

Türkiye’de Covid-19 günlük vaka sayısının makine öğrenmesi algoritmaları ile tahmin edilmesi

Ertürk SÜTCÜ¹
Parvaneh SHAMS²

Geliş tarihi / Received: 14.06.2021

Düzeltilerek geliş tarihi / Received in revised form: 23.12.2021

Kabul tarihi / Accepted: 12.01.2022

DOI: 10.17932/IAU.ABMYOD.2006.005/abmyod_v16i63004

Öz

COVID-19 enfeksiyonu 2019 yılının Aralık ayında ortaya çıkmış olup, Dünya Sağlık Örgütü tarafından 11 Mart 2020’de pandemi olarak tanımlanmıştır. Pandemi sürecinde vaka sayılarının kontrol altına alınması için sosyal alanlarda kısıtlama ve sokağa çıkma yasağı gibi önlemler alınmıştır. Pandemi sürecinde vaka sayılarına yönelik tahminlerin yapılması önemlidir. Vaka sayısı tahmininin yapılmasında kullanılan zaman serisi analizi, bölme model ve makine öğrenmesi gibi kullanılan tahminleme yöntemleri bulunmaktadır. Bu araştırmada COVID-19 vaka sayısının tahmin edilmesinde Destek Vektör Makinesi Algoritması (SVM), Karar Ağacı Algoritması (DT), Naif Bayes Algoritması (NB), K-En Yakın Komşu Algoritması (KNN) ve Rastgele Orman Algoritması (RF) olarak 5 makine öğrenme algoritması üzerinde çalışılmıştır. Algoritmaların tahmin performansı doğruluk, duyarlılık ve kesinlik değerleri ile belirlenmiştir.

¹ İstanbul Aydın Üniversitesi, Lisansüstü Eğitim Enstitüsü, Bilgisayar Mühendisliği, erturksutcu@stu.aydin.edu.tr, ORCID NO:0000-0003-4647-9814

² İstanbul Aydın Üniversitesi, Bilgisayar Mühendisliği, parvanehshams@aydin.edu.tr, ORCID NO:0000-0003-1467-3284

PCA yöntemi kullanıldığında doğruluk değeri en yüksek RF algoritmasında, duyarlılık ve kesinlik değeri en yüksek SMV algoritmasında saptanmıştır. Bu yöntemin kullanılması sonucunda en düşük doğruluk NB algoritmasında, duyarlılık ve kesinlik değerleri en düşük NB ve DT algoritmalarında elde edilmiştir. PCA yöntemi kullanıldığında tüm algoritmalarda doğruluk, duyarlılık ve kesinlik değerleri AVM'ler Açık/Kapalı veri setinde en yüksek düzeyde bulunmuştur. Okullar Açık/Kapalı, Restoranlar Açık/Kapalı ve AVM'ler Açık/Kapalı veri setinde RF algoritması ile tahminde sırasıyla %95, %88 ve %90 başarı oranı bulunmuş iken; Sokağa Çıkmak Yasak verisi için DT Algoritması ile tahminde %85 başarı oranı elde edilmiştir.

Anahtar Kelimeler: *Makine öğrenmesi, Covid-19 vaka sayısı, Sınıflandırma algoritması, Rastgele orman algoritması*

Prediction Covid-19 cases per day with machine learning algorithms in Turkey

Abstract

COVID-19 infection emerged in December 2019 and was defined as a pandemic by the World Health Organization on March 11, 2020. During the pandemic process, measures such as restrictions in social areas and curfews have been taken. It is important to make estimates for the number of cases during the pandemic process. There are estimation methods used in estimating the number of cases, such as time series analysis, divisional model, and machine learning. In this research, 5 machines as Support Vector Machine Algorithm (SVM), Decision Tree Algorithm (DT), Naive Bayes Algorithm (NB), K-Nearest Neighbor Algorithm (KNN), and Random Forest Algorithm (RF) were used to estimate the number of COVID-19 cases. The prediction performance of the algorithms was determined by

the accuracy, sensitivity and precision values. When the PCA method is used, the RF algorithm has the highest accuracy and the SMV algorithm has the highest sensitivity and precision. As a result of using this method, the lowest accuracy was obtained in the NB algorithm, and the lowest sensitivity and precision values were obtained in the NB and DT algorithms. When the PCA method was used, the accuracy, sensitivity and precision values in all algorithms were found at the highest level in shopping centers on Opened/Closed dataset. In the Schools Opened/Closed, Restaurants Opened/Closed and Shopping Malls Opened/Closed data set success rates were found respectively 95%, 88%, and 90% in estimating with the RF algorithm. In the Lockdown data, 85% success rate was achieved in the estimation with the DT Algorithm.

Keywords: *Machine learning, Number of covid-19 cases, Classification algorithm, Random forest algorithm*

Giriş

Koronavirüs ailesinde önemli insan ve hayvan patojenleri bulunur. Koronavirüs hastalığı (COVID), ilk olarak Çin'in Wuhan kentinde sebebi bilinmeyen pnömoni vakaları bildirilmesi ile Dünya Sağlık Örgütü (DSÖ) tarafından 31 Aralık 2019'da rapor edilmiştir. Ardından tanımlanan vakaların 'yeni koronavirüs 2019' (2019-nCoV) adı verilen yeni bir virüse bağlı olduğu açıklanmıştır. 30 Ocak 2020'de DSÖ'nün bu hastalığı küresel sağlık aciline neden olan bir salgın olarak duyurmasının ardından, COVID-19 Mart 2020'de pandemi evresine ulaşmıştır (Gorbalenya ve ark., 2020). Yeni Koronavirüs 2019 (2019-nCoV) daha sonra klinik özelliklerini de tanımlar nitelikte "şiddetli akut solunum sendromu koronavirüsü 2" (SARSCoV-2) olarak isimlendirilmiştir (Salehi ve ark., 2020).

Moleküler teste dayalı olarak doğrulanmış SARS-CoV-2 enfeksiyonu

geçiren insan sayısı 30 Aralık 2020-1 Şubat 2021 arasında 100 milyon üzerinde olduğu belirtilmiştir. DSÖ (2021)'ya göre COVID-19 enfeksiyonu nedeni ile dünya çapında ölen insan sayısı yaklaşık olarak 1,8 milyondur (Zhou ve ark., 2020; WHO, 2021).

Bu salgın, insan fiziksel ve zihinsel sağlığı için büyük bir tehdit oluşturarak, küresel ölçekte psikososyal etkilerle günlük yaşamı dramatik bir şekilde etkilemiştir. Tüm ülkelerde virüsün yayılma hızı, nedenleri ve tehlikeleri değerlendirilerek tam yad kısmi zamanlı karantina uygulamaları yapılmıştır (Rustam ve ark., 2020).

Makine öğrenmesi algoritmalarının verimliliği ve doğruluğu kazanılan deneyime bağlıdır. Bu durum makine öğrenme algoritmaları geliştiren uzmanların gelecek ile ilgili daha düşük hatalı ve daha iyi sonuçlar veren tahminlerde bulunmalarını sağlamıştır. Literatürde makine öğrenmesi modellerinin kullanılmasıyla değişkenler arasındaki ilişkinin ortaya konulduğu ve bu ilişkinin performans değerleri ile somut bir hale geldiği birçok araştırma bulunmaktadır (McCoy ve ark., 2018; Papacharalampous ve Tyrallis, 2018).

COVID-19 vaka sayısı tahminlemede makine öğrenmesi modellerinin kullandığı çalışmalar bulunmaktadır. Rustam ve ark. (2020)'nın araştırmasında COVID-19'dan etkilenen hastaların sayısını tahmin etme kapasitesini göstermek için makine öğrenmesi algoritmaları kullanılmıştır. Tehdit edici faktörleri tahmin etmek için doğrusal regresyon (LR), en az mutlak büzülme ve seçim operatörü (LASSO), destek vektör makinesi (SVM) ve üstel yumuşatma (ES) gibi dört standart tahmin modeli kullanılmıştır. Sonuçlar, ES'nin kullanılan tüm modeller arasında en iyi performansı gösterdiğini, yeni teyit edilmiş vakaları, ölüm oranını ve iyileşme oranını tahmin etmede LR ve LASSO'nun iyi performans gösterdiğini ortaya koymuştur. Ayrıca SVM'nin mevcut veri seti verilen tüm tahmin senaryolarında kötü performans gösterdiği belirtilmiştir.

Avan ve Aslam (2020)'nin çalışmasında bulaşıcılıkları tahmin etmek için kullanılan tahmin modellerinden biri olan otomatik ARIMA, R paketi "tahmin" yoluyla dört büyük Avrupa ülkesinde 10 gün için doğrulanmış vaka sayısını tahmin etmek için kullanılmıştır. Çalışma sonucunda ARIMA sonuçlarının bildirilen vaka sayısı ile tutarlı olduğu belirtilmiştir.

Başka bir çalışmada ABD, Brezilya, Hindistan, Rusya ve İspanya gibi ilk beş ülkeden toplam enfekte vakalarının tahmininde ARIMA model özellikleri Hannan ve Rissanen algoritması kullanılarak tahmin edilmiştir. Tahmin verileri gerçek veriler ile karşılaştırıldığında tutarlılık elde edilmiştir (Sahai ve ark., 2020).

Sevli ve Gülsoy (2020)'in araştırmasında Dünya'daki Covid-19 vakaları ile ilgili güncel veri setinde Prophet modelinin kullanılmasıyla tahminleme yapılmıştır. Araştırmada elde edilen veriler ile gerçek vakaların karşılaştırılmasında sonuçların tutarlı olduğu belirtilmiştir.

Özen ve ark. (2021)'nin çalışmasında farklı makine öğrenme modelleri kullanılarak Amerika doğrulanmış vaka sayılarını kullanılmasıyla Amerika Birleşik Devletleri'nde ilerleyen günler için vaka tahminleri yapılmıştır. Araştırmada Doğrusal Regresyon, Random Forest, Polinom Regresyon, Prophet ve ARIMA modelleri kullanılarak R ve Python dilinde tahminleme yapılmıştır. Veri performanslarının değerlendirilmesinde ortalama karekök sapması (RMSE), ortalama mutlak yüzde ve ortalama mutlak hata (MAE) kullanılmıştır. Polinom Regresyon algoritması MAPE hata metriği değerlendirmesinde en iyi tahmin sonucunu vermiştir.

Bu araştırmanın amacı makine öğrenmesi yöntemleri ile COVID-19 hastalığının vaka sayısını tahmin eden seçilen algoritmalarından en iyi algoritmanın ispatlanarak günümüzde vaka sayısı tahminlerinde kullanılmasını sağlamaktır.

Metot

Veri seti

Araştırmada veri kaynağı olarak Sağlık Bakanlığı’nın paylaştığı veriler ele alınmıştır. Araştırma dataları 11.03.2020 – 20.03.2021 tarihleri arasındaki TURCOVID19 ‘dan alınan COVID-19 verilerini kapsamaktadır. Tablo 1’de veri seti metrik bilgisi bulunmaktadır. Datalar ülkemizin toplam ve günlük olmak üzere hasta, vefat, iyileşen, test, entübe, ağır hasta ve vaka; okulların, restoranların ve AVM’lerin açık-kapalı olma durumları ve sokağa çıkma yasağı ile ilgili bilgileri içermektedir.

Tablo 1. Veri seti metrik bilgisi

Metrik	Açıklaması
Toplam Hasta	Toplam Hasta Sayısı
Günlük Hasta	Günlük Yeni Hasta Sayısı
Toplam Vefat	Toplam Vefat Sayısı
Günlük Vefat	Günlük Yeni Vefat Sayısı
Toplam iyileşen	Toplam İyileşen Hasta Sayısı
Günlük iyileşen	Günlük İyileşen Yeni Hasta Sayısı
Toplam Test	Toplam Test
Günlük Test	Günlük Test Sayısı
Entübe	Toplam Entübede Olan Hasta Sayısı
Ağır Hasta	Toplam Ağır Hasta Sayısı
Günlük Vaka	Toplam Günlük Vaka Sayısı

Veri işleme

Normalizasyon işlemi- Araştırmalarda veri setlerinde verilerin bütünlüğünün sağlanması, veri tekrarının önlenmesi ve veri bütünlüğünün korunması ile performansının artırılması için normalizasyon yapılmaktadır. Bu araştırmada kullanılan normalizasyon işlemi gereklilikleri sağlamıştır. Araştırmada uygulanan normalizasyon denklemi $X_{yeni} = \frac{(X - X_{min})}{(X_{max} - X_{min})}$ şeklindedir (X_{yeni} : Hesaplanacak değer, X : Mevcut parametre, X_{min} :

Normalize edilmiş parametrenin en küçük değeri, X_{\min} : Normalize edilmiş parametrenin en büyük değeri).

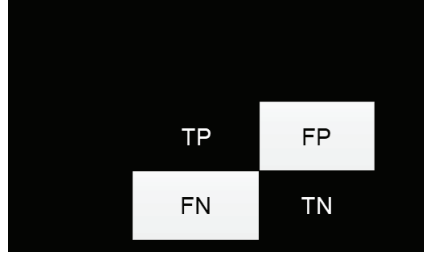
Çapraz doğrulama (CV)- Çapraz doğrulama, makine öğrenimi modellerinin başarı derecesini ortaya koymak için kullanılan yöntemdir. Çapraz doğrulama algoritma performansı hakkında bilgi verirken, verilerin daha verimli kullanılmasını sağlar. Yöntemin kullanılması için k değerinin belirlenmesi gerekir. Bu araştırmada k değeri 5 olarak alınmıştır (Şekil 1). COVID-19 vaka sayısı tahminin de %90 eğitim verisi ve %10 ise test verisi olarak kullanılmıştır.

Bütün Veri					
Eğitim seti					Test Set
Test	Eğitim	Eğitim	Eğitim	Eğitim	
Eğitim	Test	Eğitim	Eğitim	Eğitim	
Eğitim	Eğitim	Test	Eğitim	Eğitim	
Eğitim	Eğitim	Eğitim	Test	Eğitim	
Eğitim	Eğitim	Eğitim	Eğitim	Test	
Eğitim	Eğitim	Eğitim	Eğitim	Test	

Şekil 1. Kullanılan CV modeli

Karışıklık Matrisi (CM)-Bu yöntem, makine öğrenimin işleminden önce hedef veri kümelerinden elde edilen model performansının değerlendirilmesi için kullanılmaktadır. Şekil 2’de araştırmada kullanılan karışıklık matris model verilmiştir.

Gerçek Pozitif (TP): Test ve model verilerindeki değer sınıfı benzerdir. Sınıflandırma doğru yapılmıştır. Yanlış Negatif (FN): Test ve model verilerindeki değer sınıfı farklıdır. Sınıflandırma yanlış yapılmıştır. Yanlış Pozitif (FP): Gerçek değer negatif olmasına rağmen pozitif olarak değerlendirilmiştir. Yanlış sınıflandırma uygulanmıştır. Gerçek Negatif (TN): Gerçek ve sınıflandırılan değer negatif değerdedir. Sınıflandırma doğru yapılmıştır.



Şekil 2. Karışıklık matrisi modeli negatif

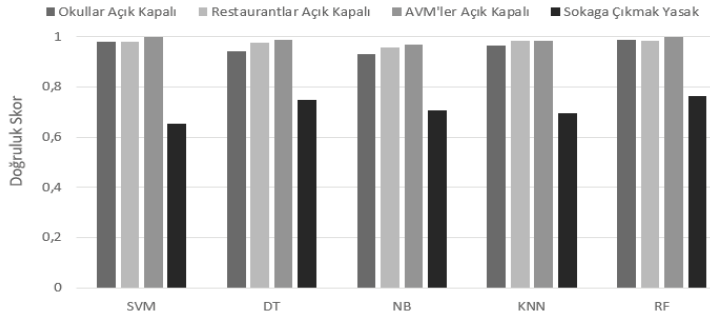
Temel bileşen analizi (PCA)- PCA, çok değişkenli bir veri kümesindeki az değişkenli ve minimum veri kaybına bilgilerin açıklanması için kullanılan matematiksel bir yöntemdir (Jin ve Bie, 2006). PCA, boyutsallığın azaltılmasını sağlar. Yöntem, boyut küçültme işlemindeki veri kümesindeki değişken sayısını azaltmayı hedeflemektedir. Dönüştürülme işleminden sonra oluşan değişkenler ilk değişkenlerin ana bileşenleri olarak adlandırılır. Varyans değeri en büyük olan değişken ilk bileşen olarak seçilir ve diğer temel bileşenler varyans değerlerini azaltmak için sıralanır.

Bulgular

PCA yöntemi kullanılarak veriler test edilmiştir. Elde edilen bulgular aşağıda verilmiştir. Araştırma başlangıcında 11 olan özellik sayısı PCA kullanılarak 5'e düşürülmüştür. Bu indirgemenin daha iyi sonuçlar ortaya koyduğu belirlenmiştir.

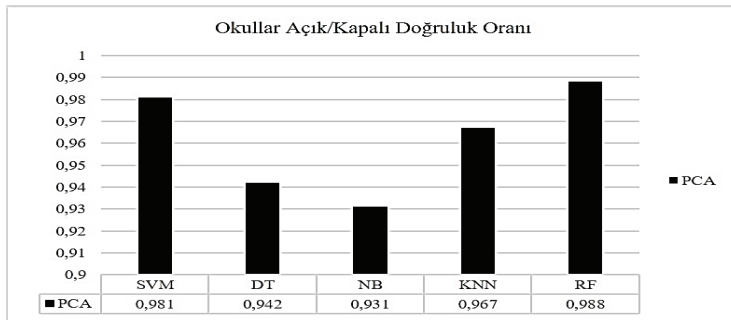
Şekil 3'te PCA yöntemi kullanılarak sırasıyla veri kümeleri Okullar Açık Kapalı, Restoranlar Açık Kapalı, AVM'ler Açık kapalı ve Sokağa Çakmak'a göre uygulamada kullanılan 5 algoritmanın doğruluk sonuç grafiği verilmektedir. Buna göre; Veri seti Okullar Açık Kapalı için en yüksek değeri veren algoritma 0.981 ile SVM iken, en düşük değer 0.931 ile NB algoritmasında bulunmuştur. Veri seti Restoranlar Açık Kapalı için en yüksek değeri veren algoritmalar 0.983 ile KNN ve RF iken, en düşük değer 0.958 ile NB algoritmasında bulunmuştur. Veri seti AVM'ler Açık

Kapalı için en yüksek değeri veren algoritmalar 0.999 ile RF ve SMV iken, en düşük değer 0.969 ile NB algoritmasında bulunmuştur. Veri seti Sokağa Çıkmak Yasak için en yüksek değeri veren algoritma 0.749 ile DT iken, en düşük değer 0.652 ile SMV algoritmasında bulunmuştur.



Şekil 3. PCA yöntemi kullanılan doğruluk sonuçları grafiği

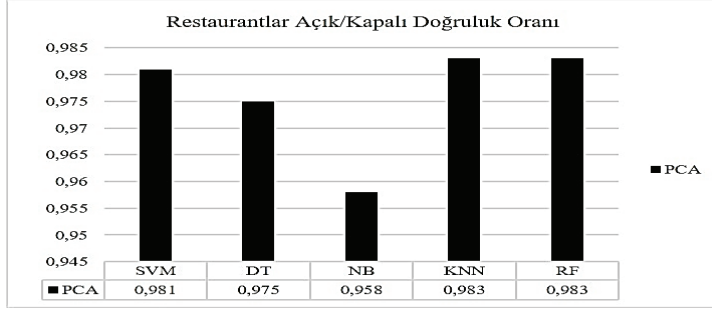
Şekil 4'te Okullar Açık Kapalı veri kümesi için PCA yönteminin kullanıldığı algoritmalarda elde edilen doğruluk sonuçları verilmektedir. Bu sonuçlara göre; PCA yönteminin kullanıldığı uygulamalarda en yüksek değer veren algoritmalar 0,988 ile RF iken, en düşük değer ise 0,931 ile NB algoritmasında bulunmuştur.



Şekil 4. Okullar açık/kapalı veri seti için doğruluk değerleri grafiği

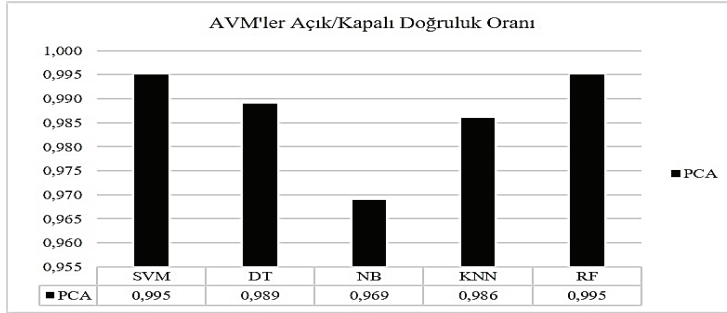
Şekil 5'te Restoranlar Açık Kapalı veri kümesi için PCA yönteminin algoritmalarda elde edilen doğruluk sonuçları verilmektedir. Bu sonuçlara

göre; PCA yönteminin kullanıldığı uygulamalarda en yüksek değer veren algoritma 0,983 ile RF ve KNN iken, en düşük değer ise 0,958 ile NB algoritmasında bulunmuştur.



Şekil 5. Restoranlar açık/kapalı veri seti için doğruluk değerleri grafiği

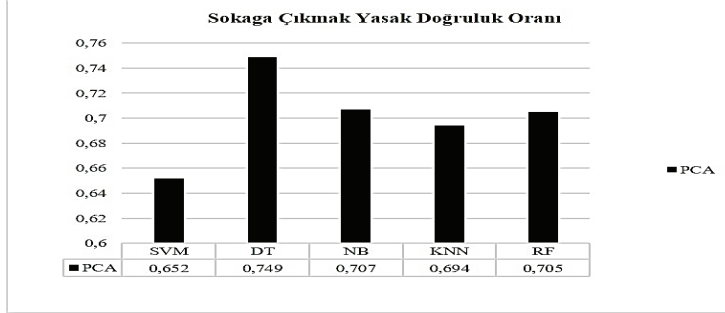
Şekil 6'da AVM'ler Açık Kapalı veri kümesi için PCA yönteminin kullanıldığı algoritmalarda elde edilen doğruluk sonuçları verilmektedir. Bu sonuçlara göre; PCA yönteminin kullanıldığı uygulamalarda en yüksek değer veren algoritma 0,995 ile SVM ve RF iken, en düşük değer ise 0,969 ile NB algoritmasında bulunmuştur.



Şekil 6. AVM'ler açık/kapalı veri seti için doğruluk değerleri grafiği

Şekil 7'de Sokağa Çıkmak Yasak veri kümesi için PCA yönteminin kullanıldığı algoritmalarda elde edilen doğruluk sonuçları verilmektedir. Bu sonuçlara göre; PCA yönteminin kullanıldığı uygulamalarda en yüksek

değer veren algoritma 0,749 ile DT iken, en düşük değer ise 0,652 ile SVM algoritmasında bulunmuştur.



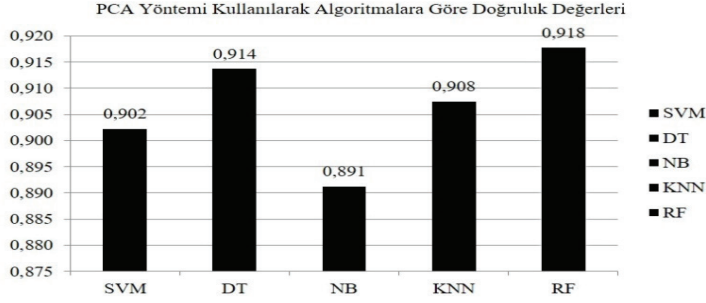
Şekil 7. Sokağa çıkmak yasak veri seti için doğruluk değerleri grafiği

Tablo 2’de çalışmada PCA yöntemi kullanılarak, algoritmalara göre ortalama değerler gösterilmektedir. Buna göre: Doğruluk değeri en yüksek RF algoritmasında, en düşük ise NB algoritmasında elde edilmiştir. Duyarlılık değeri en yüksek SVM algoritmasında, en düşük ise DT algoritmasında elde edilmiştir. Kesinlik değeri en yüksek SVM algoritmasında, en düşük ise DT algoritmasında elde edilmiştir.

Tablo 2. PCA yöntemi kullanılarak algoritmalara göre değerler

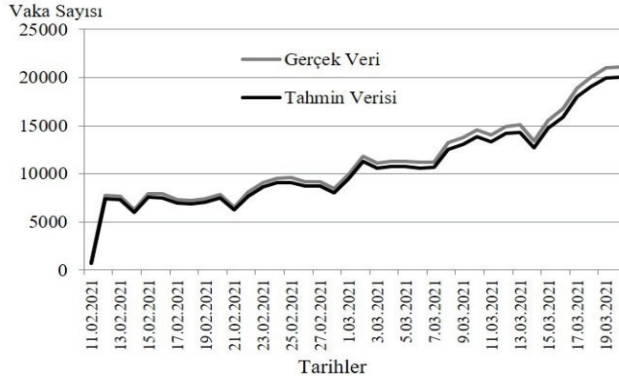
Veri Seti	Doğruluk	Duyarlılık	Kesinlik	TP	TN	FP	FN
SVM	0,902	0,767	0,695	18	47	4	3
DT	0,914	0,313	0,397	15	50	0	6
NB	0,891	0,503	0,450	15	48	2	5
KNN	0,908	0,639	0,615	16	48	2	4
RF	0,918	0,676	0,682	17	48	2	3

Şekil 8’de PCA yöntemi kullanılarak algoritmalara göre doğruluk değerleri gösterilmiştir.



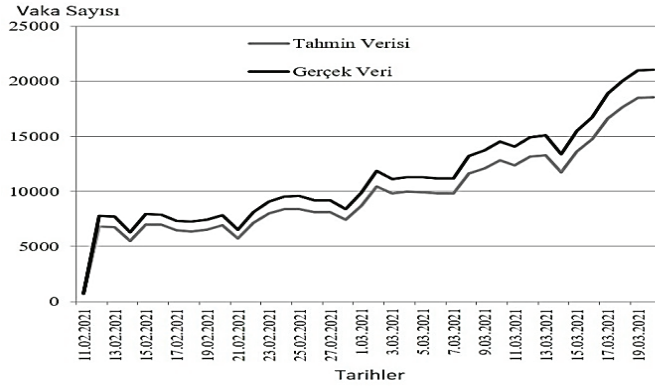
Şekil 8. PCA yöntemi kullanılarak algoritmalarla göre doğruluk değerleri grafiği

Şekil 9'da Okullar Açık / Kapalı verisi için RF Algoritması ile ilgili tahmin değerleri verilmiştir. Buna göre %90 eğitim – %10 test başarı oranı %95 olarak bulunmuştur.



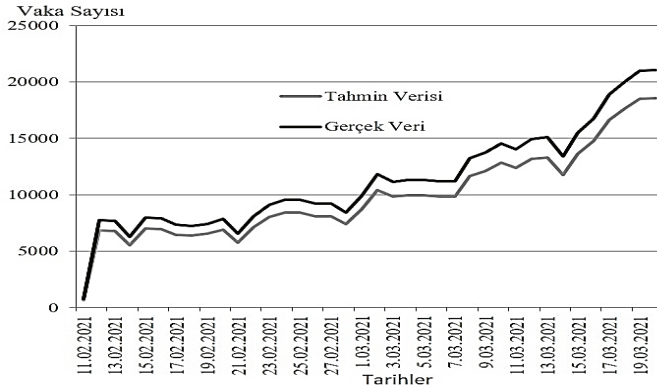
Şekil 9. Okullar açık / kapalı verisi için RF algoritması tahmin değerleri

Şekil 10'da Restoranlar açık/kapalı verisi için RF algoritması ile ilgili tahmin değerleri verilmiştir. Buna göre %90 eğitim – %10 test başarı oranı %88 bulunmuştur.



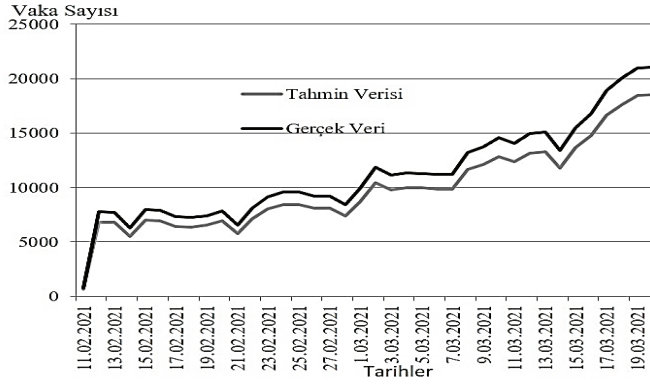
Şekil 10. Restoranlar açık/kapalı verisi için RF algoritması tahmin değerleri

Şekil 11’de AVM’ler Açık / Kapalı verisi için RF Algoritması ile ilgili tahmin değerleri verilmiştir. Buna göre %90 eğitim – %10 test başarı oranı %90 bulunmuştur.



Şekil 11. AVM’ler açık / kapalı verisi için RF algoritması tahmin değerleri

Şekil 12’de Sokağa Çıkmak Yasak verisi için DT Algoritması ile ilgili tahmin değerleri verilmiştir. Buna göre %90 eğitim – %10 test başarı oranı %85 bulunmuştur.



Şekil 12. Sokağa çıkmak yasak verisi için DT algoritması tahmin değerleri

Sonuç

COVID-19 salgınının güvencesizliği, büyük bir küresel krizi ateşleyebilir. Dünyanın dört bir yanındaki bazı araştırmacılar ve devlet kurumları, pandeminin dünya nüfusunun büyük bir bölümünü etkileyebileceğini düşünmektedir (Nash, 2020). Bu çalışmada, küresel olarak COVID-19 salgını vaka sayısını tahmin etmek için makine öğrenme tabanlı bir tahmin sistemi kullanılmıştır.

Araştırmaya göre PCA yöntemi kullanıldığında doğruluk değeri en yüksek RF algoritmasında, duyarlılık ve kesinlik değeri en yüksek SMV algoritmasında bulunmuştur. PCA yöntemi kullanılarak Okullar Açık Kapalı, AVM'ler Açık kapalı ve Sokağa Çıkmak Yasak veri setinde en yüksek doğruluk RF algoritmasında bulunmuştur. PCA yöntemi kullanılarak Okullar Açık Kapalı, Restoranlar Açık Kapalı ve AVM'ler Açık kapalı veri setinde en düşük doğruluk NB algoritmasında bulunmuştur. PCA yöntemi kullanıldığında tüm algoritmalarda doğruluk, duyarlılık ve kesinlik değerleri AVM'ler Açık/Kapalı veri setinde en yüksek düzeyde bulunmuştur.

RF algoritması ile tahminde Okullar Açık/Kapalı verisi için %95, Restoranlar Açık/Kapalı verisi için %88, AVM'ler Açık/Kapalı verisi için %90 başarı oranı bulunmuştur. Sokağa Çıkmak Yasak verisi için DT Algoritması ile tahminde %85 başarı oranı saptanmıştır. Bunun sonucunda Okullar Açık/Kapalı, Restoranlar Açık/Kapalı, AVM'ler Açık/Kapalı veri setinde RF algoritmasının; Sokağa Çıkmak Yasak veri setinde ise DT algoritmasının vaka sayısını tespit etmede başarılı olduğu belirlenmiştir.

Makine öğrenme algoritmalarıyla COVID-19 vaka tahmini çalışmalarında farklı sonuçlar bildirilmiştir. Bu farklılık kullanılan veri setleri ve algoritmaların farklı parametreler içermesinden kaynaklanmaktadır. Pandemi döneminde belirsizliklerin fazla ve sürecin aktif olması problem oluşturan durumlardandır. Halk sağlığı ve vakalar ile ilgili gerçeğe yakın tahminlerde bulunmak pandemi yönetimi süreci için önemlidir. Bu bağlamda çalışmada uygulana yöntem sonucunda elde edilen tahmin edilen veri ve gerçek verinin benzerlik oranının yüksek olmasının, mevcut algoritmanın ülkemizde vaka tahminin de kullanılabileceğini ortaya koymaktadır. Bu alanda yapılacak çalışmalar sonucunda farklı ülkelerde vaka sayısı tahmininde mevcut algoritmaların tekrarlanması ve doğruluk düzeyi en yüksek algoritmanın kullanılmasıyla pandemi yönetimi için stratejiler oluşturulabilir.

Kaynaklar

- [1] Awan, T. M., Aslam, F. (2020). Prediction of daily COVID-19 cases in European countries using automatic ARIMA model. *Journal of Public Health Research*, 9(3).
- [2] Gorbalenya, A.E., Baker, S.C., Baric, R.S. (2020). *Covid-19&Ct*. Published online.
- [3] Jin, X., Bie, R., (2006). Random Forest and PCA for Self-Organizing Maps based Automatic Music Genre Discrimination. *In DMIN*, 414-417.

- [4] McCoy, T. H., Pellegrini, A. M., Perlis, R. H., (2018). Assessment of time-series machine learning methods for forecasting hospital discharge volume. *JAMA network open*, 1(7), 87.
- [5] Nash, C. M., (2020). Harvard Professor Sounds Alarm on 'Likely' Coronavirus Pandemic: 40% to 70% of World Could Be Infected This Year, *JAMA*, 1(5).
- [6] Özen, N. S., Saraç, S., Koyuncu, M. (2021). COVID-19 Vakalarının Makine Öğrenmesi Algoritmaları ile Tahmini: Amerika Birleşik Devletleri Örneği. *Avrupa Bilim ve Teknoloji Dergisi*, (22), 134-139.
- [7] Papacharalampous, G. A., Tyrallis, H., (2018). Evaluation of random forests and Prophet for Daily stream flow forecasting. *Advances in Geosciences*, 45, 201-208.
- [8] Rustam, F., Reshi, A. A., Mehmood, A., Ullah, S., On, B. W., Aslam, W., Choi, G. S. (2020). COVID-19 future forecasting using supervised machine learning models. *IEEE access*, 8, 101489-101499.
- [9] Sahai, A. K., Rath, N., Sood, V., Singh, M. P. (2020). ARIMA modelling & forecasting of COVID-19 in top five affected countries. *Diabetes&Metabolic Syndrome: Clinical Research & Reviews*, 14(5), 1419-1427.
- [10] Salehi, S., Abedi, A., Balakrishnan, S., Gholamrezanezhad, A. (2020). Coronavirus disease 2019 (COVID-19): a systematic review of imaging findings in 919 patients. *American Journal of Roentgenology*, 215(1), 87-93.
- [11] Sevli, O., Gülsoy, V. G. B., (2020). Covid-19 Salgınına Yönelik Zaman Serisi Verileri ile Prophet Model Kullanarak Makine Öğrenmesi Temelli Vaka Tahminlemesi. *Avrupa Bilim ve Teknoloji Dergisi*, (19), 827-835.

[12] World Health Organization. Coronavirus Disease (COVID-2019) (2021). Situation Reports, *World Health Organization*, 2021.

[13] Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Cao, B., (2020). Clinical course and risk factors for mortality of adult in Patients with COVID-19 in Wuhan, China: a retrospective cohort study. *The lancet*, 395(10229), 1054-1062.