

Asansörlerde Meydana Gelen Arıza Sebeplerinin Sınıflandırılması

(Araştırma Makalesi)

Classifying the Causes of Failures in Elevators

Doi: 10.29023/alanyaakademik.632107

Meltem KARAATLI

Doç.Dr.,Süleyman Demirel Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü
meltemkaraatli@sdu.edu.tr

Orcid No: 0000-0002-7403-9587

Bu makaleye atıfta bulunmak için: Karaatlı, M. (2020). Asansörlerde Meydana Gelen Arıza Sebeplerinin Sınıflandırılması. *Alanya Akademik Bakış*, 4(3), Sayfa No.651-664.

ÖZET

Anahtar kelimeler:

Asansör,
Sınıflandırma
Algoritmaları,
Veri Madenciliği

Makale Geliş Tarihi:
11.10.2019
Kabul Tarihi:
02.09.2020

Hem işyeri hem de meskenlerde kullanılan asansörlerde güvenlik son derece önem arz etmektedir. Bu sebeple belirli periyotlarla asansörler kontrol edilmektedir. İlk kontrolü yapılan asansörler etiketlendirme işleminin ardından yönetmelikte belirtilen süreler dâhilinde son kontrolü yapılarak tekrar etiketlendirme işlemine tabii tutulur. Bu çalışmada son kontrol aşamasında yapılan etiketlendirmeye ilişkin sebepler çeşitli makine öğrenme algoritmalarından yararlanılarak sınıflandırılmış ve algoritmaların başarıları kıyaslanmıştır. Uygulamanın sonucunda İterative Classifier Optimizer ve Logitboost Algoritmaları en başarılı algoritmalar olarak belirlenmiştir.

ABSTRACT

Keywords:

Elevator,
Classification
Algorithms,
Data Mining

The earliest distinction that mankind has ever experienced throughout its Safety is highly important for elevators used in both workplaces and residential buildings. For this reason, the elevators are checked periodically. After the labeling process of initially inspected elevators, the final control takes place within the period specified in the regulation for re-labeling. In this study, the reasons for labeling in the last control stage were classified by using various machine learning algorithms and their successes were compared. Iterative Classifier Optimizer and Logitboost Algorithms were determined as the most successful algorithms.

1.GİRİŞ

Veri madenciliği istatistik, yapay zeka ve yapay zekayı kullanan makine öğrenimini birleştiren, veri dizinlerinden geçerli, mümkünse faydalı ve anlaşılır örüntülerin ortaya çıkartılabilmesi için gerçekleştirilen bir süreç olarak tanımlanabilir (Akpınar, 2017).

Veri madenciliği, nümerik veya kategorik özelliklere sahip verileri kullanarak pazarlama, finans, muhasebe, tıp ve mühendislik gibi çeşitli alanlarda kullanılmıştır. Veri madenciliği ile veri işleme, kümeleme (clustering), sınıflandırma (classification), ilkelleme (regression), veri ön işleme (data preprocessing), özellik seçimi (feature selection) gibi işlemler

gerçekleştirilebilmektedir (Şeker, 2013). Bu çalışmada da asansörlerde meydana gelen arızalara ilişkin sebeplerin sınıflandırılmasıyla ilgili bir uygulama yapılmıştır.

İşyeri ve meskenlerde kullanılan asansörler T.C. Sanayi ve Teknoloji Bakanlığı tarafından yetki verilen A tipi olarak adlandırılan muayene kuruluşlarınca belli aralıklarla muayene edilmektedir. Belirli aralıklarla yapılan kontrol sonuçları “kusursuz, hafif kusurlu, kusurlu ve güvensiz” olmak üzere dört şekilde değerlendirilir. Bu değerlendirmenin ardından (Asansör İşletme Bakım ve Periyodik Kontrol Yönetmeliği, www.resmigazete.gov.tr, 2015):

1. “Kusursuz” olarak adlandırılan asansöre, ilgili idare adına düzenli kontrolü yapan A tipi muayene kuruluşunca “yeşil renkli” bilgi etiketi yapıştırılır.
2. “Hafif kusurlu” olarak adlandırılan asansöre, ilgili idare adına düzenli kontrolü yapan A tipi muayene kuruluşunca “mavi renkli” bilgi etiketi yapıştırılır.
3. “Kusurlu” olarak tanımlanan asansöre, ilgili idare adına düzenli kontrolü yapan A tipi muayene kuruluşunca “sarı renkli” bilgi etiketi yapıştırılır.
4. “Güvensiz” olarak tanımlanan asansöre, ilgili idare adına düzenli kontrolü yapan A tipi muayene kuruluşunca “kırmızı renkli” bilgi etiketi yapıştırılır.
5. “Kırmızı renkli” bilgi etiketi yapıştırılan ve “güvensiz” olarak adlandırılan asansörün kullanımına bina sorumlusunca müsaade edilmez.

İlk kontrolü yapılan asansörler yukarıda bahsedilen etiketlendirme işlemin ardından yönetmelikte belirtilen süreler dikkate alınarak son kontrolü yapılır ve ardından tekrar etiketlendirme işlemine tabi tutulur.

Bu çalışmada son kontrol aşamasında yapılan etiketlendirme işlemine dair sebepler çeşitli makine öğrenme algoritmalarından yararlanılarak sınıflandırılmış ve algoritmaların başarıları kıyaslanmıştır.

2. LİTERATÜR TARAMASI

Literatür taraması yapıldığında çeşitli alanlarda veri madenciliğinde sınıflandırma ile ilgili birçok çalışmaya rastlamak mümkündür. Tablo 1.’de bu çalışmalardan bazılarının yer verilmiştir. Literatürde asansörlerle ilgili sınıflandırma çalışmasına rastlanılmamakla birlikte asansörlerle ilgili yapılmış çeşitli çalışmalar bulunmaktadır. Bu çalışmalardan bazıları yine Tablo 1.’de görülmektedir.

Tablo 1. Literatür Taraması

Çalışma Alanı	Yazar	Konu
<i>Veri madenciliğinde sınıflandırma ile ilgili çalışmalar</i>	Kaygın ve ark., (2016)	İşletmelerin Finansal Başarılı ve Başarısız Olma Durumlarının Veri Madenciliği Ve Lojistik Regresyon Analizi İle Tahmini
	Akçetin ve Çelik, (2014)	İstenmeyen Elektronik Posta (Spam) Tespitinde Karar Ağacı Algoritmalarının Performanslarının Kıyaslanması
	Tekin ve ark., (2011)	K-Means ve YSA Temelli Hibrit Bir Model İle Epileptik Eeg İşaretlerinin Sınıflandırılması
	Kaya ve ark., (2017)	C4.5 Karar Ağacı Temelli Öznitelik Seçimi İle Video-Okülografi (VOG) Sinyallerinden Diyabetik

		Retinopati Hastalığının Sınıflandırılması
	Daş ve Türkoğlu, (2014)	DNA Dizilimlerinin Sınıflandırılmasında Karar Ağacı Algoritmalarının Karşılaştırılması
	Koyuncuğil ve Özgülbaş, (2008)	İMKB’de İşlem Gören Kobi’lerin Güçlü ve Zayıf Yönleri: Chaid Karar Ağacı Uygulaması
	Cihan ve Kalpsız, (2015)	Öğrenci Proje Anketlerini Sınıflandırmada En Başarılı Algoritmanın Belirlenmesi
	Aşlıyan ve Günel, (2010)	Metin İçerikli Türkçe Dokümanların Sınıflandırılması
	Vernakar ve ark.,(2018).	Makine Öğrenme Yaklaşımlarını Kullanarak Motor Şanzıman Arızası Teşhisi
	Adewumi ve Akinyelu, (2016)	E-Posta Kimliği Tespiti İçin Melez Ateş Böceği ve Destek Vektör Makine Sınıflandırıcısı,
	Pei ve ark., (2019)	Hassas Karar Ağacı Yaklaşımı İle Çin Nüfusundaki Potansiyel Tip Iı Diyabetin Belirlenmesi,
	Trivedi ve Panigrahi, (2018)	Spam Sınıflandırması,
	Hussain ve ark., (2018)	Bir E-Öğrenim Sisteminde Öğrenci Katılımı Tahminleri ve Öğrenci Ders Değerlendirme Puanlarına Etkileri
	Wu ve ark., (2014)	Makine Öğrenme Tekniklerini Kullanarak İki Aşamalı Kredi Oranı Tahmini
	Krichene (2017)	Kredi Riski Değerlendirme İçin Naive Bayes Sınıflandırıcısı Kullanma Metodolojisi
	Salehzadeh (2017)	Takipçiler Hangi Tip Liderleri Tercih Eder? Bir Karar Ağacı Yaklaşımı
Çalışma Alanı	Yazar	Konu
<i>Asansörlerle ilgili yapılmış çalışmalar</i>	Bayar (2016)	Asansörlere Uygulanan Periyodik Kontrolleri İş Güvenliği Açısından İncelenmesi
	Karahan (2016)	Asansör Bakım İşlerinde Risklerin Değerlendirilmesi Açısından Çözüm Önerileri
	Bolat (2006)	Asansör Kontrol Sistemlerini Genetik Algoritma İle Simülasyonunu
	Kıral ve Gören Kıral (2008)	Ray Deformasyonlarının Asansör Kabini Titreşimlerine Etkisinin İncelenmesi
	Koç (2009)	Asansör Kılavuz Ray Konsollarının Gerilme Analizi
	Bedir (2007)	Çift Yönlü Asansör Fren Bloklarının Modellenmesi
	Dursun ve Saygın (2006)	Bir Asansör Tahrik Sistemi İçin Bulanık Mantık Denetimli Anahtarlamalı Relüktans Motor Sürücüsü Tasarımı
	Adak ve ark. (2012)	Asansör Simülatörünün Ürettiği Sonuçlar Üzerinde Bir K-Means++ Kümeleme Çalışması İle Trafik Türünün Tahmini
	Cıcık (2016)	Asansör Montaj İşlerinde İş Sağlığı ve Güvenliği Risk Faktörlerinin Değerlendirilmesi
	Kavlak (2006)	Elektrikli Asansörlerle Hidrolik Asansörlerin Karşılaştırılması
	Zarikas ve ark. (2013)	Yunanistan'da Asansör Kazalarının İstatistiksel Araştırması
	Park ve Yang	Asansör Arıza, Kaza ve Bakımlarının Düzenliliği

	(2010)	
	Çeri (2018)	Asansör Bakımı İçin Risk Bazlı Denetim Uygulaması
	İmrak ve Fetvacı (2004)	Asansörlerde Bakım Onarım Yetersizliğinden Kaynaklanan Kazalar
	Ünal ve Aykaç (2010)	Yapı İşlerinde Asansör Kazaları ve Güvenlik Önlemleri

3. VERİ MADENCİLİĞİ İLE SINIFLANDIRMA

Sınıflandırma (classification), bir veri seti verildiğinde veri üzerinde bilinen çeşitli gruplar arasında veriyi paylaşır. Veri seti eğitim seti ve test seti şeklinde ayrılır. Çeşitli sınıflandırma algoritmaları, eğitim kümesi verileri sayesinde dağılım şeklini öğrenerek sınıfının tanımı verilmeyen test verisi verildiğinde doğru şekilde sınıflandırılmaya uğraşır (Şeker, 2013). Makine öğreniminde bu tür öğrenmeye denetimli öğrenme (supervised learning) adı verilmektedir. Bu çalışmanın analiz aşamasında ücretsiz olarak sunulan Waikato Üniversitesinde Java programlama diliyle ortaya konulan WEKA (Waikato Environment for Knowledge Analysis) programı kullanılacaktır. Uygulamada kullanılacak olan algoritmalar hakkında kısaca bilgi verilecektir.

3.1. Uygulamada Kullanılan Sınıflandırma Algoritmaları

Bu bölümde uygulamada kullanılan makine öğrenme algoritmalarına kısaca yer verilmiştir.

Karar Ağaçları (Decision Tree), sınıflandırma ve tahminde kullanılan veri madenciliği tekniklerindedir. Karar ağacının ağaca benzer bir yapısı vardır. İç düğümler iki veya daha fazla çocuk düğümlerinden oluşmaktadır. Karar ağacının iç düğümleri veri kümesinin özelliklerini temsil ederken dalları da bu özelliklerin değerlerini temsil etmektedir. Her yaprak veri kümesiyle ilgili bir sınıf içermektedir. Karar ağacı, bir eğitim seti yardımıyla eğitilmektedir. Karar ağacı bilinmeyen sınıf etiketleri ile bir veri setini sınıflandırmak için kullanılmaktadır. Karar ağacı, karar vermek için öncelikle bilgiyi işlemektedir. Ağaç, çocuk düğümlerde girdi özelliklerini en iyi ayıran nitelikleri belirleyebilmek için veri setinden oluşturulmuştur (Hussain ve ark, 2018).

J48, C45 karar ağacı algoritmasının bir uygulamasıdır. J48 özyinelemeli bir bölünme ve fethetme yaklaşımını içeren yukarıdan aşağıya doğru bir strateji izlemektedir (Lemoyné ve Mastroianni, 2017). Bilgi kazancını maksimum yapmak ve bit olarak bilginin ölçülmesi, ilgili bilgi kazancının artırılması için entropi hesaplayarak en kısa (öz) karar ağacını elde etmek için bilgi kazancı teorisine başvurmuştur. J48, veriler üzerinde basit bir sınıflandırma modeli oluşturmak için anlamlı olmayan zayıf dalları keserek budama işlemi gerçekleştirmektedir (Dağ ve Varol, 2013). J48, en yüksek bilgi kazancını içeren özellik üzerinden verileri bölerek oluşturulan düğümlerden nihai bir karar verir (Akçetin ve Çelik, 2014).

Random Tree (Rasgele Ağaç) Algoritması ise, oluşturulan ağaçlar dizisinden tesadüfi olarak seçilmiş özellikleri alarak bir ağaç oluşturmaktadır. Tesadüfilik (random) ağaçlar dizisi içinde her bir ağacın eşit seçilme şansını olduğunu ifade etmektedir. **Random Forest (Rasgele Orman) Algoritması** ise, karar ağacı öğreniminde tesadüfi öznelik seçimi kullanan, budanmamış sınıflandırma veya regresyon ağaçlarının bir topluluğudur (Akınar, 2017). Veri setinde her bir düğümde yer alan tesadüfi seçilen özelliklerin en iyisinin seçimini gerçekleştirerek mevcut düğümleri dallara böler. Her veri kümesi asıl veri setinden yer

değiştirmeli olarak üretilmektedir. Rasgele özellik seçimi yoluyla ağaçlar geliştirilir ve budama işleminin olmaması sebebiyle bu algoritmanın diğerlerine göre daha çabuk ve daha doğru sınıflandırma yaptığı söylenmektedir (Akçetin ve Çelik, 2014).

Bayes Ağları ise, Bayes Teoremi diye bilinen istatistiksel prensipleri kullanarak, önceki bilgilerden yola çıkarak yeni sınıf etiketleri oluşturmaktadır. Uygulamada dikkate alınan **Naive Bayes** sınıflandırma teknikleri ise araştırılan özelliklerin bağımsızlığı varsayımına dayanarak her sınıf etiketini şartlı olasılıkla tahmin etmeye çalışmaktadır. Naive Bayes sınıflandırma basit olmasına rağmen birçok yöntemle karşılaştırıldığında iyi bir performansı bulunmaktadır (Wu ve ark, 2014). Naive Bayes sınıflandırıcısı eğitim için kısa hesaplama süresi vermektedir (Gopalakrishnan ve Sengottuvelan, 2016). Naive Bayes sınıflandırma algoritması belli olan değişkenler setinin belli bir gruba ait üyelik olasılıklarını tahminleyen istatistiksel sınıflandırıcılar olarak adlandırılır (Kılınç ve ark, 2016).

Uygulamada dikkate alınan **K-En Yakın Komşu (IBk) Algoritması** ise, 1950'lerden günümüze kadar kullanılan köklü ve sağlam bir algoritma olarak bilinmektedir. K-en yakın komşu algoritması örnek tabanlı bir öğrenme yöntemidir. Algoritma parçalı doğrusal karar sınırlarının toplamını birleştiren bir yaklaşımdır. Belirlenmiş eğitim örneklerine dayanarak, K- en yakın komşular yeni örneğe en çok benzeyen eğitim kümesini araştırmaktadır (Lemoyne ve Mastroianni, 2017). Uygulamada kullanılan diğer bir algoritma olan **K-Star Algoritması** ise, örnek tabanlı bir sınıflandırıcı olarak bilinmektedir. Entropi tabanlı bir mesafe işlevi kullanması nedeniyle diğer örnek tabanlı öğrencilerden farklı olarak görülmektedir.

Uygulamada dikkate alınan **Çok katmanlı Algılayıcı Modeli (Multi Layer Perceptron)** sinir sisteminin modellenmesi sonucu ortaya çıkan Yapay Sinir Ağlarında (YSA) kullanılan bir algoritmadır. YSA, paralel çalışabilme ve öğrenebilme yetenekleri bakımından biyolojik sinir sistemi özelliği göstermektedir (Şen, 2004). Birçok problem türü için kullanılabilen yapay sinir ağları özellikle sınıflandırma problemleri için tercih edilmektedir. Yapay sinir ağlarında kullanılan çok katmanlı algılayıcı modeli (ÇKA-MLP) gruplandırma, tanıma ve genellemeye ihtiyaç duyan problemler için geliştirilmiş algoritmadır. Bu model Delta Öğrenme Kuralı denilen bir öğrenme yöntemi kullanarak ağı beklenen çıktısı ile ortaya çıkan çıktı arasında meydana gelen hatayı minimum seviyeye indirmek suretiyle hatayı ağı yaymaktadır. Çok katmanlı ağlarda değişkenler seti girdi katmanından ağı verilerek, gizli katman aracılığıyla çıktı katmanına ulaşarak, ağı verilen girdi setine karşılık olarak ağı verdiği yanıt dışarıya iletilir (Öztemel, 2003).

Uygulamada kullanılan **Lojistik Regresyon (Logistic Regression)** ise, bağımlı değişkenin dağılımı hakkında herhangi bir varsayım içermemektedir. Bağımsız değişkenlerin normal dağılıma uyması ve bağımlı değişkenle doğrusal bir ilişkiye sahip olması ve grupların eşit varyansa sahip olması şartı taşıması gerekmektedir. Ayrıca bağımsız değişkenler kesikli, sürekli ve ikili değişken karışımı şeklinde olabilmektedir. Lojistik regresyonda iki veya daha fazla bağımlı değişken olabileceği gibi bunlar sıralı (nominal) ya da sıralı olmayabilir (Tabachnick ve Fidell, 2015).

Makine öğreniminde ensemble öğrenim (Ensemble Learning) esasına dayanan bazı algoritmalar bulunmaktadır. Bu algoritmalar elde edilen sonuçların performansını artırmak amacıyla farklı öğrenme algoritmalarının bir arada kullanılması ile ortaya çıkmışlardır. Bir işletme yönetiminin stratejik kararlar vermesi sırasında kurullarına danışması ve oylama ile

sonuçların belirlenmesi bu tip öğrenmenin bir örneği olarak verilebilir. Özetle oylama ile karar verilmesi esasına dayanmaktadır. Bu amaçla farklı algoritmalar geliştirilmiştir (Akpınar, 2017). Uygulamada dikkate alınan **Iterative Classifier Optimizer Algoritması** Çapraz doğrulama kullanarak verilen iteratif sınıflandırıcının iterasyon sayısını optimize etmektedir. **Bagging Algoritması** regresyon yöntemini kullanarak, **Classification Via Regression Algoritması** varyansı azaltarak, **LogitBoost Algoritması** ise, bir regresyon şeması yardımıyla sınıflandırma yaparak çok sınıflı problemleri çözebilmektedir. **Random Comittee Algoritmasında** ise, aynı veriye dayalı her bir temel sınıflandırıcı farklı bir rasgele sayı tohumu kullanılarak oluşturulur. Nihai tahmin her bir sınıflandırıcının ürettiği tahminlerin ortalamasıdır (WEKA2007,weka.sourceforge.net).

3.2. Model Başarım Ölçütleri

Model başarımını değerlendirirken dikkate alınan ölçütler Doğruluk Oranı(Correctly Classified Instances), Kesinlik (Precision), Duyarlılık (Recall), F-ölçütü ve Kappa istatistiğidir (Kappa Statistic). Modelin başarılı olması, doğru sınıf olarak adlandırılan örnek sayısı ve yanlış sınıf olarak adlandırılan örnek sayısı ile ilgilidir. Test sonucunda sınıflandırma modelinin başarısı hata matrisindeki doğru sınıflandırılan değişkenler ile yanlış sınıflandırılan değişkenlerin sayısı ile hesaplanır. Hataların yer aldığı matriste, sıralar test kümesinde yer alan örneklerle ilgili gerçek değerleri, sütunlarda modelin ortaya koyduğu tahminlemeyi gösterir. Tablo.2'de iki sınıflı bir veri kümesinde oluşturulmuş bir modelin hata matrisi verilmiştir. Bu hata matrisinde ana köşegen olarak adlandırılan köşegen doğru tahminlenmiş örnek sayılarını gösterirken, ana köşegen dışında yer alan matris elemanları hatalı olan değerleri ortaya koyar. TP (doğru olumlu) ve TN (doğru olumsuz) değerleri doğru sınıflandırılmış örnek sayısı olarak tanımlanmıştır. Yanlış olumlu (FP), aslında 0 (olumsuz) sınıftayken 1 (olumlu) olarak tahminlenmiş örneklerin sayısını ifade eder. Yanlış olumsuz (FN) ise 1 (olumlu) sınıftayken 0 (olumsuz) olarak tahminlenmiş örneklerin sayısını gösterir.

Tablo 2. Hata Matrisi

		Öngörülen sınıf	
		Sınıf=1	Sınıf=0
Doğru sınıf	Sınıf=1	a	b
	Sınıf=0	c	d

Doğruluk Oranı (Correctly Classified Instances): Doğruluk oranı, doğru olarak adlandırılan örnek sayısının (TP +TN), toplam örnek sayısına (TP+TN+FP+FN) oranlanmasıyla bulunur. Bu oranın 1'e tamamlayanı hata oranı olarak adlandırılır. 1 ve 2 numaralı eşitlikte görülmektedir.

$$Doğruluk = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$\text{Hata Oranı} = \frac{TP + FN}{TP + FP + FN + TN} \quad (2)$$

Kesinlik (Precision): Doğru sınıflandırılmış olumlu örneklem sayısının sınıfı olumlu olarak öngörölmüş toplam örneklem sayısına oranıdır. [0,1] Aralığında bir değer alır. 3 numaralı eşitlikte görölmektedir.

$$\text{Kesinlik} = \frac{TP}{TP + FP} \quad (3)$$

Duyarlılık (Recall): Doğru sınıflandırılmış pozitif örneklem sayısının gerçek sınıfı pozitif olan tüm örneklem sayısına oranıdır. Duyarlılık, gerçek pozitiflik oranı (True Positive Rate) olarak da adlandırılır. [0,1] Aralığında bir değer alır. 4 numaralı eşitlikte görölmektedir.

$$\text{Duyarlılık} = \frac{TP}{TP + FN} \quad (4)$$

F-Ölçütü: Çoğunlukla kesinlik ve duyarlılık arasında zıt bir ilişki vardır ki birinin değerini arttırmak diğerinin değerini düşürebilir. Bu nedenle daha kesin ve duyarlı sonuçlar elde etmek için her iki ölçütün harmonik ortalaması olan F-ölçütü kullanılır. 5 numaralı eşitlikte görölmektedir.

$$F \text{ Ölçütü} = \frac{2 * \text{Duyarlılık} * \text{Kesinlik}}{\text{Duyarlılık} + \text{Kesinlik}} \quad (5)$$

Kappa İstatistiği (Kappa Statistic): Kappa istatistiği ya da değeri beklenen ve gözlenen değerlerin karşılaştırıldığı sayısal bir değerdir. Bu değer, tesadüfi şans faktörünü de hesaba kattığından daha az yanıltıcıdır. Beklenen ve gözlenen doğruluk hesabı Kappa istatistiğinde aynı anda kullanılmakta ve karışıklık matrisi üzerinden kolaylıkla belirlenebilmektedir. Bu yöntemde satırdaki değerler gerçek değerleri, sütundaki değerler ise tahmin edilen değerleri göstermek üzere gözlenen doğruluk oranından beklenen doğruluk oranı çıkartılıp elde edilen bu değer 1-beklenen doğruluk oranına bölünmektedir. Kappa istatistiğinin yorumlanması konusunda kesin standart bir yorum olmasa da genel olarak 0.00-0.20 arası düşük, 0.21-0.40 arası kayda değer, 0.41-0.60 arası orta düzeyde; 0.61-0.80 arası önemli ve 0.81-1.00 arası mükemmel olarak tanımlanmaktadır (Aksu ve Doğan, 2019).

4. UYGULAMA

Bu bölümde çalışmanın amacı, çalışmada kullanılan veriler ve değişkenler son olarak da bulgular ve değerlendirmeler yer almaktadır.

4.1. Amaç

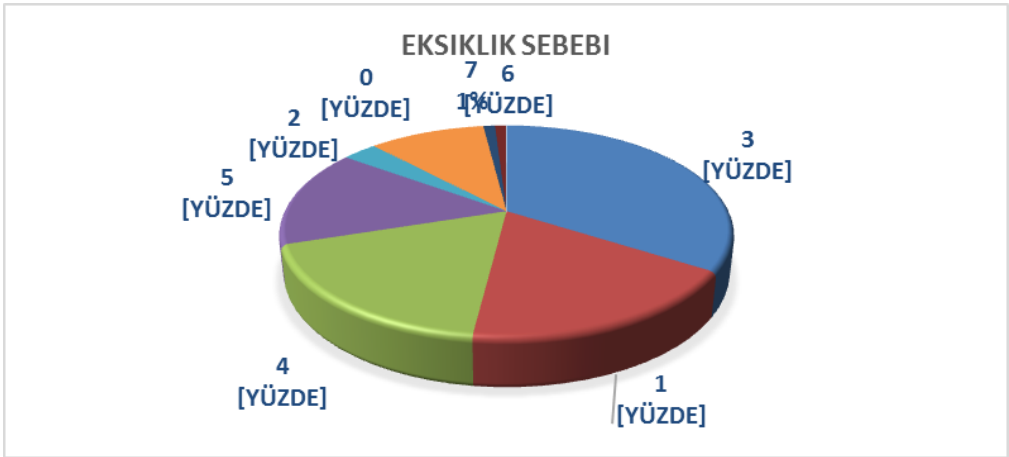
İşyeri ve meskenlerde kullanılan asansörler T.C. Sanayi ve Teknoloji Bakanlığı tarafınca görevlendirilen yetkilendirilen A tipi muayene kuruluşları tarafından düzenli olarak muayene edilmektedir. Düzenli olarak yapılan kontrol sonuçları “kusursuz, hafif kusurlu, kusurlu ve

güvensiz” olarak dört şekilde değerlendirilmektedir. İlk kontrolü yapılan asansörlerin etiketlendirme işleminin ardından yönetmelikte belirtilen süreler dikkate alınarak son kontrolü yapılır ve ardından tekrar etiketlendirme işlemi yapılır. Bu çalışmada, son kontrol aşamasında yapılan etiketlendirme işlemine dair sebepler çeşitli makine öğrenme algoritmalarından yararlanılarak sınıflandırılmış ve algoritmaların başarısı kıyaslanmıştır.

4.2. Araştırmada Kullanılan Veriler ve Değişkenler

Bu çalışmada Kayseri ilindeki A tipi muayene kuruluşu ile görüşülerek 100 adet binaya ilişkin veri temin edilmiştir. Elde edilen veriler asansörün montaj yılı, hızı (m/sn), seyir mesafesi(m), kapasitesi (kişi), fren tipi (kaymalı, ani tampon, ani), ilk güvenlik durumu (kırmızı, sarı mavi, yeşil), son güvenlik durumu (kırmızı, sarı mavi, yeşil), eksiklik sebebi (0-bina yönetimi kaynaklı, 1-firma kaynaklı, 2-inşaat firması kaynaklı, 3-problem yok, 4-firma ve inşaat firması kaynaklı, 5-firma ve bina yönetimi kaynaklı, 6-bina yönetimi ve inşaat firması kaynaklı, 7-bina yönetimi, firma ve inşaat firması kaynaklı) ait bilgilerdir.

Çalışmada asansörlerle ilgili arızalar Şekil 1.’de yer alan grafikte görülmektedir. Asansörlerin %34’ünde problem olmadığı, %18’inde firma kaynaklı, %18’inde firma ve inşaat firması kaynaklı, %15’inde firma ve bina yönetimi kaynaklı, %10’unda bina yönetimi kaynaklı, %3’ünde inşaat firması kaynaklı, %1’inde bina yönetimi ve inşaat firması kaynaklı ve son olarak %1’inde bina yönetimi, firma ve inşaat firması kaynaklı arızaların olduğu saptanmıştır.



Şekil 1. Arıza Sebeplerine İlişkin Daire Grafiği

Asansörlerin ilk kontrol aşaması tamamlandıktan sonra yönetmelikte belirtilen süreler dikkate alınarak son kontrol aşaması sürecinde etiketlendirme işlemi yapılır ve bu etiketlendirme işlemine dair sebepler belirtilir. Bu çalışmada belirtilen sebeplerin tahmini için bir sınıflandırma çalışması yapılmıştır. Sözü edilen sebepler (0-bina yönetimi kaynaklı, 1-firma kaynaklı, 2-inşaat firması kaynaklı, 3-problem yok, 4-firma ve inşaat firması kaynaklı, 5-firma ve bina yönetimi kaynaklı, 6-bina yönetimi ve inşaat firması kaynaklı, 7-bina yönetimi, firma ve inşaat firması kaynaklı) olmak üzere sekiz sınıfa ayrılmaktadır.

Veri madenciliği analizinde değerlendirilecek olan veriler WEKA Programına uygun olması için “arff formatında” hazır hale getirilmiştir. Değişkenlerden montaj yılı, seyir mesafesi, kapasitesi değişkenleri nümerik, fren tipi, ilk güvenlik durumu, son güvenlik durumu ve eksiklik sebepleri değişkenleri nominal değişkenlerdir. Nümerik yapıda olan verilerde normalizasyon işlemi WEKA programında yapıldıktan sonra analize tabi tutulmuştur. Analiz işleminde eğitim aşaması tamamlandıktan sonra sınıf etiketleri oluşturmak için “10 kat çapraz doğrulama” yapılmıştır. Bu yöntem veri setinin 10 parçaya ayrılması ve her parçanın bir defa test grubuna, geriye kalanların öğrenme grubuna atılmasıyla yapılır. Daha sonra sınıflandırmada en başarılı algoritmaları değerlendirilebilmek için Doğruluk Oranı (Correctly Classified Instances), Kappa İstatistiği (Kappa statistic), Duyarlılık (Recall), Kesinlik (Precision) ve F ölçütü model başarımlarını ölçütleri göz önüne alınmıştır.

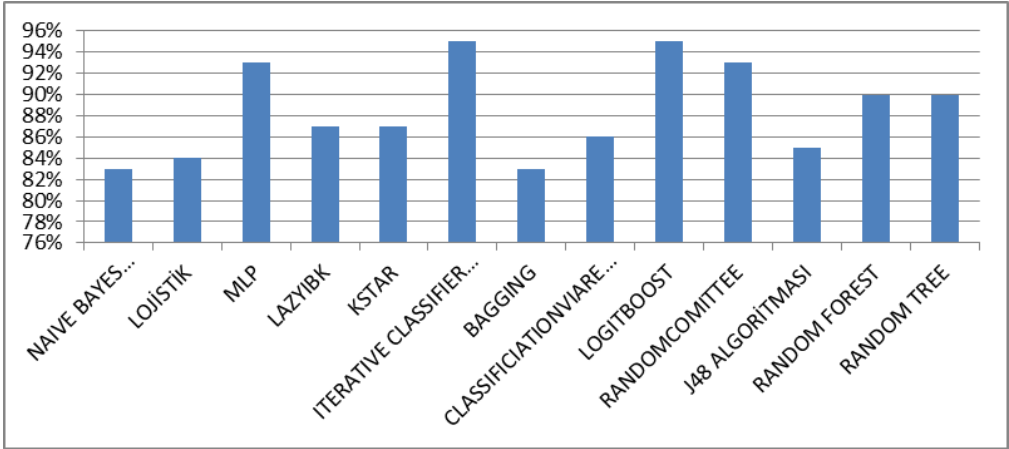
5. BULGULAR VE DEĞERLENDİRME

WEKA programı kullanılarak uygulanan algoritmaların sonuçları Tablo.3’de verilmiştir. Karşılaştırılacak algoritmaların programda yer alan parametreleri kullanılarak analiz gerçekleştirilmiştir.

Tablo 3. Algoritmaların Model Başarım Ölçütleri

	<i>Doğruluk Oranı (Correctly Classified Instances)</i>	Kappa İstatistiği (Kappa Statistic)	<i>Duyarlılık (Recall)</i>	<i>Kesinlik (Precision)</i>	F ölçütü
NAIVE BAYES ALGORİTMASI	%83	0.7827	0.845	0.830	0.828
LOJİSTİK EGRESYON	%84	0.7959	0.822	0.840	0.830
MLP	%93	0.9108	0.934	0.930	0.929
LAZYİBK	%87	0.8347	0.880	0.870	0.870
KSTAR	%87	0.8353	0.869	0.870	0.865
ITERATIVE CLASSIFIER OPTIMIZER	%95	0.9362	0.953	0.950	0.949
BAGGING	%83	0.7822	0.802	0.830	0.806
CLASSIFICATION VIA REGRESSION	%86	0.82	0.817	0.860	0.837
LOGITBOOST	%95	0.9362	0.953	0.950	0.949
RANDOMCOMITTEE	%93	0.911	0.942	0.930	0.934
J48 ALGORİTMASI	%85	0.8079	0.812	0.850	0.827
RANDOM FOREST	%90	0.8736	0.916	0.900	0.904
RANDOM TREE	%90	0.8727	0.912	0.900	0.903

Tablo 3.’de yer alan uygulama sonuçlarına göre algoritmaların performansları incelendiğinde *Doğruluk Oranı (Correctly Classified Instances)* yani doğru sınıflandırma oranı açısından bakıldığında %95 oranında Iterative Classifier Optimizer ve Logitboost Algoritmaları birinci sırada yer almıştır. Arkasından %93 oranında MLP ve Randomcommittee Algoritması yer almaktadır. Random Forest ve Random Tree Algoritmaları %90 oranında doğru sınıflandırarak üçüncü sırayı almıştır. Diğer algoritmalar ise %83 ile %87 arasında doğru sınıflandırma oranına sahiptir. Şekil 2.’de ise sınıflandırma algoritmalarının doğruluk oranlarını gösteren grafik görülmektedir.



Şekil 2. Sınıflandırma Algoritmalarının Doğruluk Oranları

Kappa istatistiği yapılan tahminin doğruluk ölçüsünü gösterir. Bu istatistiğin sonuçları incelendiğinde ise 0.9362 değerine sahip Logitboost ve Iterative Classifier Optimizer Algoritmaları birinci sırayı alırken, arkasından Randomcommittee Algoritması 0.911 değeriyle ikinci sırada, 0.9108 değeriyle MLP üçüncü sıradadır. Random Forest ve Random Tree Algoritmaları 0.8736 ve 0.8727 değerlerini alarak dördüncü ve beşinci sıraya yerleşmişlerdir. Diğer algoritmalar ise 0.7822 ve 0.8357 aralığında farklı değerler almışlardır. Kapa değeri 0.81 ile 1 arasında ise çok iyi, 0.61-0.8 aralığında ise iyi olarak yorumlanmaktadır. Dolayısıyla Logitboost, Iterative Classifier, Randomcommittee, MLP, Random Forest ve Random Tree Algoritmalarının çok iyi, diğer algoritmaların ise iyi oldukları söylenebilir.

Duyarlılık (Recall), Kesinlik (Precision) ve F ölçütü değerleri incelendiğinde sırasıyla 0.953, 0.950 ve 0.949 değerleriyle Iterative Classifier Optimizer ve Logitboost Algoritmaları ilk sırada yer almıştır. MLP 0.934, 0.930 ve 0.929 değerleriyle ikinci sırada yer alırken Random Forest Algoritması 0.916, 0.900 ve 0.904 değerleriyle üçüncü sırada, Random Tree Algoritması 0.912, 0.900 ve 0.903 değerleriyle dördüncü sıraya yerleşmiş. Diğer algoritmalarda yüksek denilebilecek oranlar almışlardır. Bu oranların ne kadar 1'e yakınsa o derecede başarılı bir performans gösterdiği söylenebilir. Dolayısıyla sözü geçen algoritmaların başarılı olduğu yorumu yapılabilir.

İlk üçte yer alan algoritmaların doğru tahminlediği örnek sayıları Tablo 4.'de yer almaktadır. Bu tablo WEKA programının ekran görüntüsünde yer alan hata matrisinin ana köşegeninde yer alan değerleri içermektedir. Hata matrisinin ana köşegeni doğru tahminlenmiş örnek sayılarını gösterir.

Tablo 4. İlk üçte Yer Alan Algoritmaların Doğru Tahminlediği Örnek Sayısı

Algoritmalar	0	1	2	3	4	5	6	7	Toplam
Iterative Classifier Optimizer	10	18	2	34	18	13	0	0	95
Logitboost	10	18	2	34	18	13	0	0	95

MLP	10	18	2	34	17	12	0	0	93
Randomcommittee	10	18	2	34	17	12	0	0	93

6. SONUÇLAR

İşyeri ve meskenlerde kullanılan asansörler T.C. Sanayi ve Teknoloji Bakanlığı tarafınca yetki verilen A tipi muayene kuruluşlarınca düzenli olarak muayene edilmektedir. Bu değerlendirme sonucunda ilk kontrol aşamasının arkasından yapılan son kontrol aşamasında etiketlendirme işlemi yapılır. Yapılan bu etiketlendirmeye dair sebepler ortaya konulur. Bu çalışmada da asansörlerde meydana gelen arızalara ilişkin ortaya konulan sebeplerin sınıflandırılması ile ilgili olarak bir uygulama yapılmış ve çeşitli algoritmaların performansları karşılaştırılmıştır.

Çalışmanın sonucunda model başarımları ölçütleri Doğruluk Oranı (Correctly Classified Instances), Kappa istatistiği, Duyarlılık (Recall), Kesinlik (Precision) ve F ölçütü dikkate alınarak algoritmalar değerlendirilmiştir. Uygulamada ücretsiz bir yazılım olarak sunulan Waikato Üniversitesinde Java programlama diliyle ortaya konulmuş olan WEKA (Waikato Environment for Knowledge Analysis) programı kullanılmıştır. Değerlendirmenin sonucunda model başarımları ölçütleri açısından Iterative Classifier Optimizer ve Logitboost Algoritmaları birinci sırada yer almışlardır. MLP ikinci sırada ve Randomcommittee Algoritması da üçüncü sırada başarılı bir performans göstermişlerdir.

Literatür taramasında gerek veri madenciliği gerekse asansörlerle ilgili birçok çalışmaya rastlanılmış ancak bu çalışmanın bir benzerine rastlanılmamıştır. Bu anlamda çalışmanın literatüre katkı sağlayabileceği düşünülmektedir.

KAYNAKÇA

- ADAK, M.F, DURU, N, DURU, H.T. (2012). “Asansör Simülasyonunun Ürettiği Sonuçlar Üzerinde Yapılan K-Means++ Kümeleme Çalışması İle Trafik Türünün Tahmini”. ELECO '2012 Elektrik-Elektronik ve Bilgisayar Mühendisliği Sempozyumu, 29 Kasım-1 Aralık, 2012, Bursa, Türkiye, 825-829.
- AEWUMI, O.A, AKINYELU, A.A. (2016). “Hybrid Firefly And Support Vector Machine Classifier For Phishing E-Mail Detection”. Kybernetes, 45(6): 977-994.
- AKÇETİN, E, ÇELİK, U. (2014). “İstenmeyen Elektronik Posta (Spam) Tespitinde Karar Ağacı Algoritmalarının Performans Kıyaslaması”. İnternet Uygulamaları ve Yönetimi, 5(2): 44-56.
- AKPINAR, H. (2017). Veri Madenciliği Veri Analizi, Papatya Yayıncılık, İkinci Baskı, İstanbul.
- AKSU, G, DOĞAN, N. (2019). “Veri Madenciliğinde Kullanılan Bir Analiz Programı: WEKA”. Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 10(1): 80-95.

- Asansör İşletme Bakım ve Periyodik Kontrol Yönetmeliği.
<http://www.resmigazete.gov.tr/eskiler/2015/06/20150624-1.htm> , 12. 03.2019.
- AŞLIYAN, R, GÜNEL, K. (2010). “Metin İçerikli Türkçe Dokümanların Sınıflandırılması”. Akademik Bilişim’10 - XII. Akademik Bilişim Konferansı Bildirileri, 10-12 Şubat, 2010, Muğla, Türkiye, 529-535.
- BAYAR, U. (2016). İş Ekipmanı Olarak Kullanılan Asansörlere Uygulanan Periyodik Kontrollerin İş Güvenliği Açısından İncelenmesi. İş Sağlığı ve Güvenliği Uzmanlık Tezi. Çalışma ve Sosyal Güvenlik Bakanlığı İş Sağlığı ve Güvenliği Genel Müdürlüğü, Ankara, Türkiye.
- BEDİR, S. (2007). Çift Yönlü Asansör Fren Bloklarının Modellenmesi ve Sonlu Elemanlarının Analizi. Yayınlanmamış Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, Türkiye.
- BOLAT, B. (2006). Asansör Kontrol Sistemlerinin Genetik Algoritma İle Simülasyonu. Yayınlanmamış Doktora Tezi, İstanbul Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, Türkiye.
- CICIK, S. (2016). Asansör Montaj İşlerinde İş Sağlığı ve Güvenliği Risk Faktörlerinin Değerlendirilmesi. İş Sağlığı ve Güvenliği Uzmanlık Tezi. Çalışma ve Sosyal Güvenlik Bakanlığı İş Sağlığı ve Güvenliği Genel Müdürlüğü, Ankara, Türkiye.
- CİHAN, P, KALIPSIZ, O. (2015). “Öğrenci Proje Anketlerini Sınıflandırmada En Başarılı Algoritmanın Belirlenmesi”. Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi, 8(1): 41-49.
- ÇERİ, G. (2018). “Asansörlerde Bakım Onarım Yetersizliğinden Kaynaklanan Kazalar”. Mühendis ve Makine Dergisi, 52(623): 67-73.
- DAĞ, B, VAROL, A. (2013). “2D:4D Sayısal Parmak Oranına Göre Bireylerin Kişilik Durumlarının Sınıflandırılması”. 1st International Symposium on Digital Forensics and Security (ISDFS’13), 20-21 Mayıs, 2013, Elazığ, Türkiye, 44-48.
- DAŞ, B, TÜRKÖĞLU, İ. (2014). “DNA Dizilimlerinin Sınıflandırılmasında Karar Ağacı Algoritmalarının Karşılaştırılması”. Elektrik-Elektronik-Bilgisayar ve Biyomedikal Mühendisliği Sempozyumu, 27-29 Kasım, 2014, Bursa, Türkiye, 381-383.
- DURSUN, M, SAYGIN, A. (2006). “Bir Asansör Tahrik Sistemi İçin Bulanık Mantık Denetimli Anahtarlamalı Relüktans Motor Sürücüsü Tasarımı”. Pamukkale Üniversitesi Mühendislik Fakültesi Dergisi, 12(2): 151-160.
- GOPALAKRISHNAN, T, SENGOTTUVELAN, P. (2016). “A Hybrid PSO With Naive Bayes Classifier For Disengagement Detection In Online Learning”. Emerald Group Publishing Limited, 50(2): 215-224.
- HUSSAIN, M, ZHU, W, ZHANG, W, ABIDI, SMR. (2018). “Student Engagement Predictions In An E-Learning System And Their Impact On Student Course Assessment Scores”. Hindawi Computational Intelligence and Neuroscience, Volume 2018: 1-21.

- İMRAK, E, FETVACI, M.C.(2004). “Asansör Tesislerinde Koruyucu Bakım Esasları”. Mühendis ve Makine Dergisi, 45(536): 37-41.
- KARAHAN, I. (2016). Asansör Bakım İşlerinde Risklerin Değerlendirilmesi ve Çözüm Önerileri. İş Sağlığı ve Güvenliği Uzmanlık Tezi. Çalışma ve Sosyal Güvenlik Bakanlığı İş Sağlığı ve Güvenliği Genel Müdürlüğü, Ankara, Türkiye.
- KAVLAK, K. (2006). “Elektrikli Asansörler İle Hidrolik Asansörlerin Karşılaştırılması”. Selçuk Üniversitesi Teknik Bilimler Meslek Yüksekokulu Teknik-Online Dergi, 5(2): 28-40.
- KAYA, C, ERKAYMAZ, O, AYAR, O, ÖZER, M. (2017). “C4.5 Karar Ağacı Temelli Öznitelik Seçimi İle Video-Okülografi (VOG) Sinyallerinden Diyabetik Retinopati Hastalığının Sınıflandırılması”. Tıp Teknolojileri Kongresi, 12-14 Ekim, 2017, Trabzon, Türkiye, 75-78.
- KAYGIN, C.Y, TAZEGÜL, A, YAZARKAN, H. (2016). “İşletmelerin Finansal Başarılı Ve Başarısız Olma Durumlarının Veri Madenciliği Ve Lojistik Regresyon Analizi İle Tahmin Edilebilirliği”. Ege Akademik Bakış, 16(1): 147-159.
- KILINÇ, D, BOZYİĞİT, F, BORANDAĞ, E, YÜCALAR, F, AKYOL, H, AKIRMAK E.A. (2016). “Uzun, Z. Sınıflandırma Tabanlı Zombi Bilgisayar Tespit Sistemi”. Akademik Bilişim Konferansı 2016, Adnan Menderes Üniversitesi, 30 Ocak-5 Şubat 2016, Aydın, Türkiye.
- KIRAL, Z, KIRAL, BG. (2008). “Ray Deformasyonlarının Asansör Kabini Titreşimlerine Etkisinin İncelenmesi”. Mühendis ve Makine, 48(578): 2-7.
- KOÇ, Z. (2009). Asansör Kılavuz Ray Konsollarının Derilme Analizi. Yayınlanmamış Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul, Türkiye.
- KOYUNCUGİL, A.S, ÖZGÜLBAŞ, N. (2008). “İMKB’de İşlem Gören Kobi’lerin Güçlü Ve Zayıf Yönleri: CHAID Karar Ağacı Uygulaması”. Dokuz Eylül Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 23(1): 1-21.
- KRICHENE, A. (2017). “Using A Naive Bayesian Classifier Methodology For Loan Risk Assessment”. Journal of Economics, Finance and Administrative Science, 22(42): 3-24.
- LEMOYNE, R, MASTROIANNI, T. (2017). “Implementation Of A Smartphone Wireless Gyroscope Platform With Machine Learning For Classifying Disparity Of A Hemiplegic Patellar Tendon Reflex”. Journal of Mechanics in Medicine and Biology, 17(6):1750083-1- 1750083-15.
- ÖZTEMEL, E. (2003). Yapay Sinir Ağları. Papatya Yayıncılık, Birinci Baskı, İstanbul, Türkiye.
- PARK, S, YANG, B. (2010). “An Implementation of Risk-based Inspection For Elevator Maintenance”. Journal of Mechanical Science and Technology, 24 (12): 2367-2376.

- PEI, D, ZHANG, C, QUAN, Y, GUO Q. (2019). "Identification of Potential Type II Diabetes In a Chinese Population With a Sensitive Decision Tree Approach". Hindawi Journal of Diabetes Research, Volume 2019: 1-7.
- SALEHZADEH, R. (2017). "Which Types of Leadership Styles Do Followers Prefer? A Decision Tree Approach". International Journal of Educational Management, 31(7): 865-877.
- ŞEKER, Ş.E. (2013). İş Zekası ve Veri Madenciliği. Cinius Yayınları, Birinci Baskı, İstanbul.
- ŞEN, Z. (2004). Yapay Sinir Ağları İlkeleri. Su Vakfı Yayınları, İstanbul.
- TABACHNICK, B.G, FIDELL, L.S. (2015). Çok Değişkenli İstatistiklerin Kullanımı. Baloğlu, M. (eds), 6. Basımdan Çeviri, Nobel Yayın Evi, Ankara.
- TEKİN, R, KAYA, Y, TAĞLUK, M.E. (2011). "K-means ve YSA Temelli Hibrit Bir Model İle Epileptik EEG İşaretlerinin Sınıflandırılması". Elektrik ve Elektronik Bilgisayar Sempozyumu, 5-6-7 Ekim, 2011, Elazığ, Türkiye, 277-283s.
- TRIVEDI, S.K, PANIGRAHI, P.K. (2018). "Spam Classification: A Comparative Analysis Of Different Boosted Decision Tree Approaches". Journal of Systems and Information Technology, 20(3): 298-320.
- ÜNAL, M.Ö, AYKAÇ, B. (2010). "Yapı İşlerinde Asansör Kazaları Ve Güvenlik Önlemleri". International Journal of Engineering Research and Development, 2(2): 13-19.
- VERNEKAR, K, KUMAR, H, GANGADHARAN, K.V. (2018). "Engine Gearbox Fault Diagnosis Using Machine Learning Approach". Journal of Quality in Maintenance Engineering, 24(3): 345-357.
- WEKA2007.<http://weka.sourceforge.net/doc.dev/weka/classifiers/meta/RandomComittee.html>, 02.04.2019.
- WU, H-C, HU, Y-H, HUANG, Y-H. (2014). "Two-Stage Credit Rating Prediction Using Machine Learning Techniques". Kybernetes, 43(7): 1098-1113.
- ZARIKAS, V, LOUPIS, M, PAPANIKOLAOU, N, KYRITSI, C. (2013). "Statistical Survey Of Elevator Accidents In Greece". Elsevier Safety Science, 59: 93-103.