



Makine Öğrenmesi Algoritmaları Kullanılarak Türkiye ve AB Ülkelerinin CO₂ Emisyonlarının Tahmini

Serpil Aydın^{1*}, Gül Aydoğdu²

¹ Ondokuz Mayıs Üniversitesi, Fen Edebiyat Fakültesi, İstatistik Bölümü, Samsun, Türkiye (ORCID: 0000-0001-6985-6120)

² Gazi Üniversitesi, Fen Edebiyat Fakültesi, İstatistik Bölümü, Ankara, Türkiye (ORCID: 0000-0002-3395-3321)

(5th International Symposium on Innovative Approaches in Smart Technologies– 28-29 May 2022)

(DOI: 10.31590/ejosat.1129958)

ATIF/REFERENCE: Aydın, S. & Aydoğdu, G. (2022). Makine Öğrenmesi Algoritmaları Kullanılarak Türkiye ve AB Ülkelerinin CO₂ Emisyonlarının Tahmini. *Avrupa Bilim ve Teknoloji Dergisi*, (37), 42-46.

Öz

Özellikle 20. ve 21. yüzyıllarda sanayi devriminin hızlanması ve bundan dolayı Dünya'nın küresel ısınması ile sıcaklık miktarlarının artması, yeryüzünde mevsimlerin döngü ve şiddetlerinin değişmesine sebep olmuştur. Hayatın her alanında ihtiyaç duyulan enerji kullanımının artması, özellikle de fosil yakıtlara yönelim, sera gazları arasında önemli bir paya sahip olan CO₂'nin salınımının artmasına ve iklim değişikliklerine neden olmuş ve olmaya da devam etmektedir. Dolayısıyla Dünya üzerindeki tüm ülkeler için insan faaliyetlerinden kaynaklı CO₂ emisyon miktarlarının yönetimi, takibi ve tahlili önem kazanmıştır. Bu çalışmada, iklim değişikliği ile ilişkili sera gazı olan CO₂'nin Türkiye ve AB ülkelerinin tahmini ve eğilimlerinin irdelenmesi hedeflenmiştir. Bu hedef doğrultusunda, makine öğrenmesi algoritmalarından karar ağacı modellemesi, yapay sinir ağları ve destek vektör makineleri kullanılarak enerji elde edilirken ortaya çıkan CO₂ salınım miktarları için tahmin yapılmıştır. Makine öğrenme metodlarının başarı performansı karar ağacı modellemesi skor değeri, ortalama hata karesi (MSE) ve ortalama hata karesinin kökü (RMSE) kullanılarak değerlendirilmiştir. Ülkeler için 2010-2019 yıllarındaki nüfus, CO₂ emisyon miktarı, enerji tüketim miktarları kullanılarak, destek vektör makineleri (DVM) ile yapay sinir ağları (YSA) metodları uygulanmış ve tahminler yapılmıştır. Sonuçlara göre DVM yönteminin, YSA yöntemine göre daha başarılı sonuçlar verdiği tespit edilmiştir. Avrupa Birliği ülkelerine toplu olarak bakıldığında, CO₂ emisyon miktarının azalarak arttığı yani artış hızının azaldığı görülmüştür. Ayrıca Almanya, Hollanda, İtalya, Fransa ülkelerinde CO₂ emisyon miktarlarında azalma beklenmektedir. Türkiye için ise durum farklıdır. Türkiye'de 2010-2019 yılları arasındaki karbondioksit salınım miktarı her yıl artmakta olup önümüzdeki yıllarda da alternatif, CO₂ salınımı olmayan enerji kaynaklarına yönelinmez ise bu artışın devam edeceği yapılan çalışmada görülmüştür.

Anahtar Kelimeler: Makine Öğrenmesi, Emisyon, Destek Vektör Makineleri, Karar Ağacı, Yapay Sinir Ağları

CO₂ Emissions in Turkey and EU Countries Using Machine Learning Algorithms

Abstract

Especially in the 20th and 21st centuries, the acceleration of the industrial revolution, the global warming of the world, and the increase in the amount of temperature caused the cycles and intensities of the seasons on earth to change. The increase in the use of energy needed in all areas of life, especially the orientation to fossil fuels, has caused and continues to cause an increase in the emission of CO₂, which has an important share among greenhouse gases, and climate changes. Therefore, the management, follow-up, and analysis of CO₂ emissions caused by human activities have gained importance for all countries in the world. This study, it is aimed to examine the estimation and trends of CO₂, a greenhouse gas associated with climate change, in Turkey and EU countries. In line with this goal, an estimation was made for CO₂ emissions that occur when energy is obtained by using decision tree modeling, artificial neural networks, and support vector machines from machine learning algorithms. Success performance of machine learning methods was evaluated using decision tree modeling score value, mean square error (MSE) and root mean square error (RMSE). Support vector

* Sorumlu Yazar: serpil.gumustekin@omu.edu.tr

machines (SVM) and artificial neural networks (ANN) methods were applied, and estimations were made by using the population, CO₂ emission amount, and energy consumption amount for the countries in the years 2010-2019. According to the results, it has been determined that the SVM method gives more successful results than the ANN method. When the European Union countries are considered as a whole, it has been observed that the amount of CO₂ emissions have decreased and increased, that is, the rate of increase has decreased. In addition, reductions in CO₂ emissions are expected in Germany, the Netherlands, Italy, and France. For Turkey, the situation is different. The amount of carbon dioxide emissions between 2010-2019 in Turkey is increasing every year, and it has been seen in the study that this increase will continue if alternative energy sources that do not emit CO₂ are not used in the coming years.

Keywords: Machine Learning, Emission, Support Vector Machines, Decision Tree, Artificial Neural Networks

1. Giriş

İklim değişikliği Dünya üzerindeki tüm canlı hayatı üzerinde yaşamı farklılaştıracak büyük bir değişimdir. Mevsimlerin süreleri, sıcaklık ortalamaları değişmektedir. Genel olarak; okyanus sıcaklık artışları, yükselen deniz seviyeleri, kutuplarda buzulların erimesi, kasırgalar, sıcak hava dalgaları, orman yangınları, kuraklıklar, seller ve yağışlar gibi aşırı hava koşullarında sıklık ve şiddet değişiklikleri ile bitki örtüsü değişikliklerini kapsar. Isınma olayları buharlaşmayı etkilemekte ve kuraklığa neden olmaktadır. Dolayısıyla bitkilerin hasat zamanı verimlilikleri değişmektedir. Yüksek CO₂ seviyeleri ürünlerin verimi üzerinde olumsuz etkilere sahip olabilmektedir. İklim değişikliği bitkileri, hayvanları ve insanları etkileyen yeni zararlı hastalıkların ortaya çıkmasına ve insan sağlığı için yeni riskler oluşturmaya neden olabilmektedir.

Küresel enerji talebi 2020'nin ilk çeyreğinde %3,8 azalmakla birlikte doğalgazın kullanımına dayalı ekonomik faaliyetler üzerinde pandemiden kaynaklı olarak, 2020'nin ilk çeyreğinde fazla bir etki yaratmamıştır. Ayrıca bazı ülkelerde tam kapanma döneminde elektrik enerji talebinde yaklaşık %20 oranında düşüşler görülmüştür [1]. COVID-19'dan kaynaklanan aksaklıklara rağmen, küresel sera gazı emisyonları 2020'de artmıştır. Küresel olarak, deniz seviyesi 2020'de zirve yaparak yılda ortalama 3,29 (+/-0,3) mm yükselmiştir [2].

Son dönemlerde yapay zekâ çalışmaları karmaşık problemler için bağlantı kurma, gruplandırma ve tahmin çalışmalarında bir hayli uygulanmakta ve güvenilir sonuçlar vermektedir. Gelişen teknoloji, hızlı bilgi kaydı hem veri miktarlarını çok arttırmış hem de çok çeşitli, karmaşık veri yığınlarının oluşmasına neden olmuştur ve olmaya da devam etmektedir. Artık veri analizlerinde sadece istatistiksel yöntemler yeterli olmamakta bunun yanında veri madenciliği yöntemlerine de ihtiyaç duyulmaktadır [3]. Makine öğrenmesi ise yapay zekâ çalışmalarının yaklaşım türlerinden biri olup, yeni bir yaklaşım türüdür. Makine öğrenmesindeki temel amaç, elimizdeki veriler ile bu verilerin sonuçları arasındaki ilişkiyi bulmaktır. Dolayısıyla, denemeler gerçekleştirilmeden mevcut verilere uygun bir makine öğrenme yöntemi belirlenmemektedir. Mevcut veri hacmi arttıkça ve veriler üzerinde farklı yöntem denemesi mümkün olduğunca artırılırsa en doğru yöntemin bulunması sağlanır. Makine öğrenimi uygulamaları ile kalite denetim süreçlerinde, reklam ve promosyon çalışmalarında, bankacılık işlemlerinde, metin çalışmalarında, görüntü ve ses tanımlamaları gibi pek çok alanda çalışabilmektedir [4].

Sözen ve ark. (2007) sektörlerden kaynaklı CO₂ emisyon miktarlarını yapay sinir ağlarıyla gelecek için CO₂ salınım miktarı tahmininde bulunmuştur. Çalışma CO₂ emisyon verilerinin trendi üzerinden yapılmıştır [5]. Yılmaz ve Yılmaz (2013) 1990 – 2009

yıllarındaki CO₂ verileri ile 2010 – 2020 yıllarındaki tahmini CO₂ salınım miktarlarını hesaplamışlardır. Çalışmada gri tahmin metodu kullanılmıştır [6]. Papuççu ve Bayramoğlu (2016) enerji üretimi ile tüketimi ile endüstriyel üretim ve ulaşım amaçlı enerji kullanım miktarlarını, CO₂ emisyon tahmini için kullanmışlardır. CO₂ salınım miktarları, yapay sinir ağları yöntemi ile AB – 28 ülkeleri ve Türkiye için tahmin edilerek 2020, 2025 ve 2030 yılları için CO₂ miktarları hesaplanmıştır [7]. Behdioğlu ve Çelik (2016), 1996 – 2013 yılları arasında OECD ülkelerine ait AR-GE harcamaları ile emisyon verileri arasındaki ilişkiyi, STIRPAT modeli üzerinden yapay sinir ağları kullanarak analiz etmişlerdir. Sonuçlara göre, OECD ülkelerinde yapılan Ar-Ge harcamaları, emisyonları negatif yönde etkilemiştir [8]. Kılınç ve Altıparmak (2020), 2020 yılında çevre vergilerinin, CO₂ emisyonu üzerindeki etkisine yönelik bir çalışma yapmışlardır. Analiz sonuçlarına göre 21 AB ülkesi ile Türkiye’de, çevre vergileri ve Ar-Ge enerji harcamaları, CO₂ emisyon salınımını negatif yönde, GSYH miktarı ile enerji tüketimi ise CO₂ emisyon salınımını pozitif yönde etkilemektedir [9].

Bu çalışmada, Türkiye ve AB ülkeleri için makine öğrenmesi yöntemleri ile çeşitli yakıt tiplerine göre tüketilen enerji miktarları ve ülkelerin nüfus miktarları kullanılarak, iklim değişikliğine sebep olan CO₂ emisyonu tahmininin yapılması amaçlanmıştır.

2. Materyal ve Metot

Çalışmada, Türkiye ve Avrupa Birliği ülkeleri incelenmiştir. Ülkelerin CO₂ emisyonlarının tahmininin yapılması için 2010-2019 yılları arasındaki “Toplam Nüfus, Katı Fosil Yakıt Tüketimi, Doğalgaz Tüketimi, Petrol Tüketimi, Güneş Enerjisi Tüketimi, Biyogaz Tüketimi, Birincil Katı Biyoyakıt Tüketimi, Yenilenebilir Belediye Atık Tüketimi, Jeo Termal Enerji Tüketimi, Hidro Elektrik Üretimi” verileri incelenmiştir. Çalışmada kullanılan değişkenlerden, tüketilen enerji miktarlarının yanması sonucu açığa çıkan CO₂ emisyonu verileri %80’i öğrenme, %20’si ise test amaçlı kullanılmıştır. Makine öğrenme metodlarından karar ağacı modellemesi, destek vektör makineleri ve yapay sinir ağları teknikleri kullanılarak uygulama yapılmıştır. Makine öğrenme metodlarının başarı performansı karar ağacı modellemesi skor değeri, MSE ve RMSE kullanılarak değerlendirilmiştir.

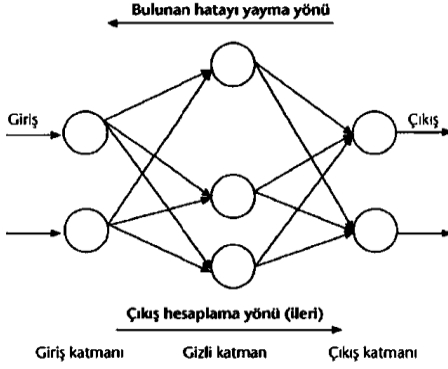
2.1. Makine Öğrenmesi

Makine öğreniminin temeli, belirli bir konudaki veriler üzerinden algoritmalar oluşturarak, çıktılarını yeni veriler çıktıkça güncelleyerek, sonuçların tahmini için istatistiksel analizleri kullanmaya dayanmaktadır. Makine öğrenimi danışmanlı (gözetimli), danışmansız (gözetimsiz), yarı danışmanlı ve takviyeli öğrenme olarak dört gruba ayrılır. Danışmanlı (gözetimli) öğrenmede, model geliştirmek için sonuçları belli bir eğitim verisi bilgisi bulunmaktadır. Danışmansız (gözetimsiz) öğrenmede ise veri setindeki bilgiler kullanılarak, değişkenler

arasındaki ilişkilerin ortaya çıkarılması sağlanır. Makine öğreniminde en çok kullanılan öğrenme algoritmaları; yapay sinir ağları, destek vektör makinaları, karar ağaçları ve lojistik regresyon olarak sıralanabilir [10].

2.2. Yapay Sinir Ağları

Temelde yapay sinir ağları, insan beyninin çalışma şeklinin örnek alınarak geliştirilmiş, bağlantılar aracılığıyla birbirleri arasında bağlanan, her biri kendi belleğine sahip, biyolojik sinir ağları gibi karar alıp hareket eden bilgisayar programlarıdır [11].



Şekil 1. Yapay Sinir Ağları

Yapay sinir ağları mevcut verilerle bilgi sağlamakta, bu bilgiler üzerinden genel bir çıkarım yapmakta sonrasında yeni durumlar ile karşılaştırılınca elde ettiği bilgilerden faydalanarak yeni durumlar için bir sonuç çıkarmaktadır. Yapay sinir ağları, öğrenbilme-genelleştirme özellikleri sayesinde pek çok alanda uygulanabilmekte ve karmaşık sorunları çözebilme imkânı sunmaktadır [12]. Yapay sinir ağlarının öğrenbilme yeteneği farklı olup, farklı farklı algoritmalarla öğrenbilirler. Bilinmeyen çıktılar için sonuç bilgisi üretebilirler. Örüntüleri kavrayıp, çözebilir ve sınıflandırma yapabilirler. Net olmayan ya da eksik bilgi ile çalışabilirler.

2.3. Destek Vektör Makineleri

Destek Vektör Makineleri (DVM), hem sınıflandırma hem de regresyon çalışmaları konusunda uygulanan, güçlü bir denetimli makine öğrenmesi algoritmasıdır. Genellikle sınıflandırma problemlerinde, örüntü tanıma çalışmalarında, hastalık teşhisi konulmasında, zaman serisi analizi gibi pek çok alanda kullanılmaktadır. DVM'nin asıl amacı iki farklı gruba göre etiketlenmiş bir eğitim verisi için geliştirilen algoritması ile yeni bir gözlem verildiğinde bu gözlemin hangi grupta olması gerektiğini belirten bir model oluşturulmasıdır. Yani iki grubun ayrımını en iyi yapabilen hiper düzlemin bulunmasıdır [13].

2.4. Karar Ağacı Algoritması

Verilerin sınıflandırılması, makine öğreniminde iki aşamalı bir süreç olup, bu aşamalar; öğrenme ve tahmin aşamalarıdır. Öğrenme aşamasında, model eğitim verileri kullanılarak geliştirilmiştir. Tahmin aşamasında ise model, verilen veriler ile tahmin için kullanılır. Karar ağacı, anlaşılması ve yorumlanması

en kolay, yaygın sınıflandırma algoritmalarından biridir. Karar ağacı algoritması, danışmanlı öğrenme algoritmaları grubuna aittir. Diğer danışmanlı öğrenme algoritmalarının aksine, karar ağacı algoritması regresyon ve sınıflandırma problemlerini çözmek için de kullanılabilir. Bir karar ağacı kullanmanın amacı eğitim verilerinden çıkarılan basit karar kurallarını öğrenerek hedef değişkenin sınıfını veya değerini tahmin etmek için kullanılacak bir eğitim modeli oluşturmaktır. Karar ağaçlarında, sınıf etiketini tahmin etmek için ağacın kökünden işleme başlanır. Kök özneliğinin değerleri, kaydın özneliğiyle karşılaştırılır. Karşılaştırmaya dayanarak, bu değere karşılık gelen dalı takip edilir ve bir sonraki düğüme geçilir. Karar ağaçlarının türleri, sahip olunan hedef değişkenin türüne göre kategorik ve sürekli değişkenli karar ağacı olmak üzere iki türdür [14].

2.5. Tahminler için Performans Metrikleri

Makine öğrenmesi çalışmaları tamamlandıktan sonra kullanılan metotların, veriler üzerindeki başarı skorunun değerlendirilmesi gerekmektedir. Yapılan tahminlerin, başarı durumlarının sayısal bir değer olarak yorumlanabilmesi için karar ağacı modellemesi skoru, 'ortalama mutlak hata (MAE), ortalama hata karesi (MSE), ortalama mutlak yüzde hata (MAPE) ve ortalama hata karesinin kökü (RMSE) gibi yöntemler bulunmaktadır.

2.5.1. Ortalama Hata Karesi (MSE)

Ortalamadan sapmaların karelerinin toplanması ve bu değerinin ortalamasının alınması ile "Ortalama Hata Karesi (MSE)" bulunup, performans değeri ifade edilmektedir. Eşitlik (1)'e göre çıkış sayısı "n" ile gösterilir. Burada, e_t , gerçek değerden tahmin edilen değer çıkarılmasıyla elde edilir.

$$MSE = \frac{1}{n} \sum_{t=1}^n e_t^2 \quad (1)$$

2.5.2. Ortalama Hata Karesinin Kökü (RMSE)

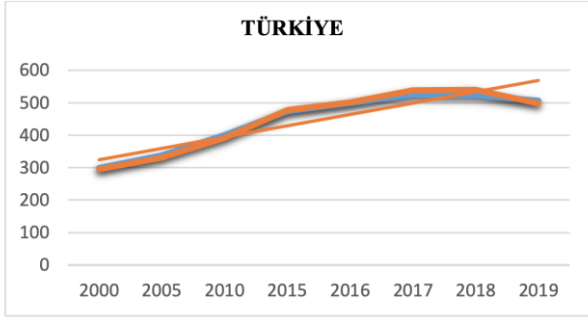
Ortalama hata karesinin kökü çok kullanılan performans ölçütlerinden biridir. Eşitlik (2)'de gösterildiği gibi hesaplanır.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2} \quad (2)$$

3. Araştırma Sonuçları ve Tartışma

3.1. Bulgular

Türkiye'nin 2010 - 2019 yıllarındaki CO₂ emisyon verilerinin tahmininde, enerji tüketim verilerine, karar ağacı modellemesi, yapay sinir ağları yöntemi ve destek vektör makineleri yöntemi uygulanmıştır. Verinin %80' i öğrenme, %20' si ise test için kullanılmıştır. Ayrıca çalışmada, tahminlerin başarı performansı değerlendirmesi için karar ağacı modellemesi skor değeri, ortalama hata karesi (MSE) ve ortalama hata karesinin kökü (RMSE) kullanılmıştır. Tablo 1'de Türkiye'nin CO₂ emisyonu performans sonuçları ve Şekil 2'de CO₂ tahmin değeri ile CO₂ gerçek değeri gösterilmektedir.

Şekil 2. CO₂ tahmin değeri ve CO₂ gerçek değeri

AB ülkelerinin CO₂ emisyonu performans sonuçları Tablo 2’de gösterilmektedir.

Tablo 1. Türkiye CO₂ emisyonu performans değerlendirilmesi

	Karar Ağacı Test Verisi Skoru	MSE	RMSE
Karar Ağacı Modellemesi	0,96	*	*
YSA	*	7477,06	86,47
DVM	*	27483,00	165,78

Tablo 2. AB Ülkeleri CO₂ emisyonu performans değerlendirilmesi

	Karar Ağacı Test Verisi Skoru	MSE	RMSE
Karar Ağacı Modellemesi	0,98	*	*
YSA	*	8608,12	92,78
DVM	*	71765,05	267,89

Ülkeler için 2010-2019 yıllarındaki nüfus, CO₂ emisyon miktarı, enerji tüketim miktarları kullanılarak, destek vektör makineleri (DVM) ile yapay sinir ağları (YSA) metotları uygulanmış ve tahminler yapılmıştır. Sonuçlara göre DVM yönteminin, YSA yöntemine göre daha başarılı sonuçlar verdiği tespit edilmiştir. Avrupa Birliği ülkelerine toplu olarak bakıldığında, CO₂ emisyon miktarının azalarak arttığı yani artış hızının azaldığı görülmüştür. Ayrıca Almanya, Hollanda, İtalya, Fransa ülkelerinde CO₂ emisyon miktarlarında azalma beklenmektedir. Türkiye için ise durum farklıdır. Türkiye’de 2010-2019 yılları arasındaki karbondioksit salınım miktarı 2019 yılına kadar her yıl artmakta olup önümüzdeki yıllarda da alternatif, CO₂ salınımı olmayan enerji kaynaklarına yönelmez ise bu artışın devam edeceği yapılan çalışmada görülmüştür.

4. Sonuç

Ülkemizde ve Dünya’da karbondioksit miktarının artışı durdurulamaz ya da azaltılmaz ise, mevsim kaymaları, çok yüksek sıcaklık değerleri, yoğun yağış sebebi ile sel ve su baskınları artarak devam edecektir. CO₂ salınım miktarındaki yükselişin temel sebebi tüm Dünya’da olduğu gibi Türkiye’de de enerji talebinin artmasıdır. En çok kullanılan ve CO₂ salınımını en çok etkileyen fosil enerji kaynakları; kömür, petrol, doğal gaz olup bunların dışında nükleer enerji, hidro elektrik enerjisi ve

Yenilenebilir enerji kaynakları bulunmaktadır. Bu nedenle, Türkiye dâhil tüm ülkelerin karbondioksit salınım miktarlarını geriye çekmek için gerekli önlemler alınmalıdır. Ülkelerin fosil yakıtları kullanmak yerine, yenilenebilir enerji kaynaklarına yönelmesi gerekmektedir.

Literatürde CO₂ emisyonları tahmini konusunda pek çok çalışma mevcuttur. Ancak bu tahminlerin büyük bir bölümü istatistiksel yöntemler ile gerçekleştirilmektedir. Bu çalışmada karbondioksit salınım miktarını ön görmek için makine öğrenmesi yöntemleri uygulanmıştır. Sonrasında ise tahmin metotlarının başarı skoruna göre metotlar kıyaslanmıştır. Gelecek çalışmalarda değişken sayısı artırılarak büyük veri çalışmaları yapılabilir.

Kaynakça

- [1] IEA 2020: <https://www.iea.org/reports/world-energy-outlook-2020>
- [2] Küresel İklim Durumu 2020; <https://www.mgm.gov.tr/FILES/Haberler/2021/WMO->
- [3] Aktan, E. (2018). Büyük veri: Uygulama alanları, analitiği ve güvenlik boyutu. Bilgi Yönetimi, 1(1), 1-22.
- [4] Garip, E. (2017). OECD ülkelerindeki karbondioksit (co2) emisyonunun makine öğrenmesi ile tahmin edilmesi (Master's thesis, Fen Bilimleri Enstitüsü).

- [5] Sözen, Adnan & Gülseven, Zafer & Arcaklıođlu, Erol, 2007. " Türkiye'de sera gazlarının sektörel enerji tüketimine dayalı tahmin ve azaltım politikaları" *Energy Policy*, Elsevier, vol. 35(12), sayfa 6491-6505, Aralık.
- [6] Yılmaz, H., & Yılmaz, M. (2013). Forecasting CO2 emissions for Turkey by using the grey prediction method. *Sigma*, 31, 141-148.
- [7] Papuçcu, H. ve Bayramođlu, T. (2016). Yapay Sinir Ağları ile CO2 Tahmini: Türkiye Örneđi, *İktisadi İdari Bilimler Dergisi*, 18, s. 762–778.
- [8] Behdiođlu, S., & Çelik, F. (2016). Ar-Ge Harcamalaeri Ve Emisyonu: Yapay Sinir Ađı Yaklaşımı. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, 136-150.
- [9] Kılınç, E. C., & Altıparmak, H. (2020). Çevre Vergilerinin Co2 Emisyonu Üzerindeki Etkisi Üzerine Bir Uygulama. *Odü Sosyal Bilimler Araştırmaları Dergisi(ODÜSOBİAD)*, 10(1), 217-227.
- [10] Alpaydın, E. (2021). *Machine learning*. MIT Press.
- [11] Elmas, Ç. (2007), *Yapay Zeka Uygulamaları*, Ankara, Seçkin Yayıncılık
- [12] Ergezer, H., Dikmen, M., & Özdemir, E. (2003). *Yapay sinir ağları ve tanıma sistemleri*. PiVOLKA, 2(6), 14-17.
- [13] Noble, W. S. (2006). What is a support vector machine?. *Nature biotechnology*, 24(12), 1565-1567.
- [14] Brijain, M., Patel, R., Kushik, M. R., & Rana, K. (2014). A survey on decision tree algorithm for classification. *International Journal of Engineering Development and Research*, 2(1), ISSN:2321 – 9939.