

**Atıf İçin:** Kuşkapan, E. (2024). Toplu Ulaşım Sistemlerinde Verimliliğin Artırılması Amacıyla Makine Öğrenmesi Yöntemlerinden Yararlanılması. *İğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 14(2), 743-754.

**To Cite:** Kuşkapan, E. (2024). Utilisation of Machine Learning Methods to Increase Efficiency in Public Transport Systems. *Journal of the Institute of Science and Technology*, 14(2), 743-754.

## **Toplu Ulaşım Sistemlerinde Verimliliğin Artırılması Amacıyla Makine Öğrenmesi Yöntemlerinden Yararlanılması**

Emre KUŞKAPAN

### **Öne Çıkanlar:**

- Kent içi ulaşım sistemleri
- Toplu taşımanın verimliliği
- Trafik güvenliği

### **Anahtar Kelimeler:**

- Toplu taşıma
- Sürüş güvenliği
- Makine öğrenmesi
- Sürdürülebilir ulaşım

### **ÖZET:**

Günlük ulaşım için toplu taşıma araçları sıklıkla tercih edilmektedir. Karayolu ulaşımında bu araçlar arasında en çok kullanılan türlerden biri otobüslerdir. Otobüsler, arabalardan çok daha fazla yolcu kapasitesine sahip olduğu için bu taşıtların sürücüleri çok daha fazla sorumluluk almaktadır. Yolculuk esnasında her yolcunun rahat, dakik ve güvenli ulaşımının sağlanması otobüs şoförlerinin sorumluluğundadır. Bu sorumlulukların fazlalığı otobüs şoförlerini psikolojik olarak etkileyebilmektedir. Bununla birlikte yaptıkları işin yorucu olması sebebiyle zaman zaman agresif tavırlar sergileyebilmektedirler. Yolculuk sırasında sürücü ile yolcu arasında çıkabilecek herhangi bir tartışma, çeşitli sorunlara neden olabilmektedir. Otobüslerinin içerisine yerleştirilen çok sayıda kamera ile sürücü ve yolcu davranışları gözlemlenebilmektedir. Bu uygulamanın yapıldığı kentlerden biri olan Erzurum'da Büyükşehir Belediyesi tarafından toplu taşıma otobüslerindeki denetimler raporlanmaktadır. Bu raporlar incelendiğinde kentte son 3 yıl içerisinde otobüs şoförleri ve yolcular arasında toplam 1068 adet tartışmanın olduğu tespit edilmiştir. Bu çalışmada, bu tartışmalar içeriklerine göre 13 kategoriye ayrılarak oluşturulan eğitim setindeki veriler makine öğrenmesi yöntemleriyle analiz edilmiştir. Analiz sonuçlarında naive bayes algoritmasının en başarılı performans ortaya koyduğu görülmüştür. Tüm verilerin bu algoritma ile dağılımı yapıldığında tartışmalarda en fazla pay sahibi olan sürücü ve yolcu kaynaklı problemler ortaya konmuştur. Sonuçlar yorumlandığında toplu taşıma ulaşımının daha verimli şekilde gerçekleştirilmesi ve sürücü-yolcu tartışmalarının azaltılabilmesi için birtakım tavsiyelerde bulunulmuştur.

## **Utilisation of Machine Learning Methods to Increase Efficiency in Public Transport Systems**

### **Highlights:**

- Urban transportation systems
- Efficiency of public transportation
- Traffic safety

### **Keywords:**

- Public transportation
- Driving safety
- Machine learning
- Sustainable transportation

### **ABSTRACT:**

Public transportation vehicles are frequently preferred for daily transportation. Buses are one of the most common types of these vehicles in road transport. As buses have a much higher passenger capacity than cars, the drivers of these vehicles take on much more responsibility. It is the responsibility of the bus driver to ensure that each passenger is transported comfortably, punctually and safely during the journey. This extra responsibility can have a psychological impact on bus drivers. Due to the tiring nature of their work, they can sometimes exhibit aggressive behaviour. Any argument between the driver and the passenger during the journey can cause various problems. The behaviour of drivers and passengers can be monitored with a large number of cameras installed in the buses. In Erzurum, one of the cities where this practice is used, the Metropolitan Municipality reports the inspections of public transport buses. When these reports were analysed, it was found that there were a total of 1068 arguments between bus drivers and passengers in the last 3 years. These arguments were divided into 13 categories according to their content, and the data in the training set created was analysed using machine learning methods. The results of the analysis showed that the naive Bayes algorithm was the most successful. When all the data was distributed with this algorithm, the problems caused by drivers and passengers, which have the highest proportion of discussions, were revealed. When interpreting the results, some recommendations are made in order to make public transport more efficient and to reduce driver-passenger arguments.

Emre KUŞKAPAN ([Orcid ID: 0000-0003-0711-5567](https://orcid.org/0000-0003-0711-5567)), Erzurum Teknik Üniversitesi, Mühendislik ve Mimarlık Fakültesi, İnşaat Mühendisliği Bölümü, Erzurum, Türkiye

**Sorumlu Yazar/Corresponding Author:** Emre KUŞKAPAN, e-mail: emre.kuskapan@erzurum.edu.tr

## GİRİŞ

Dünya genelinde her geçen gün artan taşıt sayısı ile birlikte trafik yoğunluğu da artış göstermektedir. Bu sebeple insanlar trafikte daha fazla zaman harcamak durumunda kalmaktadır. Trafikte harcanan zaman insanlarda sinir ve stres gibi olumsuz durumlara yol açabilmektedir. Bu duygular, özellikle sürücülerde daha çok ön plana çıkmaktadır. Bu durum sürücülerin trafik kuralı ihlali yapmasına, başka sürücü, yaya veya yolcular ile tartışmasına ve trafik kazası yapmalarına neden olabilmektedir. Bu nedenle trafikte güvenli ve sorunsuz yolculuk sağlanabilmesi için sürücü davranışlarının detaylı şekilde incelenmesi gereklidir.

Sürücü davranışlarını etkileyen faktörler çeşitlilik göstermektedir. Sürücü davranışını son yıllarda en belirgin olarak etkileyen faktör cep telefonu kullanımınıdır. Sürüş esnasında cep telefonu ile konuşulması sürücünün dikkatini önemli düzeyde etkilemektedir. Cep telefonu ile görüşme yapan sürücülerin intikal reaksiyon sürelerinin arttığı, trafik işaret ve levhalarının göz ardı edilebildiği ve daha agresif olabildikleri gözlemlenmiştir (Haque ve ark., 2021; Rosenberger, 2019). Sürücü davranışını etkileyen bir diğer önemli faktör ise sürüş deneyimidir. Genç sürücülerde genellikle sürüş deneyiminin daha az olmasına bağlı olarak kazaya karışma, trafikte güvenlik kurallarına uymama ve daha gergin olma gibi durumlar ile karşılaşmaktadır (Tronsmoen, 2010). Genç sürücülerde en çok görülen durumlardan birisi de trafikte heyecan yaşamak amacıyla birtakım riskleri almalarıdır. Daha deneyimli sürücüler güvenli ve risk içermeyen davranışlar sergilemektedir (Farooq ve Moslem, 2019; Hamarat ve Duran, 2020). Sürücü davranışını etkileyen bir diğer faktör ise diğer sürücülerin tutumlarıdır. Trafikte bir sürücünün gergin olması sebebiyle sergilediği davranışlar diğer sürücülerini de etkilemektedir. Uyuşturucu ve alkol tüketiminin de sürücü davranışını olumsuz yönde etkilediği oldukça bilinen bir durumdur. Özellikle araç kullanırken uyuşturucu veya alkol kullanan bireylerde trafik ihlallerinin arttığı, hataya daha çok meyilli oldukları, yolculukta gecikmeler yaşadıkları ve saldırgan yapıya sahip oldukları bilinmektedir (Farooq ve Moslem, 2020; Jomnonkwao ve ark., 2021). Benzer şekilde yorgunluk ve dikkat dağınıklığı da sürücü davranışını etkileyen önemli faktörlerdir. Sürücülerin yorgunken uzun mesafeli yolculuk yapması yolculuk esnasında uyuyakalması gibi çok ciddi riskler içermektedir. Kısa mesafeli yolculuklarda ise esneme ve gözün net görmemesi gibi durumlar riskler içermektedir. Dikkat dağınıklığı ise bireyin ruhsal yapısı veya taşıt kullanırken başka bir şey ile ilgilenmesinden kaynaklı olabilmektedir. Ruhsal olarak moral bozukluğu ve sınırlı olma durumu daha agresif sürüşe sebep olabilirken neşeli veya heyecanlı olma durumu ise yüksek hız yapma ve trafiği tehlikeye sokmaya sebep olabilmektedir. Dikkat dağınıklığına etki eden faktörler arasında ise sürüş esnasında birileri ile konuşmak, etrafı incelemek veya bir şeyler yiyip içmek yer almaktadır (Chhabra ve ark., 2017; Demircioğlu ev ark., 2019; Seth, 2020).

Sürücülerin sorumlulukları da kullandıkları taşıt türüne göre farklılık göstermektedir. Bireysel otomobil kullanan sürücülerin sorumlulukları daha az iken toplu taşıma araçlarını kullanan sürücülerin sorumlulukları çok daha fazladır. Bununla birlikte, toplu taşıma araçlarını kullanan sürücüler gün içerisinde önemli bir zamanlarını bu taşıtları kullanmakla geçirmektedirler. Bu sebeple yorgunluk ve iş yoğunluğundan kaynaklı daha gergin bir ruhsal yapıya sahip olabilmektedirler.

Toplu taşımada en sık kullanılan türlerden olan otobüslerde, sürekli yolcu inip binme durumu olduğu için taşıt içerisinde hareketlilik yaşanmaktadır. Bu durumla birlikte duraklarda da yolcu hareketliğine bağlı bir yoğunluk meydana gelmektedir. İnen veya binen yolcuların sürücüye yolculuk ile ilgili soru sormaları da sürücü için ekstra bir yoğunluk oluşturmaktadır. Tüm bu koşullar otobüs sürücülerinin davranışlarını etkilemektedir. Otobüs sürücülerinin yolculara veya trafikteki diğer sürücülere olan tutumları ise yolcu memnuniyetini önemli ölçüde etkilemektedir (Abenoza ve ark., 2017;

Cafiso ve ark., 2013). Sürücünün diğer bireylere davranışlarının yanı sıra yolculuk esnasında göstermiş olduğu davranışlar da yolcu memnuniyetini etkilemektedir. Özellikle güvenli bir yolculuk koşullarının oluşmaması önemli bir dezavantajdır. Sürücünün yüksek hızlara ulaşması ve trafik kurallarını ihlal etmesi gibi durumlar yolcu güvenliğini tehlikeye attığı için yolcu memnuniyeti düşük olmaktadır (Felleson ve Friman, 2008; Varmazyar ve ark., 2018; Sam ve ark., 2018).

### İlişkili Çalışmalar

Literatürde toplu taşıma araçlarında yolcu ve sürücü arasında meydana gelen tartışmaların sebep olduğu problemleri inceleyen çalışmalar yok denecek kadar azdır. Yapılan çalışmalar genel olarak toplu taşımada sürücü davranışının meydana getirdiği durumları ele almıştır. Örneğin Delice (2013) yaptığı çalışmada sürücülerini öfkeliendiren davranışları ve bu davranışlara verdikleri tepkileri incelemiştir. Toplu taşıma sürücülerini de kapsayan çalışmada anketlerden yararlanılmıştır. Anket verilerine istatistiksel testler uygulandığında sürücülerin diğer sürücü davranışlarına göre saldırganlıklarının arttıkları ortaya konmuştur. Mallia ve ark. (2015) otobüs şoförlerinin kişilik özelliklerinin sürüş ihlali, gecikme ve hatalara etkisini kişilik-tutum modeli kullanarak incelemiştir. İnceleme sonucunda trafikte fedakârlık, heyecan arayışı ve kural ihlali yapan sürücülerin trafik güvenliğini olumsuz etkilediği ortaya konmuştur. Useche (2017) ise otobüs şoförlerinin iş stresi ve yorgunluğunun sürüş davranışı üzerindeki etkilerini yapısal eşitlik modelleri kullanarak araştırmışlardır. Araştırmada, Kolombiya, Bogota'da 524 sürücü üzerinde anket yardımı ile inceleme yapmışlardır. Çalışma sonucunda çalışma koşullarına bağlı oluşan stres ve yorgunluğun sürüş güvenliğini riski hale getirdiği ortaya konmuştur. Murat ve Çakıcı (2017) ise bu alanda yaptıkları çalışmada toplu taşıma kullanımının artırılması ve özel araç kullanımının azaltılması amacıyla toplu taşımanın hizmet kalitesinin yeterliliğini araştırmışlardır. Toplu taşıma sistemlerindeki hizmet kalitesini araştırırken kamuoyu araştırması yapılmış ve konfor, frekans, bilgi sistemi ve ücret gibi parametrelerin öneminden bahsetmişlerdir.

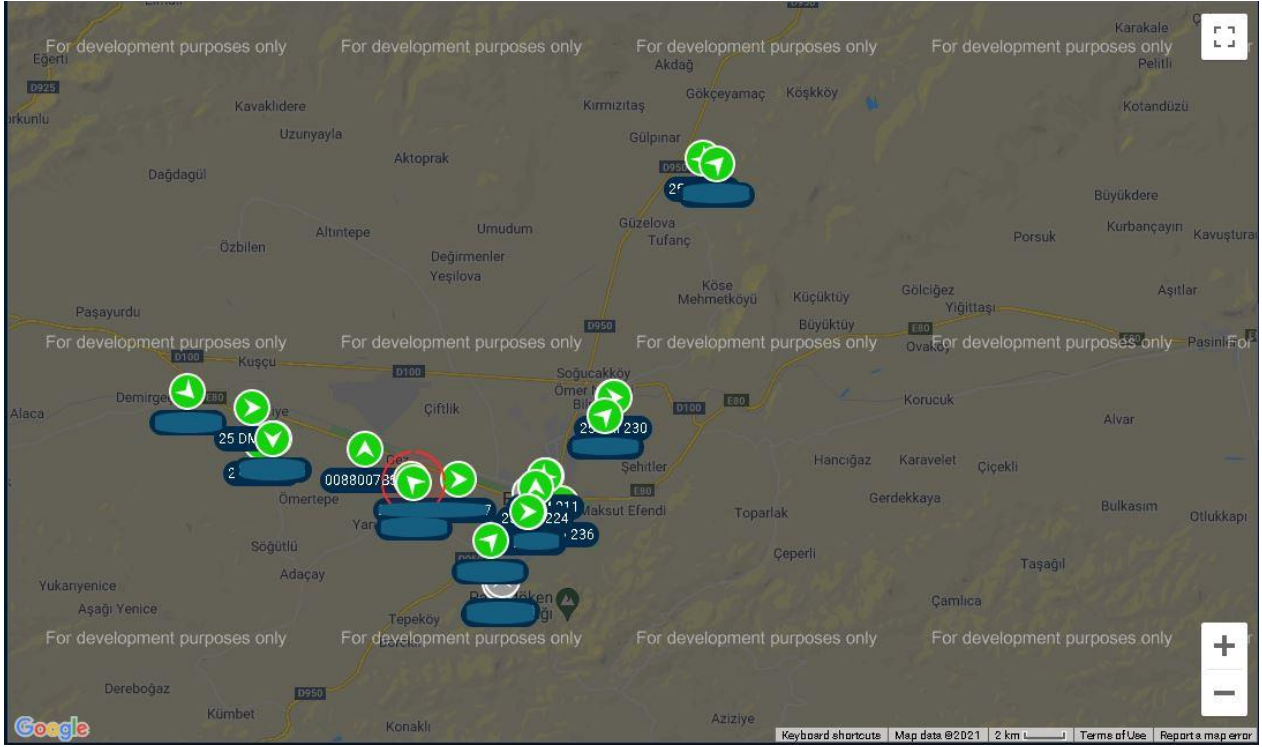
Toplu taşıma araçlarında meydana gelen problemleri incelemeye yönelik çalışmaların büyük bir kısmının anket verilerinden yararlanarak hazırlandığı görülmektedir. Literatürde toplu taşıma sürücüsü ile yolcular arasında meydana gelen tartışmaların kayıt altına alınarak daha gerçekçe sonuçlar ortaya koyan çalışmaların olmadığı tespit edilmiştir. Bu durum da sorunların daha iyi şekilde incelenerek kalıcı çözümler sunulmasına engel olmaktadır. Yapılan bu çalışmada Erzurum ilinde son 3 yılda otobüs şoförleri ile yolcular arasında meydana gelen tartışmaların kameralar ile kayıt altına alındığı verilerden yararlanılmıştır. Son 3 yılda toplu taşıma araçlarında meydana gelen 1068 adet tartışma sebepleri ana başlıklara ayrılarak makine öğrenmesi algoritmaları ile sınıflandırılmıştır. En iyi performans sergileyen algoritmaya göre oluşan sınıf yoğunlukları incelenerek sürücü ile yolcu arasındaki tartışmaya sebep olan faktörler ortaya konmuştur. Daha sonra bu faktörler incelenerek çözüm önerileri sunulmuştur.

## MATERYAL VE METOT

### Çalışma Alanı

Çalışma alanı olan Erzurum kenti Türkiye'nin doğusunda yer almaktadır. Birbirinden farklı güzergâhlara sahip toplam 37 farklı otobüs hattı kent merkezinde ulaşım olanağı sağlamaktadır. Bu 37 otobüs hattında aktif olarak 214 adet otobüs çalışmaktadır. Bunlardan 75 tanesi belediye otobüsü olup 139 tanesi ise özel halk otobüsüdür. Tüm otobüslerin çeşitli bölgelerinde yer alan kameralar vasıtasıyla anlık olarak görüntü ve ses kayıtları izlenebilmektedir. Bu kayıtlar belediye bünyesinde yer alan ulaşım kontrol merkezindeki personeller tarafından incelenip meydana gelen problemler detayları ile birlikte raporlanmaktadır. Şekil 1'de otobüslerde yer alan kamera kayıtları, görüntülerin hangi plakalı araca ait olduğu ve otobüslerin anlık konumları gösterilmektedir.

## Toplu Ulaşım Sistemlerinde Verimliliğin Artırılması Amacıyla Makine Öğrenmesi Yöntemlerinden Yararlanılması



a)



(b)

Şekil 1. a) Otobüslerin canlı olarak anlık konumları b) Otobüslerin çeşitli noktalarından kamera kayıtları (Erzurum Büyükşehir Belediyesi Ulaşım Daire Başkanlığı, 2021)

### Veri İşleme

Erzurum Büyükşehir Belediyesi Ulaşım Daire Başkanlığındaki uydu takip birimi raporlarında; olumsuzluğun meydana geldiği zamanı, hangi plakalı araçta meydana geldiği, sürücünün ismi ve yolcu ile meydana gelen tartışmanın detaylarını kayıt altına almaktadırlar. Tespit edilen problem sürücü kaynaklı ise ilgili kurum tarafından çeşitli yaptırımlar da uygulanabilmektedir. Yolcu ile otobüs sürücüsü arasındaki tartışmalar ikiye ayrılmaktadır. Bu tartışmaların büyük bir kısmı sürücü kaynaklı olup bir kısmı ise yolcu kaynaklıdır. Yolcu ile sürücü arasında en sık meydana gelen problem türleri Çizelge 1'de yer almaktadır.

**Çizelge 1.** Toplu taşıma araçlarında en sık rastlanılan sürücü ve yolcu kaynaklı problemler

Sürücü kaynaklı problemler	Yolcu kaynaklı problemler
1. Taşıtı hızlı sürme	1. Sürücüye sıkça soru sormak veya sürücüyü rahatsız etmek
2. Durakta durmama	2. Durak harici binme veya inme talebi
3. Yolcuyu araca almama	3. Ücret ödememe
4. Kazaya sebep olma	4. Diğer yolcuları rahatsız edici davranışlarda bulunmak
5. Taşıtın kapısını kapatmama	5. Otobüse zarar vermek
6. Argo ifadeler kullanma	
7. Trafik ışıklarına uymama	
8. Taşıtı kullanırken telefonla konuşma veya sigara içme	

Sürücü ve yolcu arasındaki tartışmaların büyük bir kısmı yukarıda belirtilen maddelere bağlı olarak meydana gelmektedir. Belirlenen bu 13 madde yolcu ile sürücü arasında meydana gelen toplam 1068 adet tartışmanın esas sebeplerini kapsamaktadır. Bu 13 sınıf oluştururken birbirine benzer özellikteki problemlerin aynı sınıf içerisinde yer alması amaçlanmıştır. Sürücü-yolcu tartışmalarına sebep olan maddeleri daha detaylı inceleyebilmek için tüm tartışmaların hangi sınıflarda yer aldıklarını ortaya koymak gereklidir. Çünkü toplu taşıma sistemleri kentlerin en önemli ulaşım yöntemlerinden birisidir (Özden ve ark. 2020). Toplu taşıma sistemlerindeki sürücü-yolcu etkileşimini ortaya koyabilmek için ise makine öğrenmesi algoritmalarından yararlanılmıştır. Makine öğrenmesi algoritmaları sayesinde veri kümesinin bir kısmı eğitilerek geriye kalan verilerin sınıflarının başarılı bir şekilde sınıflandırılması mümkündür. Bu sayede sınıflandırma işlemi daha pratik bir şekilde yapılmış olmaktadır.

### Makine Öğrenmesi

Bir veri kümesinin daha iyi yorumlanabilmesi için çeşitli sınıflandırmalar yapılmaktadır. Yapılan sınıflandırmalar verilerin kolayca incelenmesi ve hızlı çözümler sunulabilmesi açısından oldukça faydalıdır. Sınıflandırma işlemlerinde ise yapay zekâ tabanlı yöntemler oldukça yüksek performanslar ortaya koymaktadır. Bunlardan bir tanesi olan makine öğrenmesi insanların öğrenme şekillerini taklit etmektedir. İstatistiksel yöntemler kullanarak sınıflandırma veya tahminler yapmaya olanak sağlamaktadır. Bu işlemin gerçekleşmesini sağlayan çeşitli algoritmaları bünyesinde barındırmaktadır. Her bir algoritma veri kümesini farklı türde analiz etmektedir. Veri kümesine en uygun algoritmayı tespit ederek daha doğru sonuçlar elde edebilmek mümkündür. Bu sebeple yapılan çalışmadaki veri kümesi farklı türden algoritmalar ile sınıflandırılmıştır.

Çalışmada naive bayes, regresyon, karar ağacı ve en yakın komşuluk algoritmaları kullanılmıştır. Naive bayes algoritması karmaşık olmayan verilerde hızlı ve başarılı sonuçlar verirken veriler karmaşıklaşınca hata oranı artabilmektedir. Regresyon algoritması ise bağımlı ve bağımsız değişkenler arasında ciddi ilişki varsa başarılı bir şekilde ortaya koyabilirken değişkenler arası ilişki karmaşık ise başarısı düşmektedir. Karar ağacı algoritması ise değişkenlerin etkileşimini başarılı bir şekilde ortaya koyabilirken hassas tahmin yapabilme yeteneği düşüktür. En yakın komşuluk algoritması ise çok sınıflı problemlerde başarılı olmakla birlikte veri seti fazla olduğunda başarısı düşebilmektedir.

### Naive bayes

Naive bayes algoritmasında eğitim setinde öğretilmiş veriler sunulduktan sonra istatistiksel işlemler ile yeni test verilerinin hangi kategorilerde yer aldığı belirlenmeye çalışılır. Öğretilmiş veri sayısının fazlalığı algoritmanın daha başarılı sonuçlar vermesini sağlamaktadır. Naive Bayes algoritmasının çalışmasında Bayes teoremi kullanılır. Bayes teoremi 1 numaralı eşitlikte gösterilmektedir (Yıldırım ve Yüksel, 2017).

$$P(A/B) = (P(B/A) P(A)) / P(B) \quad (1)$$

$P(A/B)$ : B olayı bilindiğinde A olayının gerçekleşme olasılığı

$P(B/A)$ : A olayı bilindiğinde B olayının gerçekleşme olasılığı

$P(A)$ : A olayının gerçekleşme olasılığı

$P(B)$ : B olayının gerçekleşme olasılığı

### Karar ağacı

Karar ağaçları çeşitli veri madenciliği konularında sıklıkla kullanılmaktadır. Önceden belirlenmiş bir hedef çıktıya sahiptirler. Bu algoritma, birçok özeliğe sahip olan bir veri kümesini farklı özelliklere sahip kriterlere göre kümelemeyebilmektedir. Karar ağacında en temel iki kavram olan entropi ve bilgi kazancı 2 ve 3 numaralı eşitlikler ile hesaplanmaktadır (Ülgen, 2017).

$$H = -\sum p(x) \log p(x) \quad (2)$$

$$Gain(S, D) = H(S) - \sum_{V \in D} \frac{|V|}{|S|} H(V) \quad (3)$$

$p(x)$ : belirli bir sınıfa ait grubun yüzdesi

H: entropi

S: orijinal veri kümesi

D: kümenin bölünmüş parçası

V: S'nin bir alt kümesi

Gain: bilgi kazancı

### K en yakın komşuluk

K en yakın komşuluk algoritması regresyon veya sınıflandırma işlemlerinde tercih edilebilmektedir. Algoritma sınıfları belli olan bir örnek kümesindeki verilerden yararlanarak yeni verilerin Öklid uzunluk hesaplama yöntemi ile hangi sınıf aralığında yer aldığını belirlemektedir. En yakın komşuluk algoritmasının en önemli avantajı modelini daha görülmemiş veri uyarlayabilmesidir. Yani bazı algoritmalarda eğitim sürecinde tüm verilerin gösterilmesi gerekirken bu algoritmada öyle bir zorunluluk yoktur. Bu yöntemde kullanılan hesaplama 4 numaralı eşitlikte verilmiştir (Özdemir, 2019).

$$d(i, j) = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (4)$$

Burada, i ve j noktaları göstermekle birlikte  $d(i, j)$  öklit uzaklığını ifade etmektedir.

### Regresyon

Regresyon analizi, bir ya da daha fazla bağımsız değişken ile hedef değişken arasındaki ilişkiyi matematiksel olarak modelleyen bir yöntemdir. Regresyon, iki (ya da daha çok) değişken arasındaki doğrusal ilişkinin fonksiyonel şeklini, biri bağımlı diğeri bağımsız değişken olarak bir doğru denklemi olarak göstermekle kalmaz, değişkenlerden birinin değeri bilindiğinde diğeri hakkında kestirim yapılmasını sağlar. Regresyon modeli hesaplamasında 5 numaralı denklem kullanılır (Hui ve Jha 2000; Kuşkapan ve Çodur 2022).

$$y_i = \beta_0 + \sum_{j=1}^p X_{ij} \beta_j + \varepsilon_i \quad (5)$$

Burada;  $y_i$  bağımlı değişken parametrelerin bir doğrusal birleşimi olmak koşulu ile  $x_i$  bağımsız değişkenler,  $\beta_0$  ve  $\beta_1$  parametreler ve  $\varepsilon_i$  hata terimini ifade etmektedir (Mihăescu, 2011).

### Algoritmaların performans analizi

Algoritmaların performans analizleri yapılırken kesinlik, duyarlılık, f-ölçütü ve hata ölçekleri hesaplanmaktadır. Hata ölçeklerinde ise genel olarak ortalama mutlak hata (MAE), kök hata kareler ortalaması (RMSE), Kappa istatistiği ve doğruluk değerleri esas alınmaktadır. Bu terimlerden kesinlik, duyarlılık, f-ölçütü ve doğruluk değerlerinin hesaplanmasında karışıklık matrisi kullanılmaktadır. Örnek bir karışıklık matrisi Şekil 2’de verilmiştir.

KARIŞIKLIK MATRİSİ		TAHMİN EDİLEN DEĞER	
		Pozitif	Negatif
GERÇEK DEĞER	Pozitif	Doğru Pozitif (TP)	Yanlış Pozitif (FP)
	Negatif	Yanlış Negatif (FN)	Doğru Negatif (TN)

Şekil 2. Karışıklık matrisi (Dinçer ve Özkan, 2023)

Karışıklık matrisinde TP (doğru pozitif-doğruya doğru demek), TN (doğru negatif-doğruya yanlış demek), FP (yanlış pozitif-yanlış doğru demek) ve FN (yanlış negatif-yanlış yanlış demek) değerleri kullanılmaktadır. Bu değerlerin kullanılması ile kesinlik, duyarlılık, f-ölçütü ve doğruluk kriterleri aşağıda verilen 6-7-8 denklemlerinde gösterildiği gibi hesaplanır. ROC değeri ise model performansının genel olarak yorumlanabilmesi için oluşturulan eğri ile elde edilmektedir. Bu performans değerlerinin hepsi 0 ile 1 arasında değerler almaktadır.

$$\text{Kesinlik} = \frac{TP}{TP+FP} \quad (6)$$

$$\text{Duyarlılık} = \frac{TP}{TP+FN} \quad (7)$$

$$F - \text{ölçütü} = \frac{2 \times \text{Duyarlılık} \times \text{Kesinlik}}{\text{Duyarlılık} + \text{Kesinlik}} \quad (8)$$

Veri madenciliğinde performans ölçütleri olan kesinlik, duyarlılık ve f-ölçütü değerlerinin yüksek olduğu algoritma daha doğru sınıflandırma yapmaktadır. Bununla birlikte hata ölçütleri olan Kappa istatistiği ve doğruluk değerlerinin yüksek MAE ve RMSE değerlerinin düşük olduğu algoritmaların da daha doğru sınıflandırma yaptığı bilinmektedir. MAE tüm veri öngörülen değerler ile ölçüm değeri arasındaki farkın ortalaması olarak ifade edilmektedir. RMSE değeri model tarafından tahmin edilen değerler ve elde edilen ölçüm değeri arasındaki farkın ortalamasının karekökü alınarak hesaplanır. Kappa değeri ise gözlemsel arasındaki uyumayı ölçmek için ifade edilen bir terimdir. Bu üç terimin hesaplanması şekli aşağıda 9-10-11 numaralı denklemlerde verilmiştir.

$$MAE = \frac{1}{N} \sum_{i=1}^N |x_{f,i} - x_{o,i}| \quad (9)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{f,i} - x_{o,i})^2} \quad (10)$$

$$K = \frac{p_o - p_e}{1 - p_e} \quad (11)$$

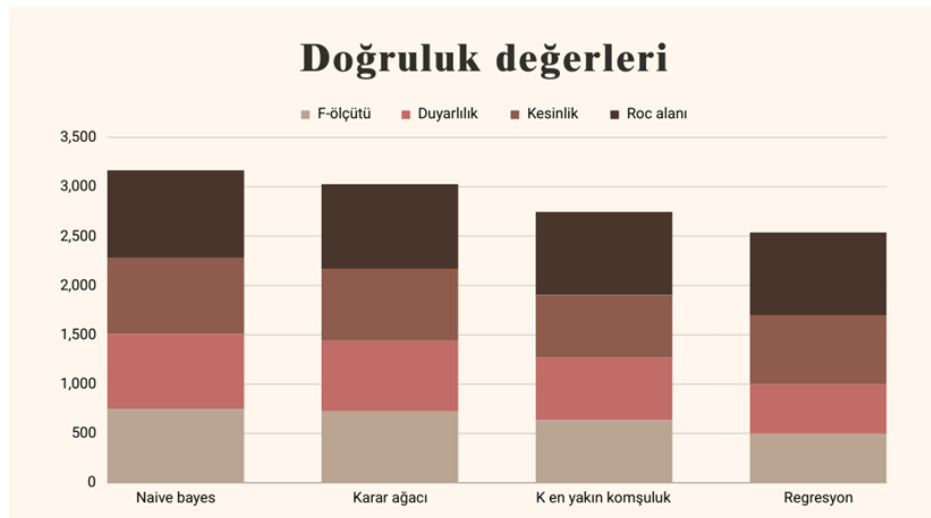
Burada  $x_{f,i}$  ve  $x_{o,i}$  sırasıyla beklenen ve gözlemlenen değerleri içermektedir.  $p_o$  gözlenen uyumluluk oranı,  $p_e$  tesadüfi uyumluluk oranı olmak üzere kappa istatistiği ( $K$ ) formülüyle hesaplanmaktadır (Shatu ve ark., 2019). Kappa istatistik değerinin yorumlanması için Çizelge 2 kullanılmaktadır.

**Çizelge 2.** Kappa değerinin yorumlanması

Kappa değeri	Uyuşma derecesi
<0	Hiç uyuşma olmaması
0.01-0.20	Önemsiz uyuşma olması
0.21-0.40	Orta derecede uyuşma olması
0.41-0.60	Ekseriyetle uyuşma olması
0.61-0.80	Önemli derecede uyuşma olması
0.81-1.00	Neredeyse mükemmel uyuşma olması

## BULGULAR VE TARTIŞMA

Toplu taşıma araçlarında meydana gelen tartışmalar sürücü ve yolcu kaynaklı olarak toplam 13 ana maddeye ayrılmıştır. Bu maddelerin her biri bir sınıfı temsil etmektedir. Toplu taşıma araçlarında meydana gelen toplam 1068 adet tartışma belirlenen bu 13 sınıftan oluşmaktadır. Belirlenen sınıflar için makine öğrenmesi algoritmaları olan Naive bayes, karar ağacı, k en yakın komşuluk ve regresyon yöntemleri kullanılmıştır. Her bir algoritma veri setini farklı yöntemlerle sınıflandırdığı için performans değerleri ve hata ölçekleri farklıdır. Mevcut veri setinin %70'i eğitim, %15'i doğrulama ve %15'i ise test için ayrılmıştır. Yolcu ile sürücü arasındaki tartışmalar sınıflandırıldığında en yüksek doğruluk değerlerine sahip algoritmanın daha başarılı olduğu söylenebilmektedir. Şekil 3'te her bir algoritmaya ait doğruluk değeri verilmiştir.

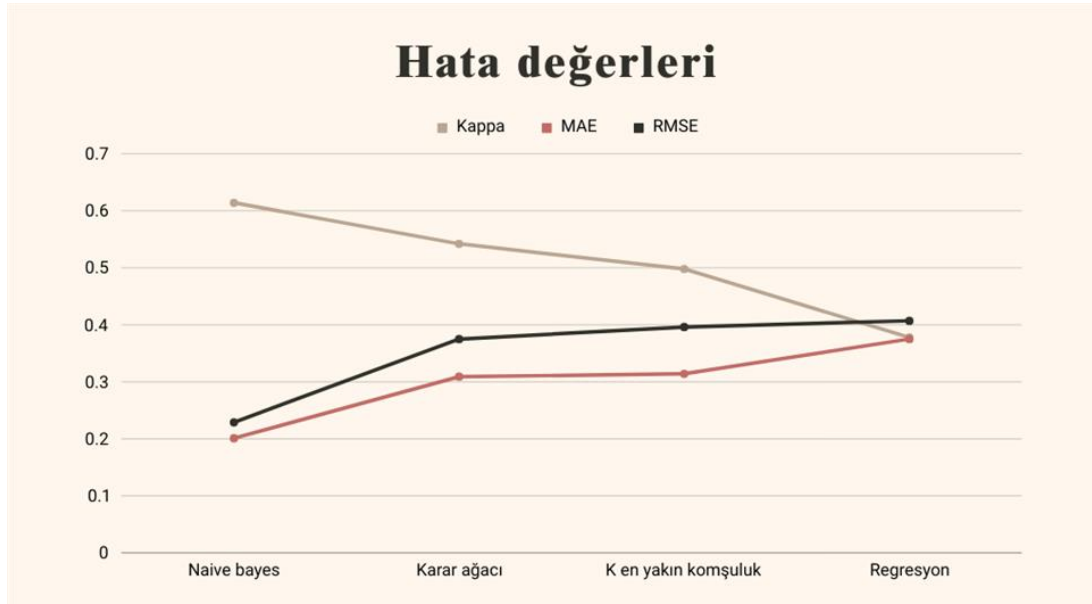


**Şekil 3.** Her bir algoritma ile elde edilen doğruluk değerleri

Doğruluk değerleri içerisinde her bir kavram farklı bir doğruluk ifadesi olsa da makine öğrenmesi için bu kavramlardan en belirleyici olan f- ölçütüdür. Diğer kavramlar tek başına bir algoritmanın başarısını ifade etmiyor olsa da her bir doğruluk değeri de birbirleri ile ilişkidir. Şekil 4'te verilen



sonuçlar incelendiğinde naive bayes algoritmasının hem f-ölçütü açısından hem de tüm doğruluk değerleri açısından en başarılı sonuçlar verdiği görülmektedir. Başarılı bir algoritmada; doğruluk değerlerinin yüksek olmasının yanı sıra makine öğrenmesindeki hata değerleri olan kappa istatistik değerinin yüksek, MAE ve RMSE değerlerinin de düşük olması beklenmektedir. Eğer bir algoritma ile elde edilen doğruluk değerleri ile hata değerleri arasında uyum yoksa bu algoritmanın mevcut veri kümesi için uygun olmadığı söylenebilir. Şekil 4'te her bir algoritma ile elde edilen hata değerleri yer almaktadır.



Şekil 4. Her bir algoritma ile elde edilen hata değerleri

Algoritmalarından elde edilen hata değerleri de incelendiğinde en düşük MAE ve RMSE, en yüksek kappa istatistik değerine sahip olan algoritmanın naive bayes algoritması olduğu görülmektedir. Öte yandan karar ağacı ve k en yakın komşuluk algoritmalarının MAE ve RMSE değerlerinin birbirine yakın olduğu görülmektedir. Regresyon algoritmasının ise doğruluk değerlerinde olduğu gibi hata değerlerinde de diğer algoritmalara göre daha az başarılı sonuçlar verdiği görülmektedir. Hem doğruluk değerleri hem de hata değerleri birlikte incelendiğinde mevcut veri kümesi için naive bayes algoritmasının en başarılı performans ortaya koyduğu görülmektedir. Meydana gelen toplam 1068 problemin naive bayes algoritması ile yapılan sınıflandırmada yüzdelik olarak dağılımları Çizelge 3'te yer almaktadır.

Çizelge 3. Sürücü ve yolcu kaynaklı problemlerin meydana gelme sıklığı

Problemin kaynağı	Tartışmaya sebep olacak içerik	Meydana gelme sayısı	Meydana gelme sıklığı (%)	
Sürücü	Taşıtı hızlı sürme	139	13	
	Durakta durmama	75	7	
	Yolcuyu araca almama	64	6	
	Kazaya sebep olma	64	6	
	Taşıtin kapısını kapatmama	75	7	
	Argo ifadeler kullanma	96	9	
	Trafik ışıklarına uymama	117	11	
	Taşıtı kullanırken telefonla konuşma veya sigara içme	128	12	
	Yolcu	Sürücüye sıkça soru sormak veya sürücüyü rahatsız etmek	117	11
		Durak harici binme veya inme talebi	117	11
Ücret ödememe		32	3	
Diğer yolcuları rahatsız edici davranışlarda bulunmak		21	2	
Otobüse zarar vermek		21	2	

Tartışmaya sebep olabilecek problemler incelendiğinde; sürücü kaynaklı problemler içerisinde trafik kurallarına uymama durumunun daha sıklıkla karşılaşıldığı görülmektedir. Bunlar içerisinde taşıtı hızlı sürme ve trafik ışıklarına uymama trafik kazalarına sebep olan temel unsurlar arasında yer almaktadır. Bunların haricinde ön plana çıkan bir diğer problem ise taşıt kullanırken telefonla konuşma veya sigara içme durumunun trafik kazasına sebep olma ihtimali daha düşükken yolcuları rahatsız edebilme durumu yüksektir. Yolcu kaynaklı problemlerde ise sürücüye sıkça soru sormak veya sürücüyü rahatsız etmek ile durak harici binme veya inme talebi öne çıkan problemler arasında yer almaktadır. Sürücüler günün önemli bir kısmını direksiyon başında geçirdikleri için normal bir insana göre daha yorgun veya sinirli olabilmesi mümkündür. Ayrıca sürüş esnasında çok fazla soru sorulması da dikkatini dağıttığı için bu durumlar tartışmalara sebep olabilmektedir. Öte yandan durak harici binme veya inme talebi de sürücüyü zor durumda bırakabileceği için bu duruma karşı gelmesi söz konusu olduğunda yolcu ile tartışmaların yaşanabilmesi mümkündür.

Toplu taşıma için yukarıda bahsedilen bu problemler birlikte değerlendirildiğinde çözüm önerisi olarak sürücülerin trafik kurallarına daha fazla riayet etmeleri hususunda yetkili kuruluşlar tarafından uyarıların yapılması, sürücüler için belirli zaman aralıkları ile bilgilendirici seminerlerin yapılması ve denetimlerin sağlanarak kurallara uyan sürücüler için ödüllendirilmeler yapılması sayesinde daha verimli sonuçlar elde edilebilecektir. Durum yaya özelinde değerlendirildiğinde ise mevcut durak noktalarında ve toplu taşıma mobil uygulamasında yeterli bilgilendirmenin sağlanması ve durak harici binme veya inme taleplerinde bulunmamaları hususunda çeşitli uyarıların yer alması sağlanırsa sürücü-yolcu tartışmalarının önemli bir kısmı engellenmiş olacaktır.

## SONUÇ

Yapılan bu çalışmada toplu taşıma sürücülerini ile yolcular arasında meydana gelen tartışmalara sebep olabilecek problemler yolcu ve sürücü kaynaklı olmak üzere iki ana kısma ayrılmıştır. Erzurum ilinde son üç yılda meydana gelen toplam 1068 adet sürücü-yolcu tartışmaları için 13 farklı sınıf belirlenmiştir. Meydana gelen tüm tartışmaların hangi sınıflara ait olduğunu belirleyebilmek için ise makine öğrenmesi algoritmalarından yararlanılmıştır. Mevcut veri seti için eğitim setinde sınıflar öğretilerek test veri setinde sınıflandırmaların tahminleri elde edilmiştir. Eğitim veri seti içinde naive bayes, k en yakın komşuluk, karar ağacı ve regresyon algoritmaları içerisinde %81 f-ölçütü değeri ile en başarılı algoritma olarak naive bayes algoritması ön plana çıkmıştır. Bu algoritma ile geriye kalan verilerin de sınıflandırılması yapıldığında sürücü kaynaklı problemler içerisinde trafik kurallarına uymama durumunun yani taşıtı hızlı sürme ve trafik ışıklarına uymama durumlarının en yaygın problemler olduğu belirlenmiştir. Yolcu kaynaklı problemlerde ise sürücüye sıkça soru sormak veya sürücüyü rahatsız etmek ile durak harici binme veya inme talebi öne çıkan problemler arasında yer aldığı tespit edilmiştir. Bu öne çıkan problemleri azaltabilmek için ise öncelikle sürücüler için trafik kurallarına uymaları yönünde bilgilendirmelerin yapılmasının yanında kural ihlali yapmayan sürücüler için ödüllendirmeler sağlanabilmesi önerilmektedir. Yolculardan kaynaklı problemlerin önüne geçebilmek için ise durak yerleri ve mobil uygulamada yolcu bilgilendirme sisteminin daha başarılı şekilde sağlanmasının yanında durak harici inme ve binme taleplerinin bulunulmaması hususunda yolcuların uyarılması önerilmiştir. Yapılan bu çalışmada elde edilen sonuçlar ve yorumlamalar dikkate alındığı taktirde kentlerde toplu ulaşımın daha verimli şekilde sağlanabilmesi mümkündür.

## Çıkar Çatışması

Yazar herhangi bir çıkar çatışmasının olmadığını beyan eder.

**KAYNAKLAR**

- Abenoza, R. F., Cats, O., & Susilo, Y. O. (2017). Travel satisfaction with public transport: Determinants, user classes, regional disparities and their evolution. *Transportation Research Part A: Policy and Practice*, 95, 64-84.
- Cafiso, S., Di Graziano, A. ve Pappalardo, G. (2013). Using the Delphi method to evaluate opinions of public transport managers on bus safety. *Safety science*, 57, 254-263.
- Chhabra, R., Verma, S. ve Krishna, C. R. (2017). A survey on driver behavior detection techniques for intelligent transportation systems. In *2017 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence* (s. 36-41). Noida, India. Erişim adresi: <https://ieeexplore.ieee.org/abstract/document/7943120>
- Delice, M. (2013). Sürücüyü Öfkeliendiren Davranışların ve Sürücülerin Bu Davranışlara Verdiği Tepkilerin İncelenmesi/The Investigation of Factors Causing Anger on Drivers and the Drivers' Response Against These Factors. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 17(2).
- Demircioğlu, A., Baş, S. S. ve Uysal, S. A. (2019). Farklı tipte araç kullanan sürücülerde ağrı, yorgunluk, uyku ve yaşam kalitesinin karşılaştırılması. *Journal of Exercise Therapy and Rehabilitation*, 6(2), 104-111.
- Dinçer, Y., Özkan, İ. (2023). Çevresel seslerin evrimsel sinir ağları ile sınıflandırılması. *Konya Journal of Engineering Sciences*, 11(2), 468-490.
- Erzurum Büyükşehir Belediyesi Ulaşım Daire Başkanlığı. (2021). Toplu taşıma kontrol merkezi. Erişim adresi: <https://ulasim.erzurum.bel.tr/>.
- Farooq, D. ve Moslem, S. (2019). A fuzzy dynamical approach for examining driver behavior criteria related to road safety. In *2019 smart city symposium Prague (SCSP)* (s. 1-7). Prague, Czech Republic. Erişim adresi: <https://ieeexplore.ieee.org/abstract/document/8805741>
- Farooq, D. ve Moslem, S. (2020). Evaluation and ranking of driver behavior factors related to road safety by applying analytic network process. *Periodica Polytechnica Transportation Engineering*, 48(2), 189-195.
- Fellessen, M. ve Friman, M. (2008). Perceived satisfaction with public transport service in nine European cities. In *Journal of the Transportation Research Forum* (s. 92-102). Erişim adresi: <http://www.trforum.org/journal>
- Hamarat, B. ve Duran, E. (2020). The factors affecting the drivers offensive behavior in the traffic: case of Çanakkale. *Journal of Awareness*, 5(4), 501.
- Haque, M. M., Oviedo-Trespalacios, O., Sharma, A., ve Zheng, Z. (2021). Examining the driver-pedestrian interaction at pedestrian crossings in the connected environment: A Hazard-based duration modelling approach. *Transportation research part A: policy and practice*, 150, 33-48.
- Hui, S. ve Jha, G. (2000). Application data mining for customer service support. *Information and Management*, 38, 1-13.
- Jomnonkwao, S., Utra, S. ve Ratanavaraha, V. (2021). Analysis of a driving behavior measurement model using a modified driver behavior questionnaire encompassing texting, social media use, and drug and alcohol consumption. *Transportation research interdisciplinary perspectives*, 9, 100302.
- Kuşkapan, E. ve Çodur, M. Y. (2022). Trafik kazalarının sınıflandırılmasında çok katmanlı algılayıcı, regresyon ve en yakın komşuluk algoritmalarının performans analizi. *Politeknik Dergisi*, 25(1), 373-380.

- Mallia, L., Lazuras, L., Violani, C. ve Lucidi, F. (2015). Crash risk and aberrant driving behaviors among bus drivers: the role of personality and attitudes towards traffic safety. *Accident Analysis & Prevention*, 79, 145-151.
- Mihăescu, M. C. (2011). Classification of learners using linear regression. In *2011 federated conference on computer science and information systems (FedCSIS)* (s. 717-721). Szczecin, Poland. Erişim adresi: <https://ieeexplore.ieee.org/abstract/document/6078214>
- Murat, Y. S. ve Cakici, Z. (2017). Comparative Analysis of Public Transport Users' Perception Targeting Sustainable Transportation. In *Engineering Tools and Solutions for Sustainable Transportation Planning*, 76-98, IGI Global.
- Özdemir, K. (2019). K-En Yakın Komşu Algoritması. Erişim adresi: <https://medium.com/batech/k-en-yakin-komşu-algoritması>
- Özden, A., Akalın, B. ve Çağdaş, K. (2020). ITS Applications in Public Transportation: What Municipalities Offer to Travelers in Turkey. *Journal of Transportation and Logistics*, 4(2), 51-64.
- Rosenberger, R. (2019). The experiential niche: Or, on the difference between smartphone and passenger driver distraction. *Philosophy & Technology*, 32(2), 303-320.
- Sam, E. F., Brijs, K., Daniels, S., Brijs, T. ve Wets, G. (2018). Public bus passenger safety evaluations in Ghana: A phenomenological constructivist exploration. *Transportation research part F: traffic psychology and behaviour*, 58, 339-350.
- Seth, I. (2020). A survey on driver behavior detection techniques. *International Journal of Scientific Research in Science and Technology*, 7, 401-404.
- Shatu, F., Yigitcanlar, T. ve Bunker, J. (2019). Objective vs. subjective measures of street environments in pedestrian route choice behaviour: Discrepancy and correlates of non-concordance. *Transportation research part A: policy and practice*, 126, 1-23.
- Tronsmoen, T. (2010). Associations between driver training, determinants of risky driving behaviour and crash involvement. *Safety Science*, 48(1), 35-45.
- Useche, S. A., Ortiz, V. G. ve Cendales, B. E. (2017). Stress-related psychosocial factors at work, fatigue, and risky driving behavior in bus rapid transport (BRT) drivers. *Accident Analysis & Prevention*, 104, 106-114.
- Ülgen, E. K. (2017). Makine Öğrenimi Bölüm-5 (Karar Ağaçları). Erişim adresi: <https://medium.com/@k.ulgen90/>
- Varmazyar, S., Mortazavi, S. B., Arghami, S. H. ve Hajizadeh, E. (2014). Factor analysis of driver behavior questionnaire (DBQ) in public transportation Bus Company: Investigation of the relationship between DBQ factors and crashes. *Scientific Journal of Review*, 3(4), 155-165.
- Yıldırım, M., ve Yüksel, C. A. (2017). Sosyal medya ile hisse senedi fiyatının günlük hareket yönü arasındaki ilişkinin incelenmesi: Duygu analizi uygulaması. *Uluslararası İktisadi ve İdari İncelemeler Dergisi*, 33-44.