırılan (tahmin edilen) değerlerin, gerçek değerlerle ne kadar örtüştüğünü ve Kappa ne kadar 1'e yakınsa algoritmanın o kadar başarılı olduğunu göstermektedir. MAE ve RMSE değerleri ise sınıflandırma işleminde yapılan hatayı belirtmektedir ve bu değerlerin 0'a yakın olması başarıyı artırmaktadır.

Çizelge 5.3'te görüldüğü gibi, veri seti üzerinde uygulanan algoritmalar için çapraz doğrulama test yöntemi de kullanılmıştır. Elde edilen sonuçlar eğitim seti test yönteminde olduğu gibi, ZeroR uygulaması temel alınarak karşılaştırıldığında diğer algoritmaların genel olarak daha başarılı sonuç verdiği görülmektedir. Uygulanan algoritmalar arasında SVM algoritması, veri setinin %66,08'ini doğru sınıflandırarak en iyi sonucu vermiştir. Çapraz doğrulama test yöntemiyle yapılan uygulamalarda da diğer başarı ölçütleri (Kappa istatistiği, MAE ve RMSE) değerlendirildiğinde, Kappa istatistiğine göre SVM

algoritması, MAE değerine göre MLP algoritması ve RMSE değerine göre de RF algoritması daha iyi sonuç vermiştir.



Şekil 5.4 Uygulanan çeyrekler açıklığı filtresi sonucu veri setindeki aykırı ve aşırı değerler

Eğitim seti ve çapraz doğrulama test yöntemleri kullanılarak yapılan algoritma uygulamaları sonucunda elde edilen bazı hata matrisleri sırasıyla Şekil 5.5 ve 5.6'da verilmiştir.

ZeroR											kNN										
a	b	c	d	e	f	g	h	i	j	k	a	b	c	d	e	f	g	h	i	j	k
0	0	0	0	0	37	0	0	0	0	0	a	37	0	0	0	0	0	0	0	0	0
0	0	0	0	0	16	0	0	0	0	0	b	0	16	0	0	0	0	0	0	0	0
0	0	0	0	0	26	0	0	0	0	0	c	0	0	26	0	0	0	0	0	0	0
0	0	0	0	0	20	0	0	0	0	0	d	0	0	0	20	0	0	0	0	0	0
0	0	0	0	0	17	0	0	0	0	0	e	0	0	0	0	17	0	0	0	0	0
0	0	0	0	0	1254	0	0	0	0	0	f	0	0	0	0	0	1254	0	0	0	0
0	0	0	0	0	358	0	0	0	0	0	g	0	0	0	0	0	0	358	0	0	0
0	0	0	0	0	263	0	0	0	0	0	h	0	0	0	0	0	0	0	263	0	0
0	0	0	0	0	24	0	0	0	0	0	i	0	0	0	0	0	0	0	0	24	0
0	0	0	0	0	74	0	0	0	0	0	j	0	0	0	0	0	0	0	0	0	74
0	0	0	0	0	78	0	0	0	0	0	k	0	0	0	0	0	0	0	0	0	78

RF											J48										
a	b	c	d	e	f	g	h	i	j	k	a	b	c	d	e	f	g	h	i	j	k
37	0	0	0	0	0	0	0	0	0	0	a	34	0	0	1	1	0	1	0	0	0
0	16	0	0	0	0	0	0	0	0	0	b	2	12	1	1	0	0	0	0	0	0
0	0	26	0	0	0	0	0	0	0	0	c	1	4	20	0	1	0	0	0	0	0
0	0	0	20	0	0	0	0	0	0	0	d	3	1	0	14	2	0	0	0	0	0
0	0	0	0	17	0	0	0	0	0	0	e	1	2	3	0	11	0	0	0	0	0
0	0	0	0	0	1254	0	0	0	0	0	f	0	0	0	0	0	1224	11	15	0	2
0	0	0	0	0	0	358	0	0	0	0	g	2	0	0	1	0	43	304	6	0	2
0	0	0	0	0	0	0	263	0	0	0	h	0	0	0	0	0	77	23	161	0	2
0	0	0	0	0	0	0	0	24	0	0	i	0	0	0	0	0	5	8	3	8	0
0	0	0	0	0	0	0	0	0	74	0	j	0	0	0	0	0	13	17	2	3	39
0	0	0	0	0	0	0	0	0	0	78	k	0	0	0	0	1	18	13	3	0	3

a: Yerüstü Karo, b: Yerüstü Tumba, c: Yerüstü Kuyu, d: Yerüstü Lavvar, e: Yerüstü Atölye, f: Yeraltı Ayak, g: Yeraltı Galeri, h: Yeraltı Üretim Bacası, i: Yeraltı Tumba, j: Yeraltı Kuyu, k: Yeraltı Nefeslik

Şekil 5.5 Eğitim seti test yöntemine göre bazı algoritmaların hata matrisleri

ZeroR											SVM										
a	b	c	d	e	f	g	h	i	j	k	a	b	c	d	e	f	g	h	i	j	k
0	0	0	0	0	37	0	0	0	0	0	a	19	2	2	4	5	4	0	0	0	1
0	0	0	0	0	16	0	0	0	0	0	b	4	5	6	1	0	0	0	0	0	0
0	0	0	0	0	26	0	0	0	0	0	c	11	7	4	1	0	0	2	0	0	1
0	0	0	0	0	20	0	0	0	0	0	d	6	0	3	8	1	1	1	0	0	0
0	0	0	0	0	17	0	0	0	0	0	e	4	1	1	2	6	0	2	0	1	0
0	0	0	0	0	1254	0	0	0	0	0	f	0	0	1	0	1	1222	25	4	1	0
0	0	0	0	0	358	0	0	0	0	0	g	0	0	1	1	1	169	160	8	3	14
0	0	0	0	0	263	0	0	0	0	0	h	0	0	0	0	0	234	23	5	0	1
0	0	0	0	0	24	0	0	0	0	0	i	0	0	0	0	0	7	15	0	0	2
0	0	0	0	0	74	0	0	0	0	0	j	0	0	0	0	1	31	37	0	1	3
0	0	0	0	0	78	0	0	0	0	0	k	0	0	0	0	0	57	17	1	0	3

NB											J48										
a	b	c	d	e	f	g	h	i	j	k	a	b	c	d	e	f	g	h	i	j	k
14	2	4	5	3	4	1	1	0	1	2	a	13	2	0	3	2	15	2	0	0	0
2	2	6	0	0	1	1	0	0	3	1	b	1	5	2	1	0	3	4	0	0	0
4	5	8	3	1	1	2	0	1	1	0	c	3	5	0	2	2	6	8	0	0	0
4	1	3	9	2	0	0	1	0	0	0	d	4	0	0	9	2	4	1	0	0	0
8	1	0	3	2	0	1	0	1	1	0	e	3	2	2	3	3	1	1	0	0	1
0	0	2	0	1	1178	64	5	3	0	1	f	0	0	0	1	0	1217	33	1	0	2
3	1	1	3	4	139	165	12	6	21	3	g	1	0	1	0	0	171	170	0	1	13
0	0	0	0	0	229	28	4	0	2	0	h	0	0	0	0	0	233	30	0	0	0
0	0	0	0	0	4	12	0	1	7	0	i	0	0	0	0	0	7	12	0	1	3
0	1	1	0	0	16	40	0	4	11	1	j	1	0	0	0	0	33	31	0	2	7
0	0	0	0	0	52	19	2	0	5	0	k	0	0	0	0	1	59	16	0	0	2

a: Yerüstü Karo, b: Yerüstü Tumba, c: Yerüstü Kuyu, d: Yerüstü Lavvar, e: Yerüstü Atölye, f: Yeraltı Ayak, g: Yeraltı Galeri, h: Yeraltı Üretim Bacası, i: Yeraltı Tumba, j: Yeraltı Kuyu, k: Yeraltı Nefeslik

Şekil 5.6 Çapraz doğrulama test yöntemine göre bazı algoritmaların hata matrisleri

İki test yöntemiyle oluşturulan hata matrisleri incelendiğinde, sınıflamaların yerüstü ve yeraltı olarak genel anlamda doğru yapıldığı görülmektedir. Yeraltında meydana gelen kazalarda ise, çok sayıda kazanın doğrudan üretim yapılan ayaklarda meydana gelmesi sınıflamalarda kaza yeri ile ilgili eğilimin üretim yapılan ayak ve üretim bacalarına, benzer açıdan galerilere kaymasına sebep olmuştur. Şekil 5.5 ve 5.6'da verilen hata matrislerinde öz niteliklerin doğru sınıflandırılıp sınıflandırılmadığını göstermektedir. Örneğin; Şekil 5.5'te verilen J48 algoritması hata matrisine göre Yeraltı Ayakta gerçekleşen kazaların 1224'ü doğru, 30'u da yanlış (11'i Yeraltı Galeri, 15'i Yeraltı Üretim Bacası, 2'si Yeraltı Kuyu ve 2'si Yeraltı Nefeslik) sınıflandırılmıştır. Diğer öz niteliklerle ilgili de benzer yorumlar yapılabilir. Yine aynı şekilde, Şekil 5.6'da J48 hata matrisinde 1217 doğru, 37 yanlış (33'ü Yeraltı Galeri, 1'i Yeraltı Üretim Bacası, 2'si Yeraltı Kuyu, 1'i ise Yerüstü Lavvar) sınıflama mevcuttur.

5.3.2 2010-2020 yılları kaza verilerinden 2021 yılı kaza verilerinin tahmini

Çalışmanın ikinci aşaması, işletme kaza verilerinin tahminine yönelik veri madenciliği algoritmalarının başarı seviyelerinin belirlenmesi amacıyla yapılan çalışmaları kapsamaktadır. Bu amaç doğrultusunda, 2010-2020 yılları arasında meydana gelen 1959 yaralanmalı kaza eğitim seti olarak belirlenmiş, bu veri setiyle eğitilen algoritmaların 2021 yılına ait 208 yaralanmalı kazayı tahmin başarısı ölçülmüştür. Bu aşamada da veri setinden kaza yerlerinin tespitine yönelik çalışmalar yapılmıştır.

Kaza yerlerinin tahmininde farklı algoritmaların tahmin başarısı belirlenmiştir. WEKA programında bu amaç için kullanılan test yöntemi, test setini kullan (supplied test set) yöntemidir. Bu yöntemde 2010-2020 yılları arası kaza verileri eğitim seti, 2021 yılı test seti olarak seçilmiştir. İlk olarak temel teşkil etmesi için ZeroR algoritması veri setine uygulanmış, daha sonra bu algoritmadaki başarının yükseltilmesi amacıyla diğer algoritmalar çalışılmıştır. Elde edilen sonuçlar Çizelge 5.4'te verilmiştir.

Çizelge 5.4 incelendiğinde, 2021 yılında oluşan kazaların yerleri ile ilgili yapılan analizde temel başarı seviyesi olarak kabul edilen ZeroR algoritmasının üzerine kısıtlı bir şekilde çıkılabildiği görülmüştür. Uygulanan algoritmalar içerisinde SVM algoritması en yüksek

başarıyı vermiş ve 2021 yılı kaza yerlerini %67,78 oranında tahmin edebilmiştir. Bu tahmine yönelik hata matrisi Şekil 5.7’de verilmiştir.

Çizelge 5.4 2021 yılı kaza verilerinin tahminine yönelik algoritmalar ve başarı parametreleri

Algoritma	Test Set Kullan				
	Değiştirilen Ayarlar	Doğru Sınıflandırılan (Toplam kaza sayısı: 208)	Kappa	MAE	RMSE
ZeroR	-	130 (% 62,5)	0	0,1103	0,231
J48	-	138 (% 66,34)	0,2391	0,0925	0,2213
kNN	Komşu sayısı:9	135 (% 64,90)	0,1484	0,0941	0,2187
SVM	-	141 (% 67,78)	0,3069	0,1514	0,2683
RF	Tekrarlama: 500	134 (%64,42)	0,1739	0,1003	0,2188

SVM												
a	b	c	d	e	f	g	h	i	j	k		
3	0	0	1	0	0	0	0	0	0	0		a
1	0	1	0	0	0	0	0	0	0	0		b
1	0	0	1	1	0	0	0	0	1	0		c
0	0	0	1	0	0	0	0	0	0	0		d
1	1	0	1	1	0	0	0	0	0	0		e
0	0	0	0	0	129	0	0	1	0	0		f
0	0	1	0	0	14	7	0	0	1	0		g
0	0	0	0	0	23	1	0	0	0	0		h
0	0	0	0	0	1	0	0	0	0	0		i
0	0	0	0	0	3	5	0	0	0	0		j
0	0	0	0	0	5	2	0	0	0	0		k

a: Yerüstü Karo, b: Yerüstü Tumba, c: Yerüstü Kuyu, d: Yerüstü Lavvar, e: Yerüstü Atölye, f: Yeraltı Ayak, g: Yeraltı Galerisi, h: Yeraltı Üretim Bacası, i: Yeraltı Tumba, j: Yeraltı Kuyu, k: Yeraltı Nefeslik

Şekil 5.7 SVM algoritması hata matrisi

SVM hata matrisinde Yeraltı Ayak öz niteliği sınıflaması incelendiğinde, Yeraltı Ayakta meydana gelen 129 kaza doğru tahmin edilmiştir. Sadece 1 kaza yanlış tahmin edilmiştir. Bunun yanında, Yeraltı Galeride meydana gelen 23 kazadan 7’si doğru tahmin edilebilmiştir.

5.3.3 2010-2021 yılları kaza verilerine yönelik birliktelik kuralları analizi

Veri madenciliği, günümüzde veri barındıran her alanda kullanımı yaygınlaşan bir bilimdir. Genel olarak, bir veri setinden bilgi çıkarma ve daha sonra kullanılmak üzere

anlaşılır bir yapıya dönüştürme sürecidir. Bu süreç içerisinde veriler toplanır, veri seti düzenlenir ve gürültülü veriler temizlenir. Daha sonra veriler üzerinde madencilik için bazı yöntem ve algoritmalar kullanılır. Birliktelik kuralları madenciliği (association rules mining), veri madenciliğinin en önemli yöntemlerinden biridir. Bu yöntemde veri setinde öz nitelikler arasında gizli kalmış ve ilginç ilişkiler, kurallar biçiminde ortaya çıkarılmaktadır.

Birliktelik kuralları analizinde çeşitli algoritmalar kullanılmaktadır. Bunlardan en yaygın olanları Apriori ve Tahminci Apriori (Predictive Apriori) algoritmalarıdır. Apriori algoritması ilk olarak Agrawal ve Srikant (1994) tarafından ortaya konmuştur. Algoritma veri setinde en sık görünen öz nitelikleri bulmakta ve bu niteliklerle birliktelik kuralları oluşturmaktadır. Öncelikle veri seti içerisinde tek elemanlı alt gruplar oluşturulmakta ve belirlenen destek ölçütünden daha düşük sıklıkla görünen alt gruplar elenmektedir. Kalan gruplar arasında ikili alt gruplar oluşturulup destek ölçütünü karşılamayan gruplar elenmektedir. Bu şekilde diğer aşamalarda da gruplar oluşturulup destek ölçütünü sağlamayan gruplar elenmektedir. Algoritma ile kurallar belirlenirken uygulamada belirlenen güven ölçütü (confidence) dikkate alınmaktadır. Bu güven ölçütüne göre oluşturulan birliktelik kuralları sıralanmaktadır.

Tahminci Apriori algoritması ise, Scheffer (2001) tarafından Apriori algoritmasının geliştirilmesi ile ortaya konmuştur. Tahmin doğruluğu (predictive accuracy) ölçütü kullanılarak yüksek doğrulukla kurallar elde edilmektedir. Diğer uygulamaya göre tüm kurallar taranmadığı için algoritma daha hızlı sonuç vermekte ve daha güvenilir kurallar oluşturulmaktadır.

Çalışmanın bu aşamasında ATİM’de 2010-2021 yıllarında meydana gelen yaralanmalı kazalar için birliktelik kuralları analizi yapılmıştır. Bu analizde en çok kullanılan Apriori ve Tahminci Apriori yöntemleri kullanılmıştır. Apriori yönteminde kaza yerleri için oluşturulan ilk 20 kural Çizelge 5.5’te verilmiştir.

Çizelge 5.5 Apriori yöntemiyle kaza yerleri için oluşturulan birliktelik kuralları

Kural No	Birliktelik Kuralı				Güven Değeri
	Kuralın Öz Nitelikleri	Kural Öz Nitelikleri Kaza Sayısı	Hedef Öz Nitelik	Hedef Öz Niteliği Kaza Sayısı	
1	Deneyim=0-5 Çalışma Saati=4-8 Çalışan Görevi=Pano-Ayak Üretim İşçisi Kaza Sebebi=Yeraltı Göçük	283	Kaza Yeri=Yeraltı Ayak	234	0,83
2	Deneyim=0-5 Çalışma Saati=4-8 Kayıp İş Günü Sayısı=4-28 Çalışan Görevi=Pano-Ayak Üretim İşçisi	332	Kaza Yeri=Yeraltı Ayak	271	0,82
3	Deneyim=0-5 Eğitim Durumu=İlkokul Çalışma Saati=4-8 Çalışan Görevi=Pano-Ayak Üretim İşçisi	339	Kaza Yeri=Yeraltı Ayak	276	0,81
4	Deneyim=0-5 Eğitim Durumu=İlkokul Çalışma Saati=4-8	345	Kaza Yeri=Yeraltı Ayak	278	0,81
5	Vardiya=0-8 Çalışma Saati=4-8 Çalışan Görevi=Pano-Ayak Üretim İşçisi	275	Kaza Yeri=Yeraltı Ayak	221	0,80
6	Deneyim=0-5 Çalışma Saati=4-8 Aylık Yevmiye=5000-5999 Çalışan Görevi=Pano-Ayak Üretim İşçisi	271	Kaza Yeri=Yeraltı Ayak	217	0,80
7	Deneyim=0-5 Çalışma Saati=4-8 Çalışan Görevi=Pano-Ayak Üretim İşçisi	577	Kaza Yeri=Yeraltı Ayak	462	0,80
8	Yaş=26-30 Deneyim=0-5 Çalışma Saati=4-8 Çalışan Görevi=Pano-Ayak Üretim İşçisi	291	Kaza Yeri=Yeraltı Ayak	233	0,80
9	Yaş=26-30 Çalışma Saati=4-8 Çalışan Görevi=Pano-Ayak Üretim İşçisi	300	Kaza Yeri=Yeraltı Ayak	240	0,80
10	Deneyim=0-5 Çalışan Görevi=Pano-Ayak Üretim İşçisi Kaza Sebebi=Yeraltı Göçük	458	Kaza Yeri=Yeraltı Ayak	365	0,80
11	Deneyim=0-5 Çalışma Saati=4-8 Kaza Sebebi=Yeraltı Göçük	314	Kaza Yeri=Yeraltı Ayak	248	0,79
12	Deneyim=0-5 Çalışma Saati=4-8 Aylık Yevmiye=5000-5999	289	Kaza Yeri=Yeraltı Ayak	227	0,79

Çizelge 5.5 (Devam) Apriori yöntemiyle kaza yerleri için oluşturulan birliktelik kuralları

13	Deneyim=0-5 Kayıp İş Günü Sayısı=4-28 Kaza Sebebi=Yeraltı Göçük	288	Kaza Yeri=Yeraltı Ayak	226	0,78
14	Çalışma Saati=4-8 Çalışan Görevi=Pano-Ayak Üretim İşçisi Kaza Sebebi=Yeraltı Göçük	417	Kaza Yeri=Yeraltı Ayak	323	0,77
15	Deneyim=0-5 Eğitim Durumu=İlkokul Kayıp İş Günü Sayısı=4-28 Çalışan Görevi=Pano-Ayak Üretim İşçisi	348	Kaza Yeri=Yeraltı Ayak	268	0,77
16	Deneyim=0-5 Çalışma Saati=4-8 Kayıp İş Günü Sayısı=4-28	372	Kaza Yeri=Yeraltı Ayak	286	0,77
17	Kayıp İş Günü Sayısı=4-28 Çalışan Görevi=Pano-Ayak Üretim İşçisi Kaza Sebebi=Yeraltı Göçük	378	Kaza Yeri=Yeraltı Ayak	290	0,77
18	Çalışan Görevi=Pano-Ayak Üretim İşçisi Kaza Sebebi=Yeraltı Göçük	686	Kaza Yeri=Yeraltı Ayak	525	0,77
19	Yaş=31-35 Çalışma Saati=4-8 Çalışan Görevi=Pano-Ayak Üretim İşçisi	332	Kaza Yeri=Yeraltı Ayak	254	0,77
20	Yaş=26-30 Kayıp İş Günü Sayısı=4-28 Çalışan Görevi=Pano-Ayak Üretim İşçisi	285	Kaza Yeri=Yeraltı Ayak	218	0,76

Çizelge 5.5'te görüldüğü gibi güven ölçütü açısından en yüksekten başlayarak birliktelik kuralları sıralanmıştır. Bu birliktelik kuralları, kaza sayıları dikkate alınarak kaza yeri ile en fazla örtüşen öz nitelikleri bir araya getirmiştir. Örneğin; Kural 1'de Yeraltı Ayakta, 0-5 yıl arasında deneyimli olan Pano-Ayak Üretim İşçisinin vardiyasının 4-8 saatleri arasında meydana gelebilecek Göçükte kazalanması yüksek doğrulukla (%83) gerçekleşebilir. Yine Kural 2'ye göre, Yeraltı Ayakta vardiyasının 4-8 saatleri arasında kazalanan 0-5 yıl deneyimli Pano-Ayak Üretim İşçisi, meydana gelen kazadan dolayı 4-28 gün iş kaybına uğramaktadır. Bu birliktelik %82 oranında gerçekleşmektedir. Dolayısıyla oluşturulan birliktelik kurallarından işletme için önemli derin çözümlemeler ve bilgiler edinilebilir.

Apriori yöntemiyle kaza yeri için oluşturulan birliktelik kuralları incelendiğinde, içerdikleri öz nitelikler açısından en fazla Deneyim (0-5), Çalışma Saati (4-8) ve Çalışan Görevi (Pano-Ayak Üretim İşçisi) öz niteliklerinin bulunduğu görülmüştür. Kuralların

tamamında Kaza Yeri olarak Yeraltı Ayak yer almaktadır. Verilen 20 birliktelik kuralında öz nitelik ve alt sınıf dağılımları Çizelge 5.6’da verilmiştir.

Çizelge 5.6 Apriori yöntemiyle oluşturulan birliktelik kurallarında öz nitelik ve alt sınıf dağılımları

Öz Nitelik (Değişken)	Birliktelik Kural Sayısı	Sıklık (%)	Alt Sınıfı ve Bulunduğu Kural Sayısı
Etkilenen Organ	-	-	
Ay	-	-	
Gün	-	-	
Yaş	4	20	26-30 (3), 31-35 (1)
Deneyim	13	65	0-5 (13)
Eğitim Durumu	3	15	İlkokul (3)
Vardiya	1	5	0-8 (1)
Çalışma Saati	14	70	4-8 (14)
Kayıp İş Günü Sayısı	6	30	4-28 (6)
Aylık Üretim	-	-	
Aylık Yevmiye	2	10	5000-5999 (2)
Çalışan Görevi	15	75	Pano-Ayak Üretim İşçisi (15)
Kaza Sebebi	7	35	Yeraltı Göçük (7)

Çizelge 5.6 incelendiğinde, Kaza Yeri ile en fazla örtüşen öz niteliklerden Deneyim toplam 20 birliktelik kuralının %65’inde, Çalışma Saati %70’inde ve Çalışan Görevi %75’inde yer almaktadır. Diğer öz niteliklerin kurallar içinde bulunma yüzdeleri %0-35 arasında değişmektedir. En fazla örtüşen öz niteliklerin alt sınıflarına bakıldığında ise, 0-5 yıl arası Deneyim, Çalışma Saati olarak ilgili vardiyanın 4-8 saatleri arası ve Çalışan Görevi olarak da Pano-Ayak Üretim İşçisi dikkati çekmektedir. Kazaların kök nedenleri hakkında detaylı yorumlamaların yapılabilmesi için bu sonuçlardan yararlanılması gerekmektedir. Kurallar içinde, 0-5 yıl deneyime sahip işçilerin kazalanma durumlarının diğerlerine oranla daha fazla olduğu, kazaların genel olarak vardiyaların son saatleri olan 4-8 saatleri arasında gerçekleştiği ve özellikle Pano-Ayak Üretim İşçilerinin kazalandığı tespit edilmiştir.

Tahminci Apriori ile oluşturulan ilk 20 birliktelik kuralı Çizelge 5.7’de verilmiştir.

Çizelge 5.7 Tahminci Apriori yöntemiyle kaza yerleri için oluşturulan birliktelik kuralları

Kural No	Birliktelik Kuralı				Tahmin Doğruluğu
	Kuralın Öz Nitelikleri	Kural Öz Nitelikleri Kaza Sayısı	Hedef Öz Nitelik	Hedef Öz Niteliği Kaza Sayısı	
1	Deneyim=0-5 Çalışma Saati=4-8 Aylık Üretim =24000-25999 Aylık Yevmiye=6000-6999 Kaza Sebebi=Yeraltı Göçük	23	Kaza Yeri=Yeraltı Ayak	23	0,9804
2	Ay=Temmuz Deneyim=0-5 Aylık Üretim =20000-21999	22	Kaza Yeri=Yeraltı Ayak	22	0,9794
3	Ay=Mart Çalışma Saati=4-8 Çalışan Görevi=Pano-Ayak Üretim İşçisi Kaza Sebebi=Yeraltı Malzeme Taşıma-Kullanma	22	Kaza Yeri=Yeraltı Ayak	22	0,9794
4	Etkilenen Organ= Ayak Parmağı Vardiya=8-16 Çalışma Saati=4-8 Çalışan Görevi=Pano-Ayak Üretim İşçisi Kaza Sebebi=Yeraltı Göçük	22	Kaza Yeri=Yeraltı Ayak	22	0,9794
5	Deneyim=0-5 Vardiya=16-24 Çalışma Saati=4-8 Aylık Üretim =24000-25999 Kaza Sebebi= Yeraltı Göçük	22	Kaza Yeri=Yeraltı Ayak	22	0,9794
6	Gün=Pazartesi Deneyim=0-5 Vardiya=16-24 Çalışma Saati=4-8 Aylık Yevmiye=5000-5999 Çalışan Görevi=Pano-Ayak Üretim İşçisi	22	Kaza Yeri=Yeraltı Ayak	22	0,9794
7	Etkilenen Organ=Ayak Parmağı Gün=Çarşamba Vardiya=16-24 Çalışan Görevi=Pano-Ayak Üretim İşçisi	21	Kaza Yeri=Yeraltı Ayak	21	0,9782
8	Ay=Mart Yaş=26-30 Çalışan Görevi=Pano-Ayak Üretim İşçisi Kaza Sebebi= Yeraltı Göçük	21	Kaza Yeri=Yeraltı Ayak	21	0,9782
9	Gün=Salı Vardiya=16-24 Kayıp İş Günü=4-28 Çalışan Görevi=Pano-Ayak Üretim İşçisi Kaza Sebebi= Yeraltı Göçük	21	Kaza Yeri=Yeraltı Ayak	21	0,9782

Çizelge 5.7 (Devam) Tahminci Apriori yöntemiyle kaza yerleri için oluşturulan birliktelik kuralları

10	Etkilenen Organ=Bel Deneyim=0-5 Çalışma Saati=4-8 Çalışan Görevi=Pano-Ayak Üretim İşçisi	20	Kaza Yeri=Yeraltı Ayak	20	0,9768
11	Gün=Pazartesi Çalışma Saati=4-8 Çalışan Görevi=Pano-Ayak Üretim İşçisi Kaza Sebebi=Yeraltı Vurma- Çarpma	20	Kaza Yeri=Yeraltı Ayak	20	0,9768
12	Gün=Cumartesi Deneyim=0-5 Vardiya=0-8 Çalışma Saati=4-8	20	Kaza Yeri=Yeraltı Ayak	20	0,9768
13	Gün=Perşembe Yaş=26-30 Vardiya=8-16 Kaza Sebebi= Yeraltı Göçük	20	Kaza Yeri=Yeraltı Ayak	20	0,9768
14	Ay=Nisan Yaş=31-35 Deneyim=0-5 Aylık Yevmiye=5000-5999	19	Kaza Yeri=Yeraltı Ayak	19	0,9754
15	Ay=Nisan Yaş=31-35 Vardiya=16-24 Çalışan Görevi=Pano-Ayak Üretim İşçisi	19	Kaza Yeri=Yeraltı Ayak	19	0,9754
16	Etkilenen Organ=El Parmağı Yaş=31-35 Çalışma Saati=4-8 Aylık Üretim =20000-21999	18	Kaza Yeri=Yeraltı Ayak	18	0,9736
17	Etkilenen Organ=Ayak Parmağı Deneyim=6-10 Vardiya=16-24 Kaza Sebebi= Yeraltı Göçük	18	Kaza Yeri=Yeraltı Ayak	18	0,9736
18	Etkilenen Organ=Ayak Parmağı Çalışma Saati=4-8 Çalışan Görevi=Pano-Ayak Üretim İşçisi Kaza Sebebi=Yeraltı Vurma- Çarpma	18	Kaza Yeri=Yeraltı Ayak	18	0,9736
19	Ay=Nisan Vardiya=0-8 Aylık Yevmiye=5000-5999 Çalışan Görevi=Pano-Ayak Üretim İşçisi	18	Kaza Yeri=Yeraltı Ayak	18	0,9736
20	Ay=Temmuz Yaş=31-35 Deneyim=0-5 Çalışma Saati=4-8	18	Kaza Yeri=Yeraltı Ayak	18	0,9736

Tahminci Apriori ile yapılan birliktelik kuralları analizinde elde edilen sonuçlar (Çizelge 5.7) diğer yönteme göre daha keskin çıkarımlar sağlamaktadır. Örneğin; Kural 1’de Aylık Üretimin 24000-25999 ton, Aylık Yevmiyenin 6000-6999 olduğu şartlarda Ayakta, vardiyanın 4-8 saatleri arasında bir Göçük olması durumunda kazalanan işçinin Deneyiminin 0-5 yıl olma durumu %98’dir. Aynı şekilde, Kural 2’ye göre Temmuz ayında 20000-21999 ton Aylık Üretim gerçekleştirildiğinde Ayakta oluşacak bir kazada, 0-5 yıl Deneyimli bir işçinin kazalanma durumu %97,9’dur. Tahminci Apriori yöntemi ile daha kesin birlikteliklerin çıkarımı, daha yüksek doğrulukla yapılabilmektedir.

Tahminci Apriori yöntemiyle Kaza Yeri için oluşturulan birliktelik kuralları incelendiğinde, öz nitelikler açısından kurallar içerisinde daha düzgün bir yayılımın bulunduğu görülmüştür. Deneyim, Vardiya, Çalışma Saati, Çalışan Görevi ve Kaza Sebebi öz nitelikleri nispeten daha fazla kuralda yer almaktadır. Kuralların tamamında Kaza Yeri, Yeraltı Ayak olarak değerlendirilmiştir. Verilen 20 birliktelik kuralında öz nitelik ve alt sınıf dağılımları Çizelge 5.8’de verilmiştir.

Çizelge 5.8 Tahminci Apriori yöntemiyle oluşturulan birliktelik kurallarında öz nitelik ve alt sınıf dağılımları

Öz Nitelik (Değişken)	Birliktelik Kural Sayısı	Sıklık (%)	Alt Sınıfı ve Bulunduğu Kural Sayısı
Etkilenen Organ	6	30	Ayak Parmağı (4), Bel (1), El Parmağı (1)
Ay	7	35	Mart (2), Nisan (3), Temmuz (2)
Gün	6	30	Pazartesi (2), Salı (1), Çarşamba (1), Perşembe (1), Cumartesi (1)
Yaş	6	30	26-30 (2), 31-35 (4)
Deneyim	9	45	0-5 (8), 6-10 (1)
Eğitim Durumu	-	-	
Vardiya	10	50	0-8 (2), 8-16 (2), 16-24 (6)
Çalışma Saati	11	55	4-8 (11)
Kayıp İş Günü Sayısı	1	5	4-28 (1)
Aylık Üretim	4	20	20000-21999 (2), 24000-25999 (2)
Aylık Yevmiye	4	20	5000-5999 (3), 6000-6999 (1)
Çalışan Görevi	11	55	Pano-Ayak Üretim İşçisi (11)
Kaza Sebebi	10	50	Yeraltı Göçük (7), Yeraltı Malzeme Taşıma- Kullanma (1), Yeraltı Vurma-Çarpma (2)

Çizelge 5.8 incelendiğinde, Kaza Yeri öz niteliği ile en fazla örtüşen öz niteliklerden Deneyim toplam 20 birliktelik kuralının %45'inde, Vardiya %50'sinde, Çalışma Saati %55'inde, Çalışan Görevi %55'inde ve Kaza Sebebi %50'sinde yer almaktadır. Diğer öz niteliklerin kurallar içinde bulunma yüzdeleri %0-35 arasında değişmektedir. En fazla örtüşen öz niteliklerin alt sınıflarına bakıldığında ise, 0-5 yıl arası Deneyim, 16-24 Vardiyası, Çalışma Saati olarak ilgili vardiyanın 4-8 saatleri arası, Çalışan Görevi olarak Pano-Ayak Üretim İşçisi ve Kaza Sebebi olarak da Yeraltı Göçük alt sınıfları öne çıkmaktadır. Tahminci Apriori yöntemiyle oluşturulan birliktelik kurallarına göre, 0-5 yıl deneyime sahip işçilerin kazalanma durumlarının diğerlerine oranla daha fazla olduğu, kazalar açısından özellikle 16-24 vardiyasının önemli olduğu, kazaların genel olarak vardiyaların son saatleri olan 4-8 saatleri arasında gerçekleştiği, en fazla Pano-Ayak Üretim İşçilerinin kazalandığı ve kaza sebeplerinde yeraltında oluşan göçüklerin öne çıktığı belirlenmiştir. Bunların yanında, birliktelik kuralları çerçevesinde Aylık Üretim ve Aylık Yevmiye öz niteliklerinin de kaza kök nedenleri arasına girmesi ve kurallar arasında yer alması dikkat çekicidir.

6. SONUÇLAR VE ÖNERİLER

Tez kapsamında gerçekleştirilen çalışmalarda, ATİM’de 2010-2021 yılları arasında gerçekleşen 2167 yaralanmalı kazalar çeşitli veri madenciliği yöntemleriyle değerlendirilmiş ve aşağıda belirtilen sonuçlar elde edilmiştir.

- 1- İlk aşamada 2010-2021 yılları arasında gerçekleşen kazaların sınıflandırılması ve tahminine dayalı veri madenciliği algoritmaları, eğitim seti ve çapraz doğrulama yöntemleri kullanılarak uygulanmıştır. Veri madenciliği istatistiki ölçütlerine göre eğitim seti yönteminde en iyi sonucu kNN algoritması vermiştir. Bu algoritma verileri %100 doğru sınıflamıştır, aynı zamanda Kappa istatistiği 1’e eşit ve MAE, RMSE değerleri de 0’a oldukça yakındır.
- 2- Çapraz doğrulama yönteminde de SVM algoritması %68 başarı sağlamıştır. İstatistiksel ölçütlere göre de SVM algoritmasının diğer algoritmalar göre kabul edilebilir seviyede olduğu görülmüştür.
- 3- İkinci aşamada 2010-2020 yılları verilerinden 2021 yılı verilerinin tahminine dayalı çalışmalar yapılmıştır. Bu çalışmada da SVM algoritmasının en iyi sonuç verdiği, temel algoritma olarak kabul edebileceğimiz ZeroR algoritması sonuçlarına göre bir miktar daha iyi sonuç elde edildiği söylenebilir. Ancak bu kısımda daha derin ve kapsamlı analizlerin yapılması gerekmektedir.
- 4- Üçüncü aşamada ise 2010-2021 yılları kaza verilerine Apriori ve Tahminci Apriori yöntemleriyle birliktelik kuralları analizi yapılmıştır. Hedef öz nitelik olarak seçilen kaza yeri ile diğer öz niteliklerin birlikte bulunma durumları araştırılmıştır. Elde edilen sonuçlara göre yüksek güven ve doğruluk değerlerinde birliktelik kuralları oluşturulmuştur. Her iki yöntemle ilgili örnek olarak verilen 20’şer birliktelik kuralları incelendiğinde Apriori yöntemi %83-76 güven değeri aralığında kurallar üretirken, Tahminci Apriori yöntemi %98-97,3 tahmin doğruluğunda kurallar türetmiştir.
- 5- Apriori uygulamasına göre Tahminci Apriori yönteminin yüksek doğruluğu olan daha kesin kurallara odaklandığı görülmektedir.
- 6- Birliktelik kuralları madenciliğiyle kaza sebepleri, meydana geliş yerleri, kaza kökeni ve kazaya sebep sayılabilecek diğer öz nitelikler hakkında daha detaylı bir bilgi çıkarımı yapılabilmektedir.

- 7- Apriori yöntemiyle oluşturulan birliktelik kurallarında Deneyim, Çalışma Saati ve Çalışan Görevi öz nitelikleri öne çıkmaktadır.
- 8- Tahminci Apriori yöntemiyle oluşturulan birliktelik kurallarında Deneyim, Vardiya, Çalışma Saati, Çalışan Görevi ve Kaza Sebebi öz nitelikleri önem kazanmaktadır. Her iki yöntemle iş kazalarının değerlendirilmesi çalışmalarında, bu öz niteliklere dikkat çekilmesi gerekmektedir.

Bu tez çalışması kapsamında yapılan araştırmalar sonucunda elde edilen veriler ışığında, oluşan iş kazalarının değerlendirilmesi ve meydana gelebilecek iş kazalarının önlenmesi için çalışmanın yapıldığı işletmeye yönelik öneriler aşağıda sunulmuştur.

- 1- Güven ölçüsü açısından en yüksekten başlayarak 20 birliktelik kuralını incelediğimizde, Apriori yönteminde %70 ve Tahminci Apriori yönteminde ise %55 oranında kazanın meydana geliş saati olan vardiyanın 4-8 saatleri, en fazla örtüşen öz nitelik alt sınıfı olarak karşımıza çıkmaktadır. Vardiyanın 4. ve 8. saatleri arasında işçinin yorgunluk ya da verilen işi bitirme endişesi ile konsantrasyonunda bozulma söz konusu olabileceği düşünülmektedir, özellikle bu saatler arasında vardiya nezaretçileri tarafından daha yoğun kontrol edilmesinin ve vardiyada yapılacak işlerin planlamalarının çalışma saati içerisinde tamamlanacak şekilde, önceki veya mevcut vardiyada oluşabilecek aksamlar düşünülmektedir yapılmasının yararlı olacağı önerilmektedir.
- 2- Güven ölçüsü açısından en yüksekten başlayarak 20 birliktelik kuralını incelediğimizde, en fazla örtüşen diğer bir öz nitelik ise çalışanın deneyim süresidir. Apriori yönteminde %65 ve Tahminci Apriori yönteminde ise %45 oranında az deneyimli işçilerin, özellikle 0-5 yıl deneyime sahip işçilerin kazalandığı görülmektedir. Bunun yanında, yine işçilerin eğitim durumlarına bakıldığında ilkokul mezunlarının kazalandıkları görülmektedir. Bu nedenle, deneyim süresi az olan işçilere daha yoğun ve nitelikli eğitimler verilmesinin, az deneyim süresine sahip çalışanların daha deneyimli işçilerin yanına yedek olarak verilerek onların nezaretinde çalıştırılmasının yararlı olacağı önerilmektedir.
- 3- Güven ölçüsü açısından en yüksekten başlayarak 20 birliktelik kuralını incelediğimizde, yine en fazla örtüşen bir öz nitelik olan kaza sebebi içerisinde yeraltında oluşan göçükler karşımıza çıkmaktadır. Klasik üretim yöntemine göre,

mekanize sistemlerde tavan daha güvenli şekilde ve herhangi bir boşluğa fırsat vermeden tahkim edilmektedir. Üretim yöntemi olarak mekanize sistemlerin kullanılmasının ve tahkimat sistemi olarak kalkan tahkimatların kullanılmasının yararlı olacağı önerilmektedir.

- 4- Güven ölçüsü açısından en yüksekten başlayarak 20 birliktelik kuralını incelediğimizde, Tahminci Apriori yöntemi ile yapılan birliktelik kuralları analizinde Aylık Üretim ve Aylık Yevmiye öz niteliklerinin de kurallar içinde bulunduğu gözlemlenmiştir. Yapılacak üretimlerin iş planına uygun şekilde, iş sağlığı ve güvenliği çalışma kurallarının öncelikli şekilde uygulanarak yapılmasının yararlı olacağı önerilmektedir.

Yukarıda belirtilen sonuçlar itibariyle, günümüzde oldukça fazla kullanım alanı bulan veri madenciliğinin iş kazası verileri gibi büyük ve karmaşık verilerde kullanılabileceği görülmektedir. ATİM için yapılan bu çalışmanın derinleştirilmesi ve diğer işletmelere de yaygınlaştırılması, kaza oluşumları hakkında detay çıkarımların yapılmasını ve ileriye dönük alınabilecek net tedbirlerin önceden ortaya konmasını sağlayacaktır. Bu anlamda, tez kapsamında yapılan çalışmaların ilerletilmesi düşünülmeli ve özellikle madencilik kapsamında faaliyet gösteren işletmelerde yaygınlaştırılması desteklenmelidir.

7. KAYNAKLAR

- Ak M F, 2019, A Neuro-fuzzy-Based Multi-criteria Risk Evaluation Approach: A Case Study of Underground Mining. In: Al-Turjman, F. (eds) Artificial Intelligence in IoT, Transactions on Computational Science and Computational Intelligence, Springer, Cham, 167–205.
- Aldana W A, 2000, Data Mining Industry: Emerging Trends and New Opportunities, Massachusetts Institute of Technology, MSc Thesis, 183p, Cambridge, USA.
- Argüden Y, Erşahin B, 2008, Veri Madenciliği Veriden Bilgiye, Masraftan Değere, ARGE Danışmanlık Yayınları, No: 10, 73s, İstanbul.
- Bouckaert R R, Frank E, Hall M, Kirkby R, Reutemann P, Seewald A, Scuse D, 2018, WEKA Manual for Version 3-8-3, University of Waikato, 343p, Hamilton, New Zealand.
- Breiman L, 2001, Random forests, Machine Learning, 45, 5-32.
- Dessureault S, Sinuhaji A, Coleman P, 2007, Data Mining Mine Safety Data, Mining Engineering, 59(8), 64–70.
- Er B, 2018, Yeraltı Kömür Madenciliğindeki Temel Risklerin Hibrit Bir Risk Analiz Yöntemi İle Değerlendirilmesi, Dokuz Eylül Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, 161s, İzmir.
- Erdoğan H H, 2016, A Quantitative Risk Assessment Methodology For Occupational Accidents In Underground Coal Mines: A Case Of Turkish Hard Coal Enterprises, Middle East Technical University, The Graduate School Of Natural And Applied Sciences, Ph.D. Thesis, 278p, Ankara.
- Erkayaoğlu M, 2020, Madencilik Sektöründe İş Sağlığı Ve Güvenliği Yönetimi İçin Veri Entegrasyonu Uygulaması, Çukurova Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 35(1), 105-114.
- Gerassis S, Saavedra Á, Taboada J, Alonso E, Bastante F G, 2020, Differentiating Between Fatal and Non-fatal Mining Accidents Using Artificial Intelligence Techniques, International Journal of Mining, Reclamation and Environment, 34(10), 687–699.

- Haliloğlu T, 2000, Amasra Bölgesi Kömürlerinin Katkısız Briketlenmesi, Bülent Ecevit Üniversitesi, Fen Bilimleri Enstitüsü, Jeoloji Mühendisliği Anabilim Dalı, Yüksek Lisans Tezi, 118s, Zonguldak.
- Han J, Kamber M, 2006, Data Mining: Concepts and Techniques, 2nd Edition, Morgan Kaufmann Publishers, 772p, San Francisco.
- Hustrulid W A, Bullock R L (Ed.), 2001, Underground Mining Methods, Engineering Fundamentals and International Case Studies, Society for Mining, Metallurgy, and Exploration, Inc. (SME), 702p, Colorado.
- İphar M, Çukuluöz A K, 2020, Fuzzy risk assessment for mechanized underground coal mines in Turkey, International Journal of Occupational Safety and Ergonomics (JOSE), 26(2), 256–271.
- Martins F F, Miranda T F S, 2012, Estimation of the rock deformation modulus and rmr based on data mining techniques, Geotechnical and Geological Engineering, 30, 787-801.
- Mevsim R, 2016, Risk Assessment By Fault Tree Analysis of Methane Explosions in Turkish Hard Coal Enterprises Underground Mines, Middle East Technical University, The Graduate School Of Natural And Applied Sciences, M.Sc. Thesis, 125p, Ankara.
- Opokua M, Opoku S K, 2021, An Adaptive Neuro-Fuzzy Inference System for Predicting Survivability Rate in Underground Mining Accident, Conference Proceedings of International Conference on Applied Science and Technology ICAST 2020, 6(1), 250–265.
- Saltoğlu S, 1988, Madenlerde Yeraltı Üretim Yöntemleri, İstanbul Teknik Üniversitesi Yayınları, No:25, 188s, İstanbul.
- Sanmiquel L, Bascompta M, Rossell J M, Anticoi H F, Guash E, 2018, Analysis of Occupational Accidents in Underground and Surface Mining in Spain Using Data-Mining Techniques, International Journal of Environmental Research and Public Health, 15, Article Number:462.
- Shuangyue L, Li P, 2013, Analysis of Coal Mine Hidden Danger Correlation Based on Improved A Priori Algorithm, Fourth International Conference on Intelligent

- Systems Design and Engineering Applications (ISDEA), Hunan, China, 112-116.
- TBMM, 2023, Bartın'ın Amasra İlçesinde Meydana Gelen Maden Kazasının Tüm Yönleriyle Araştırılarak Benzer Kazaların Önlenmesine Yönelik Tedbirlerin Belirlenmesi Amacıyla Kurulan Meclis Araştırması Komisyonu Raporu, Sıra No:440, 566s.
- Tural M, 2019, TTK Amasra Müessesesi Damarlarının Gaz İçeriğinin Ve Tahmini Metan Gelirinin Belirlenmesi, Zonguldak Bülent Ecevit Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 171s, Zonguldak.
- Witten I H, Frank E, Hall M A, 2011, Data Mining Practical Machine Learning Tools and Techniques, Third Edition, Morgan Kaufmann Publishers, 665p, Burlington.

İnternet Kaynakları

- 1- <http://www.taskomuru.gov.tr/index.php?page=sayfagoster&id=103>, 13.08.2022
- 2- <http://www.taskomuru.gov.tr/index.php?page=sayfagoster&id=104>, 13.08.2022
- 3- <https://www.datascience-pm.com/crisp-dm-2/>, 07.09.2023
- 4- https://saedsayad.com/data_mining_map.htm, 14.10.2023
- 5- <https://saedsayad.com/zeror.htm>, 09.09.2023
- 6- https://saedsayad.com/naive_bayesian.htm, 09.09.2023
- 7- https://saedsayad.com/decision_tree.htm, 09.09.2023
- 8- https://saedsayad.com/k_nearest_neighbors.htm, 09.09.2023
- 9- https://saedsayad.com/artificial_neural_network.htm, 09.09.2023
- 10-TTK İSG Veri Tabanı, 18.05.2022