Avrupa Bilim ve Teknoloji Dergisi
Sayı 20, S. 248-253, Aralık 2020
© Telif hakkı EJOSAT'a aittir
Araştırma Makalesi



www.ejosat.com ISSN:2148-2683

European Journal of Science and Technology
No. XX, pp. 248-253, December 2020
Copyright © 2020 EJOSAT
Research Article

Ağır Akut Solunum Yolu Yetersizliği Sendromu Koronavirüsü 2 (SARS-CoV-2) Dünya Çapındaki Analizi ve Gelecekteki Durumu

Mehmet Bilge Han Taş[1*], Mete Yağanoğlu[2]

[1] Atatürk Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Erzurum, Türkiye (ORCID: 0000-0001-6135-1849), bilgehantas@gmail.com
[2] Atatürk Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Erzurum, Türkiye (ORCID: 0000-0003-3045-169X), yaganoglu@atauni.edu.tr

(İlk Geliş Tarihi 6 Temmuz 2020 ve Kabul Tarihi 9 Ekim 2020)

(**DOI:** 10.31590/ejosat.764638)

ATIF/REFERENCE: Taş, M. B. H. & Yağanoğlu, M. (2020). Ağır Akut Solunum Yolu Yetersizliği Sendromu Koronavirüsü 2 (SARS-CoV-2) Dünya Çapındaki Analizi ve Gelecekteki Durumu. *Avrupa Bilim ve Teknoloji Dergisi*, (20), 248-253.

Öz

Coronaviridae ailesinin bir cinsi olan Koronavirüsler, büyük bir artı iplikli RNA genomuna sahip zarflı virüslerdir. Serolojik olarak farklı üç koronavirüs grubu tanımlanmıştır. Her grupta virüsler konakçı aralıkları ve genom dizileri ile karakterize edilir. Koronavirüsler farelerde, sıçanlarda, tavuklarda, hindilerde, domuzlarda, köpeklerde, kedilerde, tavşanlarda, atlarda, sığırlarda ve insanlarda tanımlanmıştır ayrıca gastroenterit ve solunum yolu hastalıkları gibi çeşitli ciddi hastalıklara neden olabilir. Bulaşıcılığı çok yüksektir. Solunum yoluyla bulaşabilen bu virüs, damlacık yoluyla da aktarılması yani bulaşması çok kolaydır. Yapılan çalışmaların gösterdiği gibi yakın temasta bulunan bireylerin birbirlerine çok kısa bir zaman içinde bulaştırdığı gözlemlenmiştir. Bunun yanında bulaşan virüs, solunum yollarına yerleşerek akciğerlere oturur ve orada bölünmeye devam eder. Bunun sonucunda akciğerlere büyük ve kalıcı hasar vererek konak canlıya ölümcül bir şekilde zarar vermektedir. Yeni koronavirüsün (COVID-19) küresel etkisi büyüyen bir ivmeyle dünyayı sarmaktadır. Doğrulanan vakaların doğru bir şekilde tahmin edilmesi ve sağlıklı bir çıkarım yapabilmek için yalnızca vakalar değil ayrıca ölüm ve iyileşme sayısının analizini de gerektirir. Bununla birlikte, öngörme için geniş bir geçmiş verisi gerekir. Aynı zamanda, gelecek nadiren geçmişle aynı şekilde tekrarlandığı için hiçbir tahmin kesin değildir. Ayrıca, tahminler verilerin güvenilirliğinden, kazanılmış ilgi alanlarından ve hangi değişkenlerin tahmin edildiğinden etkilenir. Psikolojik faktörler, insanların hastalıktan kaynaklanan tehlikeyi ve onları kişisel olarak etkileyebileceği korkusunu nasıl algıladığı ve bunlara nasıl tepki gösterdiği konusunda önemli bir rol oynamaktadır. Bu çalışma COVID-19'un devamını basit, ancak güçlü bir yöntemle öngörmek için objektif bir yaklaşım ortaya koymaktadır. Kullanılan veriler Dünya Sağlık Örgütü'nün (WHO) paylaşmış olduğu ve her ülkenin bu bilgileri paylaştığı güvenilir verilerdir. Hastalığın yapılan tahminler sonucunda teyit edilmiş COVID-19 vakalarında sürekli bir artış olduğunu göstermektedir. Risk günden güne artmakta ve kismi ölçüde düşüşe geçmeye meyilli bir hâl almaktadır. Bu çalışma planlama ve karar verme için büyük etkileri olan canlı bir tahmin çalışmasının zaman çizelgesini açıklamakta ve teyit edilmiş COVID-19 vakaları için objektif tahminler sunmaktadır.

Anahtar Kelimeler: Coronavirus, Covid-19, Veri Bilimi, Makine Öğrenmesi, Tahmin Etme

Worldwide Analysis and Future Status of Severe Acute Respiratory Failure Syndrome Coronavirus 2 (SARS-CoV-2)

Abstract

Coronaviruses, a genus of the Coronaviridae family, are enveloped viruses with a large plus-strand RNA genome. Three serologically different coronavirus groups have been identified. Viruses in each group are characterized by host ranges and genome sequences. Coronaviruses have been identified in mice, rats, chickens, turkeys, pigs, dogs, cats, rabbits, horses, cattle and humans, and can cause a variety of serious illnesses such as gastroenteritis and respiratory diseases. Its infectiousness is very high. This virus that can be transmitted by inhalation is also very easy to transmit via droplet. As the studies show, it has been observed that individuals who are in close contact infect each other for a very short time. In addition, the infected virus settles in the respiratory tract, sits in the lungs and continues to divide there. As a result, the host harms the creature fatally by causing great and permanent damage to the lungs. The global

[*] Sorumlu Yazar: Atatürk Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Erzurum, Türkiye, ORCID: 0000-0001-6135-1849, bilgehantas@gmail.com

impact of the new coronavirus (COVID-19) surrounds the world with a growing momentum. It requires not only the cases but also the analysis of the number of deaths and recovery in order to accurately predict the confirmed cases and make a healthy conclusion. However, extensive historical data is required for forecasting. At the same time, no prediction is certain, as the future is rarely repeated in the same way as in the past. In addition, estimations are impressed by the reliability of the data, gained interests and which variables are estimated. Psychological factors play an important role in how people perceive and react to the danger caused by the disease and the fear that it can affect them personally. This study presents an objective approach to predict the continuation of COVID-19 with a simple but robust method. The data used are reliable data that the World Health Organization (WHO) and each country shared. As a result of the estimations, it is seen that there is a continuous increase in confirmed COVID-19 cases. The risk is increasing day by day and partially tending to decrease. This study describes the timeline of a live estimation study with major impacts for planning and decision making and provides objective estimations for confirmed COVID-19 cases.

Keywords: Coronavirus, Covid-19, Data Science, Machine Learning, Forecast.

1. Giriş

Yeni koronavirüsün ortaya çıkması, aşılar ve tedavi için hedeflenen terapötiklerin kullanılamadığı bir pandemik solunum yolu hastalığına (COVID-19) neden olmuştur (Wang ve diğerleri, 2020). Salgın, dünya çapında halk sağlığı hakkında büyük endişelere neden olmaktadır. Aynı zamanda, hanenin virüsün yayılmasını yavaşlatmak için evde kalması gerektiği için ekonomik sonuçlarla ilgili endişeler artmaktadır. Ekonominin duraklamanın; tedarik zincirleri, firmaların, finans sektörünün ve hane halklarının finansal istikrarı üzerindeki etkisi büyük ölçüde bilinmemektedir. Sonuç olarak, politikacılar, işletmeler ve piyasa katılımcıları büyüme beklentilerini kısa, orta ve uzun vadede revize etmeye çalışmaktadırlar (Gormsen ve Koijen, 2020). Koronavirüsün etkileri çok geniş bir skala içinde düşünülebilir. İnsan sağlığı açısından çok tehlikeli olması sebebiyle maske ve eldiven kullanımı birçok ülkede zorunlu hale gelmektedir. Bunun yanında tüm dünya ülkerinde görülen bu pandemi sağlık sektörünü de çok büyük bir şekilde etkilemektedir. Yeterli hastane sayısı olmaması, hastanelerde yeterli sayıda solunum cihazı olmaması gibi faktörler ölüm oranlarını ciddi şekilde artırmaktadır. Bunların yanında birçok ülkede uygulanan sokağa çıkma yasağı veya sokağa çıkma tedbirleri gibi uygulamalar, tüm dünya ülkerinde ekonomik bir buhrana yol açmaktadır. Para akışı çok sınırlanmış ve kişilerin hayatta kalabilecek kadar yaşam standartlarına girmesi sebebiyle birçok ülke ekonomik anlamda sıkıntı çekmektedir. Sokağa çıkma yasağı veya kısıtlaması uygulanan ülkelerde, evde kalan hane halkı hem ekonomik hem de sağlık sıkıntıları çekmelerinin yanında bir de psikolojik anlamda bunalıma girmesi söz konusudur. Sosyal hayatı sınırlanan hane halkı büyük ölçüde zorluk çekmektedir. Tüm bu negatif yansımaların yanında bir de doğa açısından bakarsak, doğa kendini tamir etmeye başlamış ve bir nebze de olsa temizlenmeye çalışmıştır.

Bu çalışmada tüm dünya ülkelerinden toplanan verilerin istatiksel bir anlamda analizini sunmakla birlikte geleceğe dair birtakım çıkarımlarda bulunulmaya çalışılmıştır. Yapılan gözlemler makine öğrenimi yöntemiyle bir tahmin etme çalışmasına dökülmüştür. Özellikle Türkiye'nin seçilmiş birkaç ülkeyle birlikte karşılaştırması yapılmıştır. Bunun yanı sıra Türkiye'nin diğer ülkelere göre ne durumda olduğu ve gelecekte Koronavirüsün ülkede ne seviyede olacağına dair objektif bir yorum katacaktır. Bu yorumları yapabilmek adına veriler her ülkenin beyan ettiği veriler doğrultusunda, makine öğrenmesi kullanılarak ve bu veriler ışığında regresyon eğrileri çıkartarak, geleceğe dair birtakım tahminler üretmek mümkündür. Bu tahminleri değerlendirip önümüzdeki zaman doğrultusunda nasıl bir ivme kazandığı söylenebilir. Bu ivmenin birçok değişkeni içerdiği ve söz konusu tedbirlerin gevşetilmesi gibi durumlarda değişiklik gösterebilmektedir. Çünkü bu veriler geçmiş günlerden alınan kurumsal bilgilerin, geleceğe dair bir çıkarımı sonucunda yapılmaktadır.

Günümüzde artan veri miktarı ve kaynak erişiminin kolaylığından dolayı veri analizi için otomatik yöntemlere olan ihtiyaç artmaya devam etmektedir. Makine öğrenmesinin amacı, verilerdeki kalıpları otomatik olarak algılayabilen yöntemler geliştirmek ve daha sonra gelecekteki verileri veya ilgilenilen diğer sonuçları tahmin etmek için ortaya çıkarılan kalıpları kullanmaktır. Makine öğrenimi bu nedenle istatistik ve veri madenciliği alanlarıyla yakından ilişkilidir. Ancak vurgusu ve terminolojisi açısından biraz farklıdır. Çalışmada makine öğrenimi yapılarak, toplanan veriyi kendi terminolojisi içinde değerlendirerek, bilgisayarın insanlara sağlamış olduğu hesaplama hızını en yüksek seviyede kullanıp, işleyerek bizlere istenen sonuçları vermesi açısından çok önemli bir görev üstlenmektedir (Murphy, 2012).

2. Materyal ve Metot

2.1. Veri Seti Tanımı

Veri seti alınırken temelde üç farklı öznitelik ele alınarak veri seti kullanımı yapılmıştır. Bunlar; teyit edilmiş vakalar, ölümler ve iyileşenler. Bu veriler Johns Hopkins Üniversitesi Sistem Bilimi ve Mühendisliği Merkezi (CSSE) tarafından oluşturulmuş veriler olup buradan temin edildi (https://github.com/CSSEGISandData/COVID-19 14/06/2020 tarihinde erişildi). Her gün kaydedilen toplam vakalar, 22 Ocak 2020 ile 15 Mart 2020 arasındaki süreyi kapsamaktadır. Hem "laboratuvar onaylı" hem de "klinik olarak teşhis edilmiş" vakalar dahil edildi. Onaylanmış vakalar veya ölümler kadar medyada yer almayan iyileşmiş vakaların da büyük bir önemi mevcuttur. Çünkü verilerin doğruluğu medyada oluşmuş veriler değil, hükûmetlerin resmi organlarından duyurmuş olduğu verilerdir. Her üç veri modeli de devamlı artış gösterirken hem teyit edilen vakaların hem de ölümlerin eğilimleri dünya genelinde şubat ayında, Türkiye'de ise nisan ortasında azalmıştır; Güney Kore, İran ve Avrupa'daki vaka sayısındaki artışın bir sonucu olarak şubat ve mart ayı sonlarında ikinci bir üstel artış görülmektedir. Aynı zamanda, iyileşen vakaların sayısı giderek artmaktadır.

2.2. Veri Analizi

Eğitim setlerinin her biri bir tahmin üretmek için ve algoritma eğitmek için kullanıldı. Bu tahminleri üretebilmek için veri, çeşitli algoritmalar kullanılarak test edildi. Bu sınıflandırma kuralı daha sonra görülmeyen iki bileşikten hangisinin en büyük aktiviteye sahip olduğunu tahmin etmek için kullanılabilir. Böyle bir kuralın genelleme yeteneği, yani etkinlik açısından görünmeyen iki

bileşiği doğru bir şekilde sıralayabilme olasılığı tahmin edilir. Eğitimli sınıflandırıcıların her biri, karşılık gelen test setini sınıflandırmak için kullanılır. Bu hata oranının her biri gerçek hata oranının bir tahminidir ve nihai tahmini ortalamaları olarak alınır. Verilerin ortaya koyduğu bilgiler ışığında veri setleri bağımsız olarak çizilse bile, sınıflandırma test setlerinin sınıflandırma eğitim setlerinden bağımsız olmadığı unutulmamalıdır. Böylece, gerçek hata oranı tahmini aşağıya doğru eğimli olacak ve görülmeyen vakalarda hata oranının iyimser bir tahminini verecektir. Bununla birlikte, tekniklerin karşılaştırılması için yeterlidir (Burbidge ve diğerleri, 2001).

Kullanılan bu algoritmalar, her biri kendi içinde optimize edilerek en iyi sonucu vermesi ve tahmin edilecek verinin ne kadar yanından geçmesi gerektiği test edilerek eniyileştirme yapılmıştır. Bulgular sonucunda gene en yüksek değeri yani bugüne kadar gelmiş vakaların geleceğe ışık tutması için bu verilerle en iyi eşleşebilen bir doğruluk eğrisi bulunmaya çalışılmıştır. Bulunan bu eğri ile gelecekte virüsün ne durumda olacağı ve ne kadar kişiyi etkileyeceği ortaya konulmaya çalışılmıştır. Bu veriler resmi kayıtlarla tespit edilmiş günlük vaka sayıları, günlük ölüm sayıları ve günlük iyileşen kişilerin sayılarından oluşmaktadır. Bu veriler analiz edilerek en iyi tahmini yapabilmek adına kullanılmıştır. Çalışmada kullanılan algoritmalar ve eğrileri şu şekildedir; lineer regresyon eğrisi, polinominal eğrisi, destek vektör makinesi (SVM), karar ağacı ve rasgele orman algoritması ve eğrisi kullanılmıştır.

2.3. Kullanılan Algoritmalar

2.3.1 Lineer Regresyon

Lineer regresyon, gözlemlenen veya elimizde olan verilere doğrusal/lineer bir denklem yerleştirerek veya başka bir deyişle bir doğru çizen iki değişken arasındaki ilişkiyi gösterebilmek için kullanılan istatistiksel bir yöntemdir. Bir değişkenin bağımsız bir değişken olduğu, diğerinin bağımlı bir değişken olduğu kabul edilerek kullanılır. Daha önce verilen değer ile yani bağımlı değişkenle diğer değişkenler arasındaki korelasyonun sürdürülmesi ve bu korelasyonun bir bağlamda en iyi şekilde ayırabilmesini sağlar. Daha önce verilmiş bu bağımlı değişken kümesinin diğer bağımsız değişkenler arasındaki farkı ortaya koyabilmek adına en iyi şekilde ama doğrusal bir şekilde bir çizgi ile bu kümülatif değerlerin birbirinden ayrılmasına olanak sağlar (Song ve diğerleri, 2005).

2.3.2 Polinominal Regresyon

Fuller (1969) ve Gallant (1974) tarafından açıklanan segmentli polinom modelleri, bir dizi uyarlanmış alt modelden oluşur. Bir bağımsız değişken olması durumunda, alan bölümlere ayrılır ve bölümün her bir bölümünde farklı bir polinom alt modeli tanımlanır. Bu alt modeller daha sonra bölümlere ayrılmış polinom modeli oluşturmak üzere birlikte uygulanır. Bazen veriler arasında direkt bir korelasyon mevcut olmayabilir. Bu durumda polinomial regresyon eğrisi bu korelasyon olmayan veriler arasında yani düzensiz ve karmaşık verilerden bir eğri uydurma konusunda kullanılabilir.

2.3.3 Destek Vektör Makinesi (Support Vector Machine (SVM))

SVM, sınıflandırma ve örüntü tanıma için kullanılan çok basit ve etkili bir sınıflandırma algoritmasıdır. Destek Vektör Makineleri algoritması 1995 yılında Vladimir Vapnik tarafından tanıtıldı. Bu algoritmanın temel amacı hiper düzlemleri veya sınırları inşa eden bir fonksiyon elde etmektir. Bu hiper düzlemler, farklı girdi veri noktaları kategorilerini ayırmak için kullanılır. SVM ikili sınıflandırma kullanır ve değerleri belirli bir spesifikasyona göre ayırt etmek için yüksek boyutlardaki özellik uzayında hiper düzlemler kullanan sistemlerdir. Hiper düzlemler, istatistiksel öğrenmeyi kullanmak için belirli algoritmalar ile eğitilir. SVM sınıflandırma yöntemi, özellik çıkarmayı içeren ve istenen çıktıları üreten denetimli öğrenmeye benzer. SVM'nin avantajı, eğitilmesinin çok kolay olmasıdır. Yüksek boyutlu verileri sinir ağlarından daha iyi ölçekleyebilir. Tipik olarak iki tip SVM sınıflandırıcısı vardır bunlar doğrusal ve doğrusal olmayandır. Ayrıca SVM, ortam denetiminde kullanılabilen çekirdek işlevlerine sahiptir (Jain ve diğerleri, 2020; Yağanoğlu ve Köse, 2017).

2.3.4 Karar Ağacı

Karar ağacı, denetimli öğrenme algoritmalarına ait bir makine öğrenme algoritmasıdır. Sınıflandırma problemlerinin çözümünde kullanılır. Bu çalışmada, önceki verilerden (eğitim ve test aşamaları) elde edilen bazı türetilmiş karar kurallarına dayalı olarak bağımlı değişkeni tahmin etmek için bir karar ağacı sınıflandırıcısı kullanılmıştır. Örneklerin özelliklerini sınıflandırmak için kök düğümlerin kullanıldığı düğümler ve düğümler olarak temsil edilir. Yaprak düğümleri (çocuksuz düğümler) kararları veya sınıflandırmaları gösterir. Her aşamadaki diğer tüm özellikler arasında en yüksek kazancı (en homojen dallar) değerlendirmek, her bir düğümdeki bir karar ağacının temel seçimidir. Karar ağacının performansı bir karışıklık matrisi kullanılarak değerlendirilir (Qawqzeh ve diğerleri, 2019).

2.3.5 Rastgele Orman

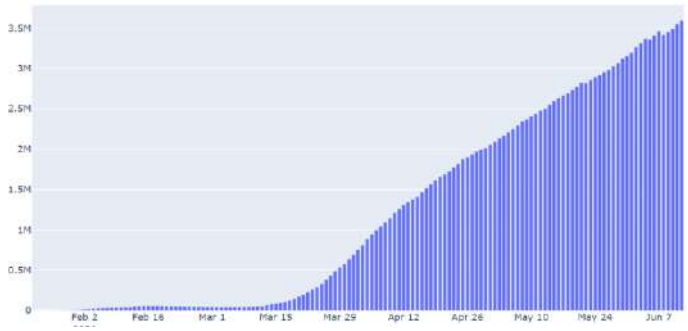
Sınıflandırma yöntemi olarak Rastgele orman (Random Forest) sınıflandırıcısı kullanılmıştır. Rastgele orman sınıflandırıcısı, her sınıflandırıcının girdi vektöründen bağımsız olarak örneklenen rastgele bir vektör kullanılarak oluşturulduğu, ağaç sınıflandırıcıların bir kombinasyonundan oluşur ve her ağaç, bir giriş vektörünü sınıflandırmak için en popüler sınıf için bir birim oyu verir yani sayımlama yapar (Breiman, 1999). Bir karar ağacının tasarımı, bir nitelik seçim ölçüsü ve budama yönteminin seçimini gerektirir. Karar ağacı indüksiyonu için kullanılan özelliklerin seçiminde birçok yaklaşım vardır ve çoğu yaklaşım doğrudan niteliğe bir kalite ölçüsü katar. Karar ağacı indüksiyonunda en sık kullanılan nitelik seçimler bilgi kazanç oranı kriteri (Quinlan, 2014) ve Gini Endeksi'dir (Brieman ve diğerleri, 1984). Rastgele orman sınıflandırıcısı, Gini Dizini'ni, özniteliklerin sınıflara göre safsızlığını ölçen bir öznitelik seçim ölçüsü olarak kullanır. Rastgele orman kısaca sınıflandırıcısı ormanlar arasında gezerek, dallara doğru bir ilerleme yapar. Dalların sonuna geldiğinde bir oylama yapar. Burada "n" gezilmesi gereken ağaç sayısını ifade eder. Örneğin n sayısını 4 seçersek, 4 adet ağaç arasından sonuca gidilir ve bir oylama yapılır. Oylama sonucunda en çok oy almış sınıf veya tahmin sonucu elde eldilir. Bu sayede başarılı bir tahmin veya sonuç fırsatı yakalanmış olur.

3. Araştırma Sonuçları ve Tartışma

3.1. Dünyadaki Veriler

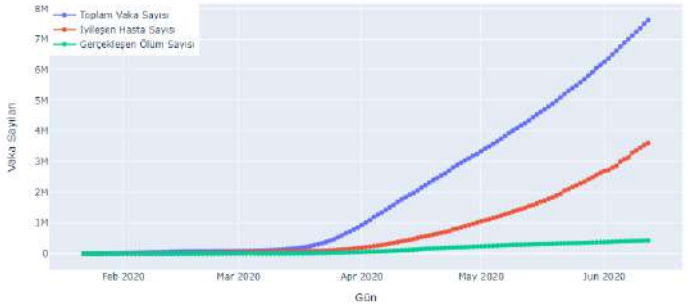
Dünyada gitgide tırmanmakta olan Koronavirüs salgını, son günlerde biraz duraklamaya geçmiş olsa da Şekil-1'de görüldüğü üzere dünyada toplam aktif vaka sayısı 3.5 milyonu geçmektedir.

Tüm dünya ülkeleri bu konuda önlem almasına rağmen virüsün kolay bulaşması ve insanları kolay bir şekilde konak olarak ölmesine neden olmaktadır. Dünyada toplam vaka sayısı ise 7.5 milyonu geçmiş durumdadır.



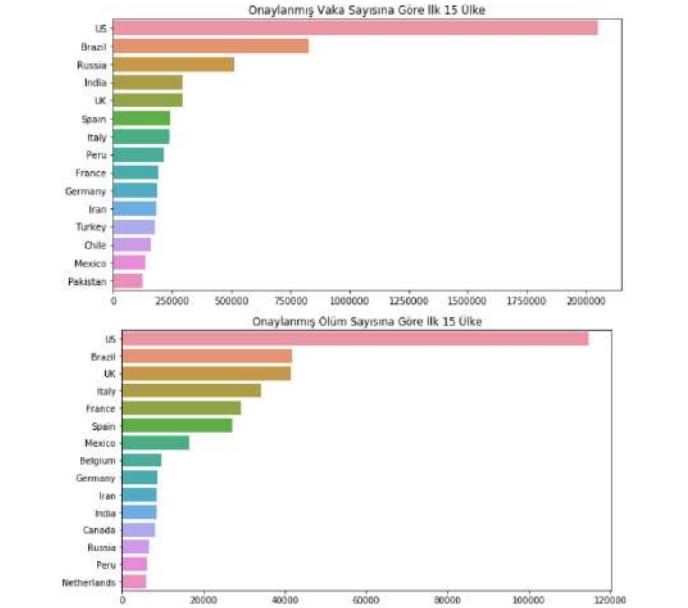
Şekil-1 (Dünyada Görülen Toplam Aktif Vaka Sayısı)

Dünyada etkisini gösteren bu pandemide iyileşen hasta sayısı 3.5 milyonu geçmektedir. Öte yandan ölen insan sayısı 425 bin üzerindedir. Tüm dünyada bu sayılar durağan hale gelirken eğer önlemler aynı şekilde alınmaya devam edilmezse bu değerler artan bir grafikte devam edecektir.



Şekil-2 (Dünyadaki Toplam Vaka, İyileşen Hasta, Ölüm Sayıları)

Dünyadaki her ülkede görülen koronavirüs gün geçtikçe artan bir grafikle ilerlemektedir. Virüsün çıktığı ülke olan Çin en yüksek hasarı almamasına rağmen, Amerika bu virüsten çok büyük bir ölçüde etkilenmiş durumdadır. Son günlerde Brezilya ve Rusya'daki vaka artışı da koronavirüsün henüz kontrol altına alınamadığının en büyük kanıtıdır.



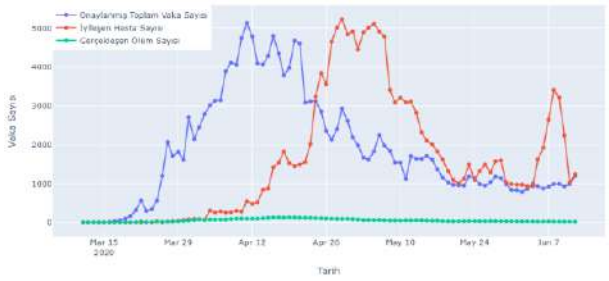
Şekil-3 (Dünyada en çok görülen 15 ülkenin vaka sayıları ve ölüm sayıları)

3.2. Türkiyedeki Veriler

Türkiye, Dünya Sağlık Örgütü'nün de (WHO) dikkatini çeken bir şekilde bu pandemi dönemini geçirmektedir. Birçok ülke turizm için Türkiye'yi güvenli ülke olarak düşünmektedir. Bunun en önemli sebeplerinden biri Türkiye'de yapılan özverili çalışmalardır. Bu çalışmaların yapılmasında öncelikli olarak Türk doktorlarının özverisi daha sonra medikal malzemelerin üretimi noktasında yapılmış çalışmalar ile birçok dünya ülkesine bu pandemi sürecinde çeşitli yardımlarda da bulunulmuştur. Şekil-4'te görüldüğü gibi görülen vaka eğrisinin burnu aşağı doğru ivme almış ve duraklama aşamasına gelmiştir. Bunun yanında iyileşen hasta sayısıyla birlikte aktif vaka sayısı 30 bin civarlarındadır. Ölüm oranının diğer ülkelere nazaran çok düşük olmasının en temel sebebi Türk doktorları ve sağlık çalışanlarının göstermiş olduğu özveriyle direkt olarak bağlantılıdır. Pandemi döneminde uygulanan sağlık politakasının da etkisi büyüktür. Dünya ülkelerinde genel olarak yapılan uygulama, yoğun bakıma alınan hastalara ilaç tedavisi başlanmaktadır. Ancak Türkiye'de alınan kararla durumu ciddiye gittiği düşünülen hastalara direkt olarak ilaç tedavisi başlanmaktadır. Bu da ölüm oranının çok düşük olmasının en temel sebeplerinden birisidir. Günlük vaka sayıları Şekil-5'te verildiği üzere çok düşük seviyelere gelmişken son günlerde yukarıya doğru bir artış trendi görülmektedir.

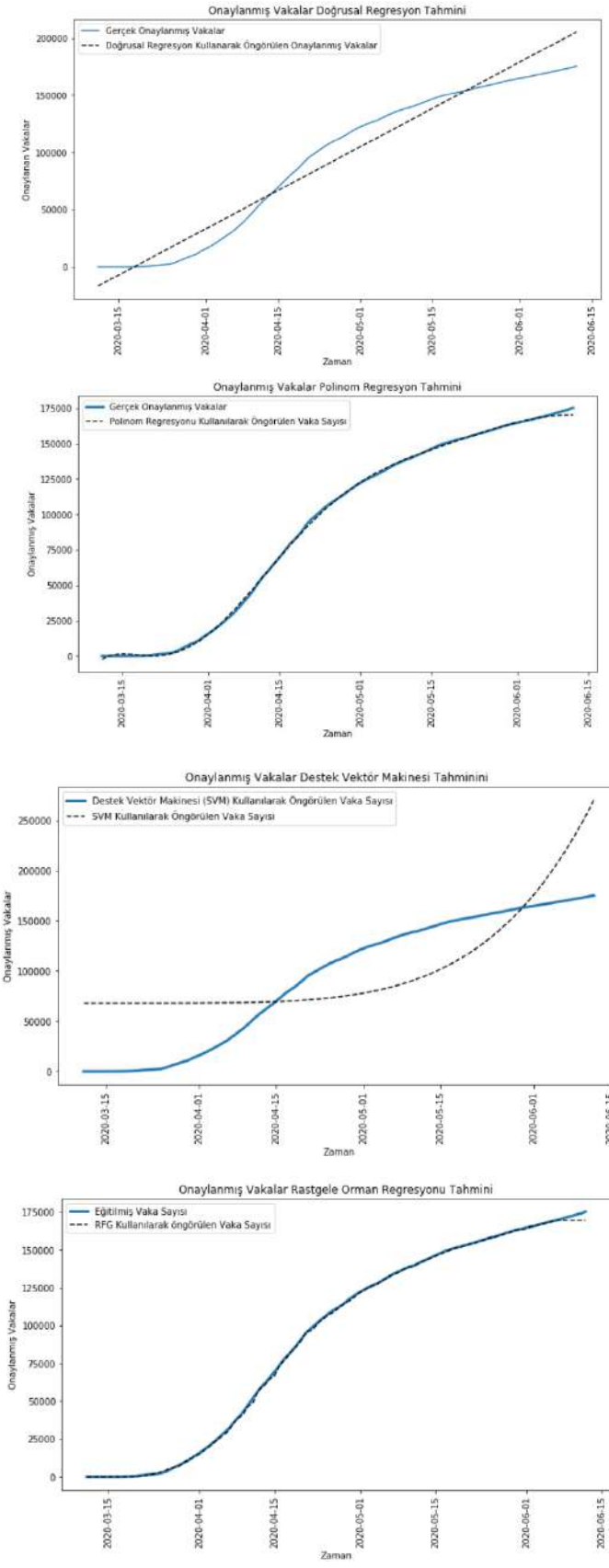


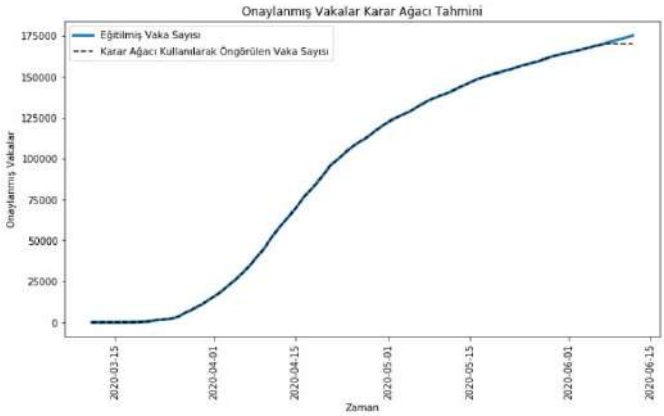
Şekil-4 (Türkiye'deki toplam vaka, iyileşme oranı, ölüm oranları)



Şekil-5 (Türkiye'de görülen günlük vaka, iyileşen, ölüm sayıları)

3.3. Veriler ile Gelecekteki Durumunun Tahmin Edilmesi





Şekil-6 (Türkiye'deki vakaların soldan sağa doğru sırasıyla Lineer, Polinomial, SVM, Rastgele Orman, Karar Ağacı eğrileriyle belirlenmesi)

Kullanılan çeşitli algoritma karşılaştırmaları sonucunda, gerçek rakamlarla oluşturulmuş eğri grafiği ile tahmin edilmiş eğri grafiği göz önünde bulundurulduğunda en çok örtüşme görülen eğri tahmin için kullanılabilmektedir. Bu bilgi ışığında kullanabileceğimiz üç adet grafik mevcuttur. Şekil-6'da verildiği üzere Polinom Regregyon, Rastgele Orman ve Karar Ağacı eğrileri bu veri seti için en iyi tahmin yapabilen eğriler olduğu görülmektedir. Küçük farklılar gösterse de en çok örtüşen algoritma, Karar Ağacı algoritmasıdır. Karar Ağacı regresyon eğrisine göre 178 bin olan toplam vaka sayısı gün geçtikçe azalarak yaklaşık 2 aylık bir dilimde 254 bin civarına gelerek hızla biteceği çıkarımını yapabilmekteyiz. Dünyadaki artış trendinin aksine belli başlı birkaç ülke dışında artık vaka sayıları duraklamaya girmiş olup, aynı şekilde Türkiye'de de duraklamaya girmiştir. Bu çalışmada kullanılan eğri uydurma yöntemleri ile geleceğe dair bir çıkarım yapılabilmekte olup, bu çıkarımların şu anki günümüz perspektifine göre objektif bir şekilde değerlendirilmiştir. Bu sonuçlar eğer günümüz tedbirleri devam ettiği takdirde gerçekleşecek sonuçlardır. Eğer ki ilerleyen günlerde tedbiri kararları değişirse ve insanlar bu tedbirlerin aksine hareket ederse bu sonuçlardan çok daha fazla şekilde etkilenebilme durumu söz konusudur.

4. Sonuç

Koronavirüs dünyanı hızla saran ve bulaşıcılığı çok yüksek bir virüs türüdür. Bu bulaşıcılığı solunum yolundan bulaşabildiği için çok hızlı yayılım göstermektedir. Aynı zamanda hayvanlardan da geçebildiği için çok geniş bir konak yelpazesi vardır. Ancak gerekli tedbirle alınırsa ve bu tedbirler sabırlı bir şekilde uygulanabilirse bu bulaşıcılığı azalmaktadır. Bulaştığı canlılarda çeşitli kronik hastalığı olanlar daha fazla olmak üzere kalıcı hasarlar bırakmaktadır. Bunun yanında bu virüs geniş yayılma ağı sayesinde çok fazla canlıya ve insana erişimi mümkündür. Mümkün mertebe bu virüsten korunma yoluna başvuran ülkeler genellikle daha az can kaybıyla veya daha az kronik hastayla pandemi dönemini atlatabilecektir. Türkiye'nin bu konudaki durumu çok açık bir şekilde göstermektedir ki gerekli önlemler ve sağlık protokolleriyle başarıya ulaşabileceği yadsınamaz bir gerçektir.

Bu çalışmada makine öğrenmesi kullanılarak ve bunu destekleyecek yaygın algoritmalar kullanılarak, geleceğe dair bir çıkarımda bulunulmuştur. Bu çıkarımlar günümüz perspektifi ve

gelecekteki şartların tahayyülü ile daha doğru bir çıkarımda bulunulmuştur. Günümüz şartlarındaki tedbirlerin devamı ile daha hızlı bir iyileşme aşamasına dönülebilir. Kullanılan algoritmalar kendi içlerinde optimize edilerek, diğer algoritmalarla kıyası yapılmıştır. Bu eniyileme işlemleri sonucunda ve geçmişteki veriler ışığında en uygun sonuçlara götürecek biçimde eğriler oluşturulmuştur. En uygun gelen algoritmayla yetinmeyip, diğer algoritmaları da bu problemin çözümüne uygun bir eğri çıkarabilmek için hepsini eniyileyerek kullanılmıştır. Daha sonra bu kıyasların arasında bir seçim yaparken de en iyi sonuca varabilmek için hep diğer algoritmalarla karşılaştırma yapılmıştır.

Kaynakça

Wang, Chunyan, Wentao Li, Dubravka Drabek, Nisreen M.A. Okba, Rien van Haperen, Albert D.M.E. Osterhaus, Frank J.M. van Kuppeveld, Bart L. Haagmans, Frank Grosveld, and Berend-Jan Bosch, "A human monoclonal 1 antibody blocking SARS-CoV-2 infection," bioRxiv, 2020.

Gormsen, N. J., & Koijen, R. S. (2020). Coronavirus: Impact on stock prices and growth expectations. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2020-22).

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

Burbidge, R., Trotter, M., Buxton, B., & Holden, S. (2001). Drug design by machine learning: support vector machines for pharmaceutical data analysis. *Computers & chemistry*, *26*(1), 5-14.

Song, K. B., Baek, Y. S., Hong, D. H., & Jang, G. (2005). Short-term load forecasting for the holidays using fuzzy linear regression method. *IEEE transactions on power systems*, *20*(1), 96-101.

Fuller, W.A. 1969. Grafted polynomials as approximating functions. Aust J Agric Econ 13:35-46.

Gallant, A. R. 1974. The theory of nonlinear regression as it relates to segmented polynomial regressions with estimated join points. Institute of Statistics Mimeograph Series No. 925, 25 p. Raleigh, N.C.

Jain, M., Narayan, S., Balaji, P., Bhowmick, A., & Muthu, R. K. (2020). Speech emotion recognition using support vector machine. *arXiv preprint arXiv:2002.07590*.

Qawqzeh, Y. K., Otoom, M. M., Al-Fayez, F., Almarashdeh, I., Alsmadi, M., & Jaradat, G. (2019). A Proposed Decision Tree Classifier for Atherosclerosis Prediction and Classification. IJCSNS, 19(12), 197.

Breiman, L. (1999). Random forests. *UC Berkeley TR567*.

Quinlan, J. R. (2014). *C4. 5: programs for machine learning*. Elsevier.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.

Yağanoğlu, M., & Köse, C. (2017). Wearable vibration based computer interaction and communication system for deaf. Applied Sciences, 7(12), 1296.