

## Amerika’da Meydana Gelen Trafik Kazalarının Yapay Sinir Ağları ve Çok Değişkenli Regresyon Yöntemleriyle Tahmini

Ömer Faruk CANSIZ<sup>1</sup>, İbrahim ERGİNER<sup>2\*</sup>, Ebru DOĞRU<sup>3</sup>

<sup>1</sup>İskenderun Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, İnşaat Mühendisliği Bölümü, Hatay

<sup>2</sup>İskenderun Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, İnşaat Mühendisliği Bölümü, Hatay

<sup>3</sup>İskenderun Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, İnşaat Mühendisliği Bölümü, Hatay

<sup>1</sup><https://orcid.org/0000-0001-6857-2513>

<sup>2</sup><https://orcid.org/0000-0002-6001-3124>

<sup>3</sup><https://orcid.org/0000-0002-7353-1115>

\*Sorumlu yazar: ibrahim.erginer.mfbel8@iste.edu.tr

### Araştırma Makalesi

#### Makale Tarihiçesi:

Geliş tarihi: 24.12. 2020

Kabul tarihi: 17.04.2021

Online Yayınlanma: 15.12.2021

#### Anahtar Kelimeler:

Kaza sayısı

Yapay sinir ağları

Çok değişkenli regresyon

### ÖZET

Dünya sağlık örgütü (WHO) dünyada 1,35 milyon kişinin trafik kazaları nedeniyle hayatını kaybettiğini açıklamaktadır. Amerika’da da dünyada olduğu gibi Amerika’da da trafik kazası nedeniyle hayatını kaybeden kişi sayısı çok fazladır. Çok sayıda insanın hayatını kaybetmesine sebep olan trafik kazalarının birçok nedeni olabilir. Yapılan bu çalışmada trafik kazalarına sebep olan değişkenler incelenmektedir. Araç sayısı, sürücü sayısı, nüfus ve seyahat edilen araç mili değişkenleri kullanılarak bir veri seti oluşturulmaktadır. Veri seti oluşturulurken Amerika Birleşik Devletleri Ulaştırma Bakanlığı veri tabanından elde edilen veriler kullanılmaktadır. Veri seti 1994-2018 yılları arasındaki verileri kapsamaktadır. Oluşturulan veri seti kullanılarak istatistiksel yöntemlerden çok değişkenli regresyon (MR) ve yapay zekâ tekniklerinden yapay sinir ağları (YSA) kullanılarak kaza sayısı tahmin modelleri geliştirilmektedir. Geliştirilen modeller birbiriyle kıyaslanmaktadır. Modeller kıyaslanırken hataların kareleri ortalaması (HKO), ortalama yüzde hata (OYH) ve korelasyon katsayısı (R) kriterleri kullanılmaktadır. Yapılan analizler sonucunda YSA kaza sayısı tahmin modelinin HKO, OYH ve R değerleri sırasıyla 265228,79-%0,829-0,988 olarak hesaplanmaktadır. Yapılan bu çalışmada YSA modelinin regresyon modellerine göre daha iyi sonuçlar verdiği gözlemlenmektedir. Buna benzer çalışmalarda yapay zekâ metodlarından YSA’nın kullanılması önerilmektedir.

## Prediction of Traffic Accidents in The America by Using Artificial Neural Networks and Multivariate Regression Methods

### Research Article

#### Article History:

Received: 24.12.2020

Accepted: 17.04.2021

Published online: 15.12.2021

#### Keywords:

Number of accidents

Artificial neural networks

Multivariate regression

### ABSTRACT

World Health Organization (WHO) announced that 1.35 million people die worldwide due to traffic accidents. In America, the number of people who lost their lives due to traffic accidents is very high, as in the world. There may be many reasons for the occurrence of traffic accidents. In this study, the variables that cause traffic accidents are examined. A data set is created using the variables of number of vehicles, the number of drivers, population, the number of vehicles traveled and the number of vehicle miles traveled. Data obtained from the database of the United States Department of Transportation are used while creating the data set. The data set includes data between the years 1994-2018. Accident number prediction models are developed by using multivariate

regression (MR) from statistical methods and artificial neural networks (ANN) from artificial intelligence techniques by using the data set. The models developed are compared with each other. While comparing the models, the criteria such as mean squares of errors (MSE), mean of percentage errors (MPE) and correlation coefficient (R) are used. As a result of the analyzes made, the MSE, MPE and R values of the ANN accident number prediction model were calculated as 265228.79-% 0.829-0.988, respectively. In this study, it is observed that ANN model gives better results than regression models. It is recommended to use ANN, one of the artificial intelligence methods, in similar studies.

**To Cite:** Cansız ÖF., Erginer İ., Doğru E. Amerika'da Meydana Gelen Trafik Kazalarının Yapay Sinir Ağları ve Çok Değişkenli Regresyon Yöntemleriyle Tahmini. *Osmaniye Korkut Ata Üniversitesi Fen Bilimleri Enstitüsü Dergisi* 2021; 4(3): 342-354.

## Giriş

Amerika Birleşik Devletleri yüz ölçümü açısından dünyanın en büyük 3. ülkesidir. Geniş topraklara sahip olan bu ülkede karayolları şehir içi ve şehirlerarası seyahat ve taşımacılık için en çok tercih edilen ulaşım türüdür. Karayollarının yoğun olarak kullanılması kazaları beraberinde getirmektedir. Trafik kazaları çoğunlukla ölüm ve yaralanma ile sonuçlanmaktadır. Giderek artan ulaşım ağı karayollarında kazaların sayısını da arttırmaktadır. Karayollarında taşınan yolcu ve yük emniyeti için meydana gelen kazaların sayısını azaltmak gerekmektedir. Kazaların sayısını azaltmak için kazalara neden olan faktörlerin etkilerini minimuma indirmek gerekmektedir.

Kazalara sebep olan faktörlerin hangi oranlarda araç sayısına, sürücü sayısına, nüfusa ve seyahat edilen araç miline bağlı olduğunu belirlemek kaza riskini azaltmak açısından önemli olmaktadır. Ayrıca bu faktörlerin kaza sayısını hangi oranda etkilediği bilinirse, gerçekleşecek kazaların sayısı tahmin edilebilir ve bu doğrultuda ulaştırma politikaları geliştirilebilir.

Literatürde yapay zekâ yöntemleri ve istatistiki testler kullanılarak yapılan bir çok çalışma yer almaktadır. Ataseven (2013) büyüyen teknolojiyle beraber artan hesaplama ve işleme gücüyle karmaşık simülasyonların uygulanması ve gelişmiş yapay zeka teknolojilerinden yararlanarak temel kriterlere dayalı yapılacak tahmin modellerinin uygulanmasının olağan hale geldiğini aktarmaktadır. Tahmin tekniklerinden zaman seri yöntemlerinde bulunan Box Jenkins (ARIMA) metodolojisi ve YSA yöntemlerinin tahmin performanslarını karşılaştırarak YSA kullanımının daha başarılı sonuçlar verdiğini belirtmektedir (Ataseven, 2013; Çelik, 2013).

Ramli (2011), çalışmada Ft050 Batu Palattan Kluang'a giden karayolunda Neuro Shell2 adlı YSA metodu kullanarak tahmin modeli geliştirmektedir. Model için kaza sayısının fazla olduğu yerleri inceleyerek genel kaza örüntülerini oluşturmaktadır. Modelin performansını ortalama mutlak yüzde hata kullanarak değerlendirmektedir. Sonuç olarak bir YSA modeli önermektedir (Ramli, 2011).

Cansız ve ark. (2009), YSA metodunu kullanarak motorlu taşıt kazalarında ölümcül yaralanma sayısını tahmin etmektedir. YSA kullanarak yaptıkları tahmin modeli için trafikle ilgili değişkenleri, veri seti olarak kullanmaktadır. Yaptıkları tahmin sonuçlarına göre tanjant sigmoid (tansig) transfer fonksiyonuna sahip, 14 adet nöronu ve Levenberg-Marquardt eğitim algoritmasına sahip olan ANNEFA'nın en iyi sonuçları verdiğini belirtmektedir (Cansız ve ark., 2009). Erginer ve ark. (2020),

Türkiye’de insana, araca ve yola bağlı olarak meydana gelen trafik kazalarını incelemektedir. Trafik kaza sayısını ve bu kazalardaki yaralanma sayılarının tahmini için, YSA’yı ve çok değişkenli regresyon metotlarını kullanmaktadır. Daha sonra bu iki metodu karşılaştırarak YSA ile oluşturulan tahmin modelinin daha başarılı sonuçlar elde ettiğini belirtmektedir (Erginer ve ark., 2020). Cansız ve ark. (2017), Türkiye İstatistik Kurumu (TÜİK) ile Emniyet Genel Müdürlüğü’nden yararlanarak elde edilen verilerle regresyon ve YSA metotları kullanarak tahmin modelleri oluşturmaktadır. Tahmin modelleri için bağımsız değişken olarak nüfus, sürücü sayısı, taşıt sayısı ve taşıt-km verilerini, bağımlı değişken için ise yaralı sayısı verilerini kullanarak 2’li 3’lü ve 4’lü bağımsız değişken kombinasyonları yapmaktadır. YSA ve regresyon modelleri için HKO, OYH ve R’yi değerlendirme kriteri olarak kullanmaktadır. Sonuç olarak YSA ve regresyon modellerinin kıyasında YSA modellerinin gerçeğe daha yakın sonuçlar verdiğini gözlemlemektedir (Cansız ve ark., 2017). Akgüngör ve Doğan (2009), Türkiye’nin İstanbul, Ankara ve İzmir illerinde meydana gelen trafik kazası sonucu ölenlerin sayısını 1986-2005 yılları arasındaki verilerden yararlanıp, YSA, Smeed ve Andreasen modellerini kullanarak tahmin modelleri oluşturmaktadır. Model oluşturmak için bu şehirlerdeki nüfus ve araç sayısını bağımsız değişken olarak, bağımlı değişken olarak ise ölü sayısını kullanmaktadır. YSA’da çok katmanlı, ileri beslemeli ve geri yayımlı ağ yapısı kullanmaktadır. Kullanılan modeller kıyaslandığında YSA modelinin diğer modellerden daha iyi sonuçlar verdiğini belirtmektedir (Akgüngör ve Doğan, 2009). Aghayan (2013), Kıbrıs ve İran’da bulunan otoyol kaza verilerinden yararlanarak kaza şiddetlerini tahmin etmek için YSA, genetik algoritma ve kombine genetik algoritma yöntemlerini kullanmaktadır. Kullandıkları tahmin yöntemlerinin karşılaştırılması için tepki süreleri ve doğru tahmin kriterlerini kullanmaktadır. Çalışma sonucunda YSA’nın en iyi sonuçları verdiğini belirtilmektedir (Aghayan, 2013).

Mussone ve ark. (1999), Milan kentinde olan kazaları analiz etmek için YSA metodunu kullanmaktadır. Değişik kombinasyonlar kullanarak kent için meydana gelen kavşak kazalarının tehlike derecelerini belirlemektedir. Çalışmanın sonucunda kavşak karmaşasının, kavşak düzenlemesine bağlı olan yüksek kaza indeksini çözebileceği ve karşıdan karşıya geçen yayalar için en yüksek kaza indeksinin, gece vakti sinyalle olamayan kavşaklarda meydana geldiği belirtmektedir (Mussone ve ark., 1999).

Kibar (2015), tarafından Ankara-Aksaray-Ereğli arasında kamyon trafiği yüzdesinin yaklaşık %40-50 olduğu bölünmüş kırsal yollar dikkate alınmakta ve kamyon kazaları ile trafik ve karayolu geometrik özellikleri arasındaki bağlantı YSA ve istatistiksel yöntemlerle modellenmektedir. Çalışmada öncelikle, kamyon kazaları ile trafik ve karayolu geometrik özellikleri arasındaki bağlantıyı Poisson, sıfır değer ağırlıklı negatif binom regresyon yöntemleri kullanılmaktadır. Yapılan analizler sonucunda en iyi istatistiksel yöntemin negatif binom regresyon olduğunu belirtmektedir. İkinci olarak en iyi istatistiksel sonucu veren negatif binom regresyon yöntemiyle, YSA yöntemini birbirleriyle karşılaştırıp en yakın değerlerin YSA metoduyla elde edildiğini belirtmektedir (Kibar ve ark., 2015).

Altun ve ark. (2005), 1990'lı yıllar itibariyle ulaştırma mühendisliği alanında YSA kullanımını özetleyen çalışmalarında özellikle ulaşım planlaması, bakım onarım ulaşım sistemlerinin işletilmesi, ulaşım parametrelerinin tahmini, yol üst yapısının dizaynı ve trafik mühendisliği problemlerinin çözümünde başarılı sonuçlar elde edildiğini belirtmektedir.

Cansız (2007), Türkiye'nin 1988-2005 yılları arasında ulaştırma sektörü için enerji analizi yaparak türlere göre enerji yoğunluğu değerlerini belirlemektedir. Bu değerleri kullanarak 1970-1987 ile 2006-2020 yılları arasındaki enerji tüketim tahminlerini yapmaktadır. Çalışmada YSA yönteminden yararlanarak enerji tahmin modeli geliştirmektedir. 1970-2020 arasındaki enerji tüketiminin şu anki durumunu, gelişimini ve değişimini göstermektedir (Cansız, 2007). Es ve ark. (2014), yaptıkları çalışmada YSA ve çok değişkenli regresyon kullanarak Türkiye'nin net enerji talebini tahmin etmeye çalışmaktadır. Bu tahmin için 1970-2010 yılları arasındaki nüfus, ithalat ihracat, bina yüz ölçümü, taşıt sayısı ve gayri safi hasıla verilerini kullanmaktadır. Çalışmanın sonucunda yapılan YSA modelinin tahmin performansı ve çok değişkenli regresyon yöntemiyle kıyaslanarak; YSA modelinin tahmin değerlerinin daha iyi sonuçlar verdiğini gözlemlenmektedir. Daha sonra YSA modelini kullanarak 2011-2015 yılları arasındaki net enerji talep tahmini gerçekleştirilmektedir (Es ve ark., 2014). Cansız ve ark. (2020), Türkiye karayolu taşımacılığında enerji tüketimini incelemektedir. Yapay zeka metotları içerisinde bulunan uygulamalı sinirsel bulanık çıkarım sistemi (ANFIS) ve YSA, regresyon metotlarından olan çok değişkenli lineer regresyon yöntemlerinden yararlanarak karayolu taşımacılığında enerji tüketimi için tahmin modeli oluşturmaktadır. Modellemede bağımsız değişken olarak; ağırlıklı ortama günlük trafik, motorlu taşıt sayısı, taşıt-km, karayolu yol uzunluğu ve nüfus değişkenlerini, bağımlı değişken olarak ise enerji tüketimini belirlemektedir. Tahmin modellerinin kıyaslanmasında; HKO, OYH ve determinasyon katsayısı performans kriterlerini kullanmaktadır. Çalışmanın sonucunda en iyi tahmin modelinin lineer regresyon ile elde edildiği belirtilmektedir (Cansız ve ark., 2020).

Bu çalışmada, araç sayısı, sürücü sayısı, nüfus ve seyahat edilen araç miline bağlı olarak kaza sayısı tahmini yapılmaktadır. Bu tahmin modelleri için YSA ve çok değişkenli regresyon yöntemleri kullanılmaktadır.

### **Materyal ve Metot**

Bu çalışmada nedenlerine göre Amerika karayollarında meydana gelen kaza sayısı için tahmin modelleri oluşturulmaktadır. Tahmin modellerini elde etmek için YSA ve çok değişkenli regresyon kullanılmaktadır. Tahmin modellerinde, 1994-2018 yılları arasındaki araç sayısı, sürücü sayısı, nüfus ve seyahat edilen araç mili, kaza sayısı verilerinden yararlanarak veri seti oluşturulmaktadır. Veri seti kullanılarak Matlab R2020b programında YSA ve çok değişkenli regresyon analizleri yapılmaktadır. Analizlerde veri setinin %65'i eğitim %35'i test için kullanılmaktadır. Bu veri seti Tablo 1'de verilmektedir.

**Tablo 1.**Veri seti tablosu

Yıl	Seyahat edilen araç mili	Nüfus	Araç sayısı	Lisanslı sürücü sayısı	Kaza sayısı
1994	2358	260327	192497	175403	36254
1995	2423	262803	197065	176628	37241
1996	2484	265229	201631	179539	37494
1997	2552	267784	203568	182709	37324
1998	2628	270248	208076	184861	37107
1999	2690	272691	212685	18717	37140
2000	2747	282162	217028	190625	37526
2001	2796	284969	221230	191276	37862
2002	2856	287625	225685	194602	38491
2003	2890	290108	230633	196166	38477
2004	2965	292805	237949	198889	38444
2005	2989	295517	245628	200549	39252
2006	3014	298380	251415	202810	38648
2007	3031	301231	257472	205742	37435
2008	2977	304094	259360	208321	34172
2009	2957	306772	258958	209618	30862
2010	2967	309326	257312	210115	30296
2011	2950	311580	265043	211875	29867
2012	2969	313874	265647	211815	31006
2013	2988	316058	269294	212160	30202
2014	3026	318386	274805	214092	30056
2015	3095	320743	281312	218084	32538
2016	3174	323071	288034	221712	34748
2017	3212	325147	290387	225346	34560
2018	3240	327167	297043	227558	33654

### Çok Değişkenli Lineer Regresyon

Çok değişkenli lineer regresyon, birden çok değişken kullanarak yapılan regresyon analizidir. Problemdeki bağımlı ve bağımsız değişkenler arasında bir bağlantı kurar ve bu bağlantı denklemlerle kurulur (Cansız ve ark., 2009).

Oluşturulan regresyon modellerinde çok değişkenli regresyon metotlarından lineer, interaction, quadratic ve pure-quadratic yöntemleri kullanılmaktadır. Regresyon yöntemlerinden lineer, interaction, quadratic ve pure-quadratic yöntemlerinin genel formu sırasıyla Denklem 1, Denklem 2, Denklem 3 ve Denklem 4'te verilmektedir.

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon_i \quad (1)$$

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_1 x_2 + \beta_6 x_1 x_3 + \beta_7 x_1 x_4 + \beta_8 x_2 x_3 + \beta_9 x_2 x_4 + \beta_{10} x_3 x_4 + \varepsilon_i \quad (2)$$

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_1 x_2 + \beta_6 x_1 x_3 + \beta_7 x_1 x_4 + \beta_8 x_2 x_3 + \beta_9 x_2 x_4 + \beta_{10} x_3 x_4 + \beta_1 x_1^2 + \beta_1 x_2^2 + \beta_1 x_3^2 + \beta_1 x_4^2 + \varepsilon_i \quad (3)$$

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_1 x_1^2 + \beta_1 x_2^2 + \beta_1 x_3^2 + \beta_1 x_4^2 + \varepsilon_i \quad (4)$$

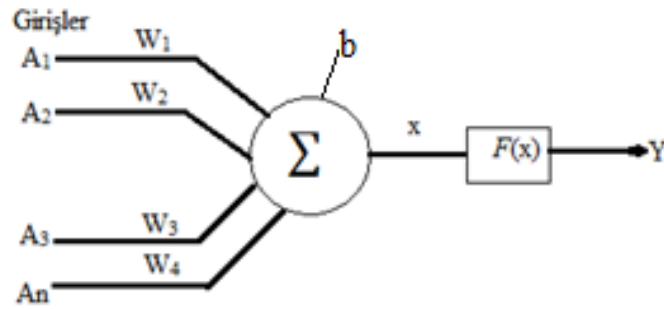
Bu denklemlerde  $x_i$  bağımsız değişken olan nüfus, seyahat edilen araç milli, kayıtlı araç ve lisanslı sürücü sayısını temsil etmektedir. “Y” ise bağımlı değişken olan kaza sayısını ifade etmektedir. “ $\beta$ ” değişken katsayılarını ve “ $\varepsilon$ ” hata bileşenini temsil etmektedir.

### Yapay sinir ağları (YSA)

Yapay zeka yöntemlerinden biri olan YSA, insan beynini modelleyerek insan beyninin öğrenme yeteneğinden ve beyindeki biyolojik sinir sisteminden yararlanılarak geliştirilen bir yöntemdir.

YSA’yı kullanarak karmaşık problemler hızlı bir şekilde çözümlenebilir. YSA öğrenme özelliği sayesinde problemleri tanıma, tahmin sunma gibi olanaklara da sahiptir (Es ve ark., 2014).

Sinir hücresi taklit edilerek programlanmış YSA, nöronlardan oluşan bir ağ yapısına sahiptir. YSA modellerinde deneme-yanılma yoluyla öğrenme işlemi uygulanmaktadır. Öğrenme bağımsız değişkenler kullanılarak bağımlı değişkenlerin tahmin edilmesi şeklinde yapılır.



Şekil 1. YSA nöron mekanizması

Şekil 1’de gösterilen YSA nöron mekanizmasında “A” girdileri, ”W” ağırlıkları, “ $\Sigma$ ” toplama fonksiyonunu, “f” aktivasyon fonksiyonunu, “b” bias değerlerini ve y değeri nöronun çıktısını temsil etmektedir.

YSA’da girdiler ağırlıkla çarpılarak sonuçları bias değeriyle toplanır ve çıkan değer transfer fonksiyonundan geçer. İşlem sonucunda elde edilen sonuç nöronun çıktısıdır. Şekil 1’ de verilen nöron mekanizması denklem olarak yazıldığında Denklem 5 elde edilir.

$$Y = f\left(\sum_{i=1}^N A_i \times W_i + bias_i\right) \quad (5)$$

Bu çalışmada YSA metodu kullanılırken veri seti -1 ile +1 arasında normalize edilmektedir. Normalize işlemi her değişken için ayrı ayrı yapılmaktadır. Normalize edilmiş değerler YSA metoduyla analiz edilerek kaza tahmin modelleri oluşturulmaktadır. Oluşturulan tahmin modelleri daha sonra denormalize edilerek gerçek kaza verileriyle karşılaştırılmaktadır. Kullanılan normalize ve denormalize işlemlerinin formülü Denklem 6 ve Denklem 7’de verilmektedir.

$$X = \frac{(y_{\max} - y_{\min}) \cdot (x - x_{\min})}{(x_{\max} - x_{\min})} - y_{\min} \quad (6)$$

$$X = \frac{(\text{Normalize deęer} + y_{\min})(x_{\max} - x_{\min})}{(y_{\max} - y_{\min})} + x_{\min} \quad (7)$$

Ayrıca bu alıřma iin tansig transfer fonksiyonu ve Levenberg-Marquardt eęitim algoritmasından yararlanılmaktadır. Tansig transfer fonksiyonunun formülü Denklem 8’de verilmektedir.

$$\text{Tansig}(N) = \frac{2}{(1 + \exp^{-2N})} - 1 \quad (8)$$

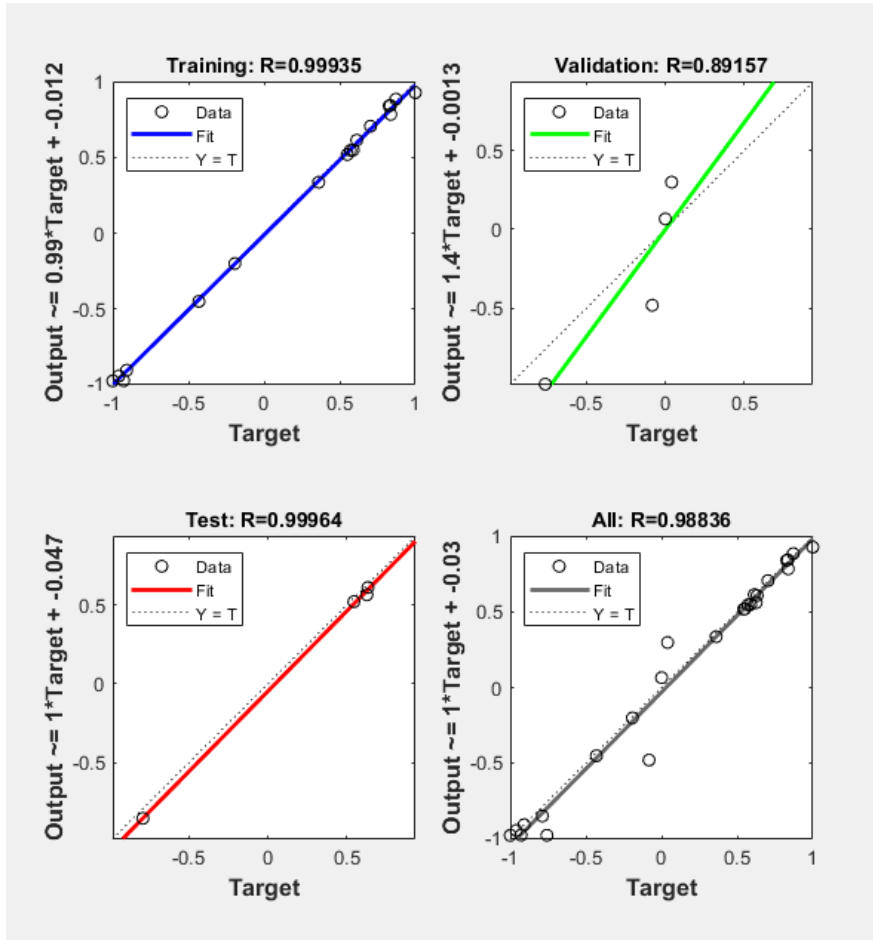
**Tablo 2.** Normalize edilen deęerler

Yıl	Seyahat edilen araç mili	Nüfus	Ara sayısı	Lisanslı sürücü sayısı	Kaza sayısı
1994	-1,000	-1,000	-1,000	-1,000	0,361
1995	-0,853	-0,926	-0,913	-0,953	0,571
1996	-0,714	-0,853	-0,825	-0,841	0,625
1997	-0,560	-0,777	-0,788	-0,720	0,589
1998	-0,388	-0,703	-0,702	-0,637	0,543
1999	-0,247	-0,630	-0,614	-0,549	0,550
2000	-0,118	-0,347	-0,531	-0,416	0,632
2001	-0,007	-0,263	-0,450	-0,391	0,704
2002	0,129	-0,183	-0,365	-0,264	0,838
2003	0,206	-0,109	-0,270	-0,204	0,835
2004	0,376	-0,028	-0,130	-0,099	0,828
2005	0,431	0,053	0,016	-0,036	1,000
2006	0,488	0,139	0,127	0,051	0,871
2007	0,526	0,224	0,243	0,163	0,613
2008	0,404	0,310	0,279	0,262	-0,083
2009	0,358	0,390	0,271	0,312	-0,788
2010	0,381	0,466	0,240	0,331	-0,909
2011	0,342	0,534	0,388	0,399	-1,000
2012	0,385	0,602	0,399	0,396	-0,757
2013	0,429	0,668	0,469	0,410	-0,929
2014	0,515	0,737	0,575	0,484	-0,960
2015	0,671	0,808	0,699	0,637	-0,431
2016	0,850	0,877	0,828	0,776	0,040
2017	0,937	0,940	0,873	0,915	0,000
2018	1,000	1,000	1,000	1,000	-0,193

## Bulgular ve Araştırma

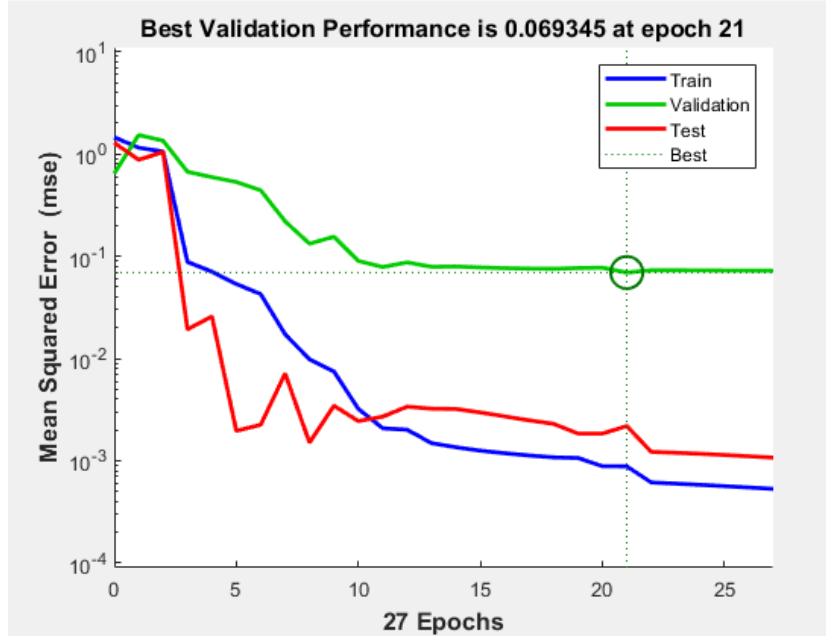
### Kaza sayısı modeli

Araç sayısı, sürücü sayısı, nüfus ve seyahat edilen araç mili değişkenleri kullanılarak oluşturulan YSA kaza sayısı tahmin modelinin korelasyon ve eğitim grafiği sırasıyla Şekil 2 ve Şekil 3'te verilmektedir. Bu modelde epoch sayısı 1000 ile sınırlandırılmaktadır. HKO değerinin 6 defa üst üste artması durumunda eğitim durdurulmakta ve performans değeri belirlenmektedir.



Şekil 2. Kaza sayısı modeli R grafiği





Şekil 3. Kaza sayısı modeli eğitim süreci

Yapılan çok değişkenli regresyon analizleri sonucunda lineer, interaction, quadratic ve purequadratic regresyon denklemleri sırasıyla Denklem 9, Denklem 10, Denklem 11 ve Denklem 12’de oluşturulmaktadır.

Bu denklemlerde “VT” seyahat edilen araç mili, “P” nüfus, “V” araç sayısı, “D” lisanslı sürücü sayısını ifade etmektedir.

Lineer regresyon denklemi;

$$C=103065,63+28675,17*VT-402,95*P+155,22*V-342,63*D \quad (9)$$

Interaction regresyon denklemi;

$$C=108052,92-514725,68*VT-4029,43*P+113,84*V+12790,87*D+6981,32*VT*P-2039,29*VT*V-088,16*VT*D-25,06*P*V-50,95*P*D+66,71*V*D \quad (10)$$

Quadratic regresyon denklemi;

$$C=-73116,30-289965,04*VT+2776,00*P-689,50*V+2290,78*D+4163,13*VT*P-1128,25*VT*V-2830,48*VT*D-0,17*P*V+21,45*P*D+14,37*V*D-11381,89*VT^2-32,53*P^2+2,61*V^2-11,58*D^2 \quad (11)$$

Purequadratic regresyon denklemi;

$$C=-289994,19-24971,41*VT+3731,51*P-1706,71*V+420,60*D+9943,17*VT^2-6,57*P^2+3,67*V^2-2,49*D^2 \quad (12)$$

Çalışmada tahmin modelleri oluşturulurken veri setinde kullanılan değişkenlerin en büyük ve en küçük değerleri Tablo 3’de verilmektedir.

**Tablo 3.** Verilerin en büyük ve en küçük değerleri

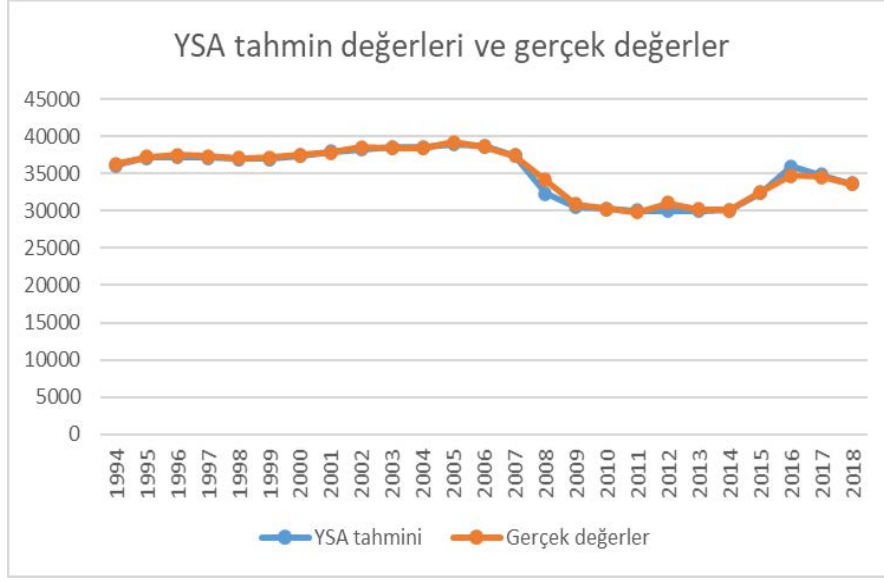
Değişkenler	En Büyük Değer	En Küçük Değer
Seyahat edilen araç mili	3,24	2,358
Nüfus	327,167	260,327
Araç sayısı	297,043	192,497
Lisanslı sürücü sayısı	227,558	175,403
Kaza sayısı	39,252	29,867

YSA ve çoklu regresyon yöntemleri ile yapılan analizlerin R, HKO, OYH kullanılarak karşılaştırılması Tablo 4’te verilmektedir.

**Tablo 4.** Model karşılaştırması

		R	HKO	OYH(%)
Kaza Sayısı	Lineer Regrasyon	0,845824	2219668	3,760
	Interaction Regrasyon	0,895189	1611269	3,294
	Quadratic Regrasyon	0,927922	1172709	2,606
	Purequadratic Regrasyon	0,940816	1158293	2,585
	YSA	0,988360	265229	0,829

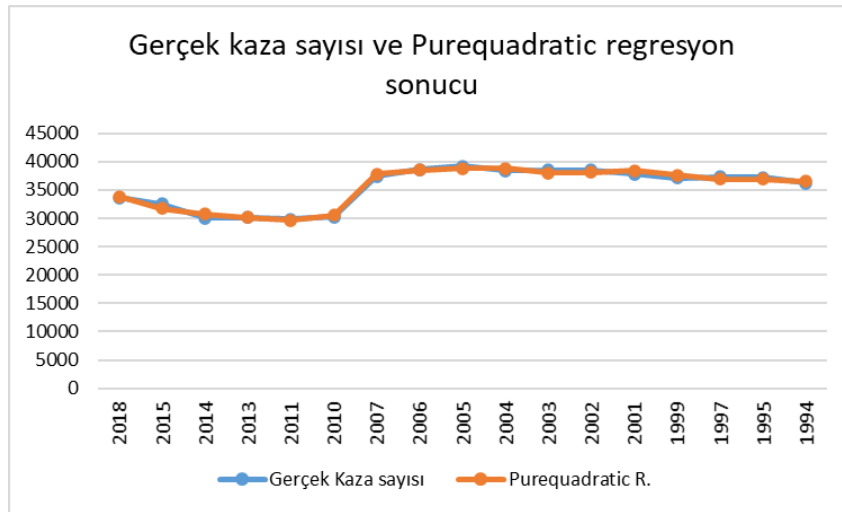
Tablo 4’te kaza sayıları için oluşturulan tahmin modellerinin kıyaslaması gösterilmektedir. Bu sonuçlara göre YSA modeli çok değişkenli regresyon türlerinden daha başarılı sonuçlar vermektedir. Elde edilen modellerde YSA modelinin R değeri 0,98’dir. HKO değerinin en düşük olduğu model quadratic regresyon modeli olup, OYH değerinin en düşük olduğu model ise YSA modelidir. Oluşturulan YSA modelinde kaza sayısı değişkenini en çok etkileyen değişken seyahat edilen araç mili değerinin olduğu görülmektedir.



Şekil 4. YSA modeli tahmini- Kaza sayısı değeri ilişkisi

Şekil 4'te en iyi tahmin sonuçlarını veren YSA'nın tahmin değerleri ve gerçek değerlerin tablo ile karşılaştırması verilmektedir. Şekil 4 incelendiğinde, tahmin sonuçları ve gerçek değerlerin uyumlu olduğu görülmektedir. 2007-2009 ve 2015-2017 yılları arasında küçük sapmalar olduğu görülmektedir. YSA modelinin performansının artırılmasında ve tahmin sonuçlarına yakın değerlere ulaşılmasında, kullanılan epoch sayısının 1000 ile sınırlandırılması ve 6 defa ard arda HKO değerinin artmasıyla eğitimin durdurulmasının önemli olduğu görülmektedir. Ayrıca model eğitimi yapılırken tanjant sigmoid transfer fonksiyonunun kullanılmasıyla model gücünün arttığı görülmektedir.

Şekil 5'te regresyon modelleri içerisinde en iyi performans değerine sahip pure-quadratic regresyon modelinin tahmin sonuçları ile gerçek kaza sayısı değerleri arasındaki ilişki verilmektedir.



Şekil 5. Pure-Quadratic regresyon modeli tahmini-Kaza sayısı değeri ilişkisi

Şekil 5 incelendiğinde kaza sayısı değerleri ile pure-quadratic regresyon modelinin tüm yıllar için örtüştüğü ve büyük sapmaların oluşmadığı görülmektedir.

### **Sonuçlar**

Yapılan bu çalışmada araç sayısı, sürücü sayısı, nüfus ve seyahat edilen araç mili bağımsız değişkenleri kullanılarak ABD kaza sayısı tahmin modelleri geliştirilmektedir. Bu modeller YSA ve çok değişkenli regresyon modelleri kullanılarak oluşturulmaktadır. Oluşturulan kaza sayısı tahmin modellerinin sonuçları HKO, OYH ve R değerleriyle kıyaslandığında en iyi sonuçları YSA tahmin modelinin verdiği görülmektedir.

Çok değişkenli regresyon yöntemleri kullanılarak yapılan kaza sayısı tahmin sonuçları kendi içerisinde karşılaştırıldığında en iyi tahmin sonuçlarını pure-quadratic regresyon modelinin verdiği görülmektedir. Pure-quadratic regresyon yönteminde değişkenlerin karelerinin de alınarak farklı bir değişken olarak modele girtilmesi, modelin gücünün artmasına sebep olmaktadır.

Kaza sayısı tahmin modellerinde kullanılan değişkenler değerlendirildiğinde en etkili değişkenin seyahat edilen araç milli olduğu görülmektedir.

Bu çalışma sonucunda görüldü ki; araç sayısı, sürücü sayısı, nüfus ve seyahat edilen araç mili değişkenlerinden yararlanarak yapılan kaza sayısı tahminleri başarılıdır.

### **Çıkar Çatışması Beyanı**

Makale yazarları herhangi bir çıkar çatışması olmadığını beyan eder.

### **Araştırmacıların Katkı Oranı Beyan Özeti**

Yazarlar makaleye eşit oranda katkı sağlamış olduğunu beyan eder.

### **Kaynakça**

Aghayan I. Improved traffic crash modeling through accuracy and response time using classification algorithms: A model comparison approach, Eastern Mediterranean University (EMU), Doctoral dissertation 2013.

Akguuml AP., Doğan E. An application of modified Smeed, adapted Andreassen and artificial neural network accident models to three metropolitan cities of Turkey. Scientific Research and Essays 2009; 4(9): 906-913.

Altun İ., Dündar S., Yöntem K. Yapay sinir ağları ile trafik akım kontrolü. Deprem Sempozyumu, Kocaeli 2005; 1335-1344.

Ataseven B. Yapay sinir ağları ile öngörü modellenmesi 2013.

- Cansız ÖF., Çalışıcı M., Ünsalan K. Türkiye karayollarında meydana gelen kazalarda oluşan yaralı sayısı için tahmin modellerinin oluşturulması. 2. Uluslararası Mühendislik ve Tasarım Kongresi, Kocaeli 2017.
- Cansız ÖF. Enerji politikalarının ulaştırma sistemlerinin optimizasyonu ile geliştirilmesi ve uygulamadan elde edilen getirilerin ortaya konması. Fen Bilimleri Enstitüsü, İnşaat Mühendisliği Anabilim Dalı, Doktora Tezi 2007; 352s.
- Cansız ÖF., Ünsalan K., Erginer İ. Karayolları enerji tüketiminin yapay zeka ve regresyon yöntemleri ile modellenmesi. Uludağ Üniversitesi 2020.
- Cansız ÖF., Çalışıcı M., Miroglu MM. Use of artificial neural network to estimate number of persons fatally injured in motor vehicle accidents. In Proceedings of the 3rd International Conference on Applied Mathematics 2009; 136-142.
- Çelik Ş. Zaman serileri analizi ve trafik kazası verilerine uygulanması. Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi 2013; 3(4): 43-51.
- Erginer M., Cansız ÖF., & Erginer İ. Trafik kaza sayısının ve yaralı sayısının yapay sinir ağları ile tahmini. Osmaniye Korkut Ata Üniversitesi Fen Bilimleri Enstitüsü Dergisi 2020; 3(1): 29-35.
- Es HA., Kalender FY., Hamzaçebi C. Yapay sinir ağları ile Türkiye net enerji talep tahmini. Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi 2014; 29(3): 495-504.
- Kıbar FT. Türkiye'de kamyon kazaları ile trafik ve karayolu geometrik özellikleri arasındaki ilişkinin istatistiksel ve yapay sinir ağları yöntemleri ile modellenmesi. Doktora tezi, Karadeniz Teknik Üniversitesi/Fen Bilimleri Enstitüsü 2015.
- Mussone L., Ferrari A., Oneta M. An analysis of urban collisions using an artificial intelligence model. Accident Analysis & Prevention 1999; 31(6): 705-718.
- Ramli MZ. Development of accident prediction model by using artificial neural network (ANN). Doctoral dissertation, Universiti Tun Hussein Onn Malaysia 2011.