

Covid-19 Aşılama Sürecinin Tahminine Yönelik Derin Öğrenme Tabanlı Bir Model

Anıl Utku*, Ümit Can

Munzur Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Tunceli, Türkiye

*anilutku@munzur.edu.tr , ucan@munzur.edu.tr 

Makale gönderme tarihi:03.06.2022, Makale kabul tarihi:15.08.2022

Öz

Covid-19 pandemisi, insanlığın son zamanlarda karşılaştığı en büyük zorluklardan biridir. Henüz tedavi edici bir ilaç geliştirilemediği için tüm dünyayı sosyal ve ekonomik anlamda olumsuz etkilemektedir. Covid-19'un etkilerini ve vücutta bıraktığı hasarı en aza indirmek için farklı aşı çalışmaları yapılmıştır. Dünya genelinde insanlar aşılana salgının seyri kontrol altına alınmaya çalışılmaktadır. Bu noktada kullanılacak günlük aşı miktarının belirlenmesi, ihtiyaç duyulacak aşı ve enjektör gibi malzemelerin miktarına ve bunlarla beraber sağlık hizmetlerinin planlanmasına kadar önemli birçok alanda belirleyici olacaktır. Bununla birlikte birçok araştırmacı, virüs yayılım modeli oluşturmak ve Covid-19'un gidişatını tahmin etmek için farklı tahmin yöntemleri önermiştir. Bunlar arasında yapay zekâya dayalı yöntemler en ilgi çekici ve yaygın olarak kullanılan yöntemlerdir. Bu çalışmada, dünyada en yüksek aşılama oranına sahip ilk 20 ülke için günlük yapılan aşı sayılarının tahmin edilmesi amaçlanmıştır. Bu amaçla DT, kNN, LR, RF, SVM, MLP, CNN, RNN ve geliştirilen LSTM tabanlı derin öğrenme modelinin karşılaştırmalı bir analizi sunulmuştur. Uygulanan modeller için RMSE, MAE ve R^2 metriklerine göre elde edilen deneysel sonuçlar karşılaştırmalı olarak analiz edilmiştir. Deneysel sonuçlar, geliştirilen LSTM tabanlı modelin uygulanan ülkelerin tamamına yakınında 0.90'ın üzerinde R^2 değerine sahip olduğunu göstermiştir.

Anahtar Kelimeler: Aşı, covid-19, derin öğrenme, LSTM, makine öğrenmesi

A Deep Learning Based Prediction Model for Predicting the Covid-19 Vaccination Process

Abstract

The COVID-19 pandemic is one of the biggest challenges humanity has faced lately. As a therapeutic drug has not yet been developed, it negatively affects the entire world in social and economic terms. Various vaccine studies have been conducted to minimize the impact of COVID-19 and its harm to the body. People around the world are trying to control the course of the epidemic by vaccinating them. Determining the daily amount of vaccines to be used at this point will be decisive in important points such as the number of materials such as vaccines and injectors that will be needed, as well as the planning of health services. However, many researchers have proposed different predictive methods to build a model for the spread of the virus and predict the course of COVID-19. Of these, artificial intelligence methods are the most attractive and widely used. This study, it has been aimed to predict the number of daily vaccinations for the top 20 countries with the highest vaccination rate in the world. In this regard, a comparative analysis of DT, kNN, LR, RF, SVM, MLP, CNN, RNN, and the LSTM-based deep learning model was presented. The experimental results obtained according to the RMSE, MAE, and R^2 metrics for the applied models have been analyzed comparatively. Experimental results showed that the developed LSTM-based model has an R^2 value of over 0.90 in almost all of the applied countries.

Keywords: Vaccine, covid-19, deep learning, LSTM, machine learning

GİRİŞ

Covid-19, şiddetli akut solunum sendromuna sebep olan SARS-CoV-2 virüsünün neden olduğu, küresel pandemi haline gelen bulaşıcı bir hastalıktır (Lalmuanawma vd., 2020). Covid-19 ilk olarak

Aralık 2019'da Çin'in Wuhan şehrinde ortaya çıkmış ve 17 Kasım 2019'da ilk vaka teyit edilmiştir (Cucinotta ve Vanelli, 2020) ve daha sonrasında milyonlarca insanın ölümüne sebep olmuştur

(Şengür, 2021). Ayrıca Covid-19 pandemisi, ülkeleri sınırlarını kapatmak ve sokağa çıkma yasağı almak gibi önlemler almaya zorlayarak dünya ekonomisi üzerinde yıkıcı bir etkiye sahip olmuştur. Dünya Sağlık Örgütü (DSÖ), Covid-19'un ortaya çıktığı ilk tarihten 28 Nisan 2022 tarihine kadar dünya genelinde 509.531.232 doğrulanmış vaka, 6.230.357 Covid-19 kaynaklı ölüm ve yaklaşık 12 milyon doz aşı yapıldığını belirtmiştir.

DSÖ, Covid-19 bulaşmanın ilk aşamalarında pandemi olarak ilan ederek Covid-19'un çok şiddetli ve ölümcül bir hastalık olduğunu belirtmiştir (Jamal vd., 2021). Covid-19'un insan sağlığını önemli ölçüde etkilediği, doğrudan veya dolaylı olarak önceden var olan sağlık problemlerini tetikleyerek ölümlere neden olduğu belirtilmektedir. Ortaya çıkan bu virüsün başlıca semptomları yüksek ateş, sürekli ve inatçı öksürük, tat ve koku kaybı, yorgunluk, kas ağrısı ve nefes almada zorluktur. Bilim insanları, virüsün yayılmasını kontrol etmenin yollarını geliştirmek için virüsün morfolojik yapısını anlamaya çalışmaktadır (Zhang ve Yan, 2020). Bununla birlikte, devam eden birçok klinik çalışmanın Covid-19 vakalarının tedavisinde potansiyel olarak etkili olacağı düşünülmektedir.

Covid-19'a karşı geliştirilen aşilar, kişilerin hastalığa yakalanmadan virüse karşı bağışıklık geliştirmelerine yardımcı olmaktadır (Rubin, 2021). Farklı aşı türleri, koruma sağlamak için farklı şekillerde çalışır. Ancak her türlü aşı ile vücudun T-lenfositleri ve B-lenfositleri üretmesi amaçlanmaktadır (Royer ve Reinherz, 1987). Geliştirilen aşilar mRNA, protein alt birim aşiları ve vektör aşiları olmak üzere kategorize edilebilir (Samaranayake vd., 2021). mRNA aşiları, Pfizer-BioNTech ve Moderna tarafından geliştirilen, hücrelere virüse özgü zararsız bir proteinin nasıl üretileceğine ilişkin talimatlar veren virüs materyallerini içermektedir (Bisgin vd., 2021). Hücreler, proteinin kopyalarını elde ettikten sonra, aşidaki genetik materyali yok eder. Vücut, proteinin orada olmaması gerektiğini kabul eder ve gelecekte enfekte olursa Covid-19'a neden olan virüsle nasıl savaşılacağını hatırlayacak T-lenfositleri ve B-lenfositleri üretir.

Protein alt birim aşiları, virüsün zararsız parçalarını (proteinleri) içerir. Bir kez aşilanan vücut, proteinin orada olmaması gerektiğini anlar ve gelecekte enfekte olduğu zaman Covid-19'a neden

olan virüsle nasıl savaşılacağını hatırlayacak olan T-lenfositleri ve antikorları oluşturur (Wang vd., 2020).

Vektör aşiları, Johnson & Johnson's Janssen tarafından geliştirilmekte ve virüsün değiştirilmiş bir versiyonunu içermektedir. Modifiye edilmiş virüsün kabuğunun içinde, Covid-19'a neden olan virüsten bir materyal olan viral vektör bulunmaktadır. Viral vektör, hücrelerin içine girdiğinde genetik materyal, hücrelere Covid-19'a neden olan virüse özgü bir protein yapma talimatı verir. Bu talimatları kullanarak hücreler proteinin kopyalarını oluşturur (Gupta vd., 2021).

Dünya genelinde Covid-19 pandemisinden çok sayıda insan etkilendiği ve hastalığın tedavisi bulunmadığı için mevcut veriler kullanılarak virüs yayılımının tahmin edilmesi önem kazanmaktadır. Literatürdeki çalışmalar, hastalığın tahmin edilmesi için karmaşık algoritmalar ve modellerin geliştirilmesine odaklanmaktadır. Araştırmacılar, virüsün olası davranışlarını belirlemeye ve virüsün yayılımının doğru bir şekilde tahmin edilmesine yardımcı olabilecek tahmin modelleri geliştirmeye çalışmaktadır.

Arora vd., (2020) tarafından yapılan çalışmada, Hindistan'ın 32 eyaleti için pozitif Covid-19 vakalarının sayısını tahmin etmek için LSTM varyantlarının karşılaştırmalı bir analizi sunulmuştur. Günlük ve haftalık vakaları tahmin etmek için minimum hataya sahip LSTM modeli seçilmiştir. Önerilen yöntemin, günlük tahminler için %3'ten az ve haftalık tahminler için %8'den az hata ile kısa vadeli tahmin için yüksek doğruluk sağladığı gözlemlenmiştir.

Alazab vd., (2020) tarafından yapılan çalışmada, Covid-19 vaka sayısını tahmin etmek için CNN tabanlı bir derin öğrenme modeli sunulmuştur. Önerilen model ile hastaları tespit edebilmek için 1000 adet göğüs röntgeni görüntüsü analiz edilmiştir. Deneysel sonuçlar, önerilen sistemin Covid-19 tespitinde %95.99 F-skor değerine sahip olduğunu göstermiştir. Ayrıca, günlük vaka, ölüm ve iyileşme sayısını tahmin etmek için Prophet, AutoRegressive Integrated Moving Average (ARIMA) ve LSTM kullanılmıştır. Deneysel sonuçlar, Avustralya ve Ürdün için ortalama doğruluk oranlarının %94.80 ve %88.43 olduğunu göstermiştir.

Che Azemin vd., (2020) tarafından yapılan çalışmada, Covid-19 hastalarına ait radyolojik görüntüler kullanılarak derin öğrenme tabanlı bir hastalık tespit sistemi geliştirilmiştir. Çalışmada, bir

Research article/Araştırma makalesi
 DOI:10.29132/ijpas.1125729

milyon göğüs röntgeni görüntüsü kullanılarak anomalileri tespit etmek için ResNet-101 konvolüsyonel sinir ağı mimarisine dayalı bir derin öğrenme modeli kullanılmıştır. Deneysel çalışmalar, geliştirilen modelin %77.3 hassasiyet, %71.8 duyarlılık ve %71.9 doğruluk değerine sahip olduğunu göstermiştir.

Wang vd., (2020) tarafından yapılan çalışmada, Covid-19 salgınının epidemik trendini tahmin etmek için LSTM tabanlı bir model geliştirilmiştir. Çalışmada Rusta, Peru ve İran için salgının yükselen trendinin tahmin edilmesi ve sosyal izolasyon ile sokağa çıkma yasağı gibi önleyici tedbirlerin Covid-19'un yayılması üzerindeki etkinliği analiz edilmesi amaçlanmıştır. Ülkeler için salgının seyrinin azalacağı zamanlar belirlenerek vaka sayıları tahmin edilmiştir.

Pinter vd., (2020) tarafından yapılan çalışmada, Macaristan'daki Covid-19 vakalarını tahmin etmek için hibrit bir makine öğrenmesi yaklaşımı önerilmiştir. Adaptive Network-based Fuzzy Inference System (ANFIS) ve Multi-Layered Perceptron-Imperialist Competitive Algorithm (MLP-ICA) modelleri kullanılarak oluşturulmuş hibrit makine öğrenmesi yöntemi kullanılarak, vaka ve ölüm sayıları tahmin edilmiştir. Deneysel sonuçlar, önerilen modelin başarılı bir tahmin performansına sahip olduğunu göstermiştir.

Younis (2021) tarafından yapılan çalışmada, VGG, LetNet-5, AlexNet ve Resnet-50 mimarileri kullanılarak akciğer röntgeni görüntüleri üzerinde Covid-19 ve SARS_MERS virüslerinin tespit edilmesi amaçlanmıştır. Ayrıca LSTM modeli kullanılarak İtalya'daki 10 günlük Covid-19 vaka sayısının tahmini gerçekleştirilmiştir. Deneysel sonuçlar, VGG modelinin diğer modellere kıyasla %91 doğruluk oranına sahip olduğunu, LSTM'in ise Covid-19 vaka sayısını %99 doğrulukla tahmin ettiğini göstermiştir.

Abbasimehr ve Paki, (2021) tarafından yapılan çalışmada, Covid-19 vakalarının tahminine yönelik MultiHead Attention, LSTM ve CNN modellerini Bayes optimizasyon algoritması ile birlikte kullanarak hibrit tahmin modelleri geliştirilmiştir. Geliştirilen modeller, çoklu zaman noktalarının tahminine izin veren çoklu çıktı tahmin stratejisine dayalı olarak tasarlanmıştır. Bayes optimizasyon algoritması, her model için en iyi hiperparametreleri otomatik olarak seçmek ve tahmin performansını geliştirmek için kullanılmıştır. Geliştirilen hibrit

model ile 10 günlük kısa vadeli tahminler için 0.25 SMAPE değeri elde edilmiştir. Uzun vadeli tahminler için ise 2.59 SMAPE değeri elde edilmiştir.

Alassafi vd., (2022), tarafından yapılan çalışmada, Covid-19 salgınının Malezya, Fas ve Suudi Arabistan'daki yayılımının belirlenmesi için derin öğrenme tabanlı bir tahmin modeli geliştirilmiştir. Vaka ve ölüm sayılarını tahmin etmek için RNN ve LSTM modelleri kullanılmıştır. LSTM %98.58 hassasiyet değerine sahip olurken RNN %93.45 hassasiyet değerine sahip olmuştur.

Bu çalışmada, dünya genelinde en yüksek aşılama oranına sahip ilk 20 ülke için günlük yapılan aşı sayılarının tahmin edilmesi amaçlanmıştır. Çözölmeye çalışılan bu problem çok boyutlu ve karmaşık bir problemdir (Kızılloluk ve Can, 2021). Bu amaçla Decision Tree (DT) (Karar Ağacı), k-Nearest Neighbour (kNN) (k- En Yakın Komşu), Linear Regression (LR) (Lineer Regresyon), Random Forest (RF) (Rastgele Orman), Support Vector Machine (SVM) (Destek Vektör Makinesi), Multilayer Perceptron (MLP) (Çok Katmanlı Algılayıcılar), Convolutional Neural Network (CNN) (Evrışimli Sinir Ağları), Recurrent Neural Networks (RNN) (Tekrarlayan Sinir Ağları) ve geliştirilen Long Short-Term Memory (LSTM) (Uzun Kısa Süreli Bellek Ağları) tabanlı derin öğrenme modelinin karşılaştırmalı bir analizi sunulmuştur. Uygulanan modeller için RMSE, MAE ve R² metriklerine göre elde edilen deneysel sonuçlar karşılaştırmalı olarak analiz edilmiştir. Günlük yapılacak aşı sayısının tahmini sağlık çalışanlarının görev planlaması ve ekipman tedarigi için önemli bir konudur. Aşı ve enjektör gibi malzemelerin tedarigi ile sağlık personellerinin çalışma saatlerinin planlanmasında ve aşı yaptıracak kişilerin gruplar halinde aşılmasında stratejiler geliştirmek için günlük hayata uyarlanabilir. Bu çalışma, günlük yapılan aşı sayısının tahminine yönelik ilk çalışmadır. Literatürde yaygın olarak kullanılan popüler makine öğrenmesi ve derin öğrenme modellerinin uygulamalı bir analizi sunulmuştur. Bu çalışma vasıtasıyla Türkçe literatüre de katkıda bulunabilmek amaçlanmıştır.

TAHMİN MODELLERİ

Bu çalışmada günlük yapılan aşı sayısının tahminine yönelik popüler makine öğrenmesi ve derin öğrenme modelleri uygulamalı olarak karşılaştırılmıştır. DT, kNN, RF ve SVM makine

Research article/Araştırma makalesi
 DOI:10.29132/ijpas.1125729

öğrenmesi algoritmaları ile MLP, CNN, RNN ve LSTM derin öğrenme modelleri karşılaştırmalı olarak analiz edilmiştir.

DT, regresyon veya sınıflandırma problemlerinde kullanılan bir makine öğrenmesi algoritmasıdır (İbrahim ve Abdulazez, 2021). Bir veri kümesini giderek daha küçük alt kümelere ayırırken aynı zamanda ilgili ağacı aşamalı olarak geliştirilir. Nihai sonuç, karar düğümleri ve yaprak düğümleri olan bir ağaçtır. DT, kök düğüm, dallar ve yaprak düğümlerden oluşur. Kök düğüm, tüm örneği temsil eden ilk düğümdür ve daha fazla düğüme bölünebilir. İç düğümler bir veri kümesinin özelliklerini, dallar ise karar kurallarını temsil eder. Yaprak düğümler ise sonucu temsil eder. Bir karar düğümünün, her biri test edilen öznitelik için değerleri temsil eden iki veya daha fazla dalı vardır. Yaprak düğümü sayısal hedefle ilgili bir kararı temsil eder. DT hem kategorik hem de sayısal verileri işleyebilir (Song ve Ying, 2015).

kNN, tahmin edilecek değerlerin bağımsız değişkenlerinin oluşturduğu vektörün sınıfının, en yakın komşuların hangi sınıfta yoğun olduğu bilgisinden yola çıkarak tahmin edilmesine dayanır. kNN, uzaklık ve komşuluk sayısına (k) dayalı tahminler yapar (Franco-Lopez, 2001). Mesafe, tahmin edilecek noktanın diğer noktalara olan uzaklığı ile hesaplanır. Uzaklığı hesaplamak için Minkowski, Öklid veya Manhattan fonksiyonları kullanılır. Komşuluk sayısı üzerinden en yakın kaç komşunun hesaplanacağı belirlenir. k değeri sonucu doğrudan etkileyecektir. k değeri 1 ise, aşırı öğrenme olasılığı çok yüksek olacaktır. Çok büyükse, çok genel sonuçlar verecektir. Bu nedenle, k'nin optimum değerini tahmin etmek problemin ana konusudur. kNN, basit ve gürültülü eğitim verilerine karşı direnci nedeniyle en popüler makine öğrenmesi algoritmalarından biridir. Ancak aynı zamanda bir dezavantajı da vardır. Örneğin, mesafeleri hesaplarken tüm durumları sakladığından, büyük veriler için kullanıldığında çok fazla bellek alanına ihtiyaç duyar (Tomppo ve Halme, 2004).

LR, gözlemlenen verilere doğrusal bir denklem uydurarak iki değişken arasındaki ilişkiyi modellemeye çalışır. Regresyon, bağımsız değişkenlere dayalı bir hedef tahmin değeri modeller (Hazra ve Gogtay, 2016). Çoğunlukla değişkenler ve tahmin arasındaki ilişkiyi bulmak için kullanılır. Farklı regresyon modelleri, dikkate aldıkları bağımlı ve bağımsız değişkenler arasındaki ilişkinin türüne ve

kullanılan bağımsız değişkenlerin sayısına göre farklılık gösterir. LR, belirli bir bağımsız değişkene (x) dayalı olarak bir bağımlı değişken değerini (y) tahmin etme görevini gerçekleştirir. $y=ax+b$ şeklinde bir fonksiyon kullanarak, x (girdi) ve y (çıkıtı) arasında doğrusal bir ilişki bulur (Hou vd., 2019).

RF, topluluk öğrenme yöntemini kullanan bir denetimli öğrenme algoritmasıdır. Topluluk öğrenme yöntemi, daha doğru bir tahmin yapmak için birden fazla makine öğrenmesi algoritmasından gelen tahminleri birleştiren bir tekniktir (Rahman, 2019). Hem regresyon hem de sınıflandırma problemlerine uygulanabilmesi, hiperparametre tahmini olmadan bile iyi sonuçlar vermesi nedeniyle popüler makine öğrenme modellerinden biridir. RF, eğitim süresi boyunca birkaç karar ağacı oluşturarak ve tüm ağaçların tahmini olarak sınıfların ortalamasını çıkararak çalışır (Prasad vd., 2006). Ağaçlar, aralarında hiçbir etkileşim olmadan paralel olarak çalışmaktadır. RF, ağaçları büyüttükçe, modele ek rastgelelik ekler. Bir düğümü bölerken, en önemli özelliği aramak yerine, rastgele bir özellik alt kümesi arasında en iyi özelliği arar (Chan ve Paelinckx, 2008). Bu nedenle, bir düğümü bölmek için algoritma tarafından yalnızca özelliklerin rastgele bir alt kümesi dikkate alınır.

SVM, veri noktalarını belirgin bir şekilde sınıflandıran N boyutlu bir uzayda bir hiperdüzlem bulmayı amaçlayan bir denetimli öğrenme algoritmasıdır (Sen vd., 2020). İki veri noktası sınıfını ayırmak için seçilebilecek birçok olası hiperdüzlem vardır. Amaç, her iki sınıfın veri noktaları arasındaki maksimum mesafeye sahip bir düzlem bulmaktır. Marj mesafesini en üst düzeye çıkarmak, gelecekteki veri noktalarının daha güvenle sınıflandırılabilmesi için önemlidir. Destek vektörleri sadece gözlemin koordinatlarıdır. SVM, iki sınıfı (hiper düzlem/çizgi) en iyi ayıran bir sınırdır. SVM, bir düzlemde yer alan noktaları ayırmak için bir çizgi çizer (Shivanna ve Agrawal, 2020). Bu çizginin her iki sınıfın noktaları için maksimum mesafede olmasını amaçlar.

MLP, bir ileri beslemeli sinir ağı modelidir. Giriş katmanı, çıkış katmanı ve gizli katman olmak üzere üç tip katmandan oluşur (Ghritlahre ve Prasad, 2018). Giriş katmanı işlenecek giriş sinyalini alır. Tahmin ve sınıflandırma gibi gerekli görevler çıktı katmanı tarafından gerçekleştirilir. Giriş ve çıkış katmanı arasına yerleştirilen rastgele sayıda gizli katman, MLP'nin hesaplama birimidir. Bir MLP'deki ileri beslemeli ağa benzer şekilde, veriler girişten çıkış

Research article/Araştırma makalesi
 DOI:10.29132/ijpas.1125729

katmanına doğru akar. MLP'deki nöronlar, geri yayılım algoritması ile eğitilir (Liu vd., 2018).

CNN, düşük seviye özelliklerin uzamsal hiyerarşilerini otomatik ve uyarlanabilir bir şekilde öğrenmek için tasarlanmış bir derin öğrenme modelidir (Laroui vd., 2019). CNN, temel olarak evrişim, havuzlama ve tamamen bağlantılı katmanlardan oluşmaktadır. Evrişim ve havuzlama katmanları, öznelik çıkarma işlemini gerçekleştirirken, tamamen bağlantılı katmanlar, çıkarılan öznelikleri sınıflandırma işlemini gerçekleştirir. Evrişim katmanı, özel bir doğrusal işlem türü olan evrişim gibi bir dizi matematiksel işlemden oluşur (Bodapati ve Veeranjanyulu, 2019). Dijital görüntülerde, piksel değerleri iki boyutlu (2B) bir ızgarada saklanır. Bir özellik, görüntünün herhangi bir yerinde ortaya çıkabileceğinden, CNN görüntü işleme için oldukça verimlidir. Bir katman çıktısını bir sonraki katmana besledikçe, çıkarılan özellikler hiyerarşik ve aşamalı olarak daha karmaşık hale gelebilir. Geri yayılım ve gradyan azaltma olarak adlandırılan optimizasyon algoritmaları aracılığıyla parametreler optimize edilir (Zhang vd., 2020).

RNN, zaman serisi verileri veya dizi içeren veriler için uyarlanmış özel bir yapay sinir ağı türüdür. RNN çeşitli zaman indeksleri arasındaki bağımlılıkları tanımlayarak bağlamsal bilgileri çıkarmak için tasarlandığından, dizi analizi işlemini gerçekleştirmek için yaygın olarak kullanılan bir sinir ağı modelidir. RNN, birbirini izleyen çok sayıda tekrarlayan katmandan oluşur ve bu katmanlar, diziyi diğer dizilerle eşleştirmek için sırayla modellenir. RNN, diziden bağlamsal verileri yakalamak için güçlü bir yeteneğe sahiptir. Ancak, ağ yapısındaki bağlamsal bilgiler sabittir ve veri sınıflandırma sürecini başarmak için etkin bir şekilde kullanılır (Liu ve Guo, 2019). RNN'in işlem döngüsünde, gizli katmandan gelen sonuç hem çıktı üretir hem de içerik birimlerine yazılır. Bu şekilde her yeni girdi, önceki girdilerin işlenmesi sonucunda üretilen içerik birimleriyle birlikte işlenir. Farklı zamanlarda depolanan veriler arasında bir ilişki varsa buna uzun süreli bağımlılık denir. RNN, bu uzun vadeli bağımlılıklar arasındaki ilişkiyi hesaplayabilen bir ağıdır (Liu vd., 2020).

LSTM, gelişmiş bir tekrarlı sinir ağı modelidir. LSTM, RNN'de yaşanan kaybolan gradyan problemini çözmek için kullanılabilir. LSTM, hücre, giriş kapısı, çıkış kapısı ve unutma kapısı birimlerinden oluşur (Nurcahyanto vd., 2022).

Hücreler belirli zaman aralıklarında değerlerin hatırlanmasını, kapılar ise hücreye giren ve çıkan bilgi akışını düzenler. LSTM, önemli olaylar arasındaki bilinmeyen boyut ve süredeki zaman gecikmeleri göz önüne alındığında, zaman serilerini sınıflandırmak, işlemek ve tahmin etmek için çok uygundur. LSTM bellek bloğu, ağırlıklı girdiler, aktivasyon fonksiyonları, önceki bloklardan girdiler ve nihai çıktılar gibi çeşitli bileşenlere sahip karmaşık bir birimdir (Staudemeyer ve Morris, 2019).

GELİŞTİRİLEN DERİN ÖĞRENME TABANLI TAHMİN MODELİ

Bu çalışmada DT, kNN, LR, RF, SVM, MLP, CNN, RNN ve geliştirilen LSTM tabanlı derin öğrenme modelinin karşılaştırmalı bir analizi sunulmaktadır. Modeller, aşılama oranı en yüksek olan ilk 20 ülke için uygulanmıştır. Her ülke için günlük aşılama sayılarına göre hata metrikleri kullanılarak deneysel çalışmalar yapılmıştır. Her bir model için RMSE, MAE ve R2 metriklerine göre elde edilen sonuçlar karşılaştırmalı olarak analiz edilmiştir.

Veri seti

Bu çalışmada, Our World in Data tarafından Kaggle üzerinden açık erişime sunulan, ülkelerin aşılama verileri kullanılmıştır (Daily and Total Vaccination for COVID-19 in the World from Our World in Data, 2021). Verisetinde Country, Country ISO code, Date, Total number of vaccinations, Total number of people vaccinated, Total number of people fully vaccinated, Daily vaccinations, Total vaccinations per hundred, Total number of people fully vaccinated per hundred, Number of vaccinations per day, Daily vaccinations per million ve Source name öznelikleri bulunmaktadır. Verisetinde bulunan özneliklerin açıklamaları Tablo 1'de verilmiştir

Tablo 1. Verisetindeki özneliklerin açıklamaları

Country	Aşı bilgilerinin sağlandığı ülke
Country ISO code	Ülke için ISO kodunu
Date	Veri girişi yapılan tarih
Total number of vaccinations	Ülkede yapılan aşılama toplam sayısı
Total number of people vaccinated	Aşılanan toplam kişi sayısı
Total number of people fully vaccinated	Aşı şemasına göre tüm aşılama yaptırmış kişi sayısını
Daily vaccinations	Belirli bir tarihte yapılan aşı sayısı

Research article/Araştırma makalesi
DOI:10.29132/ijpas.1125729

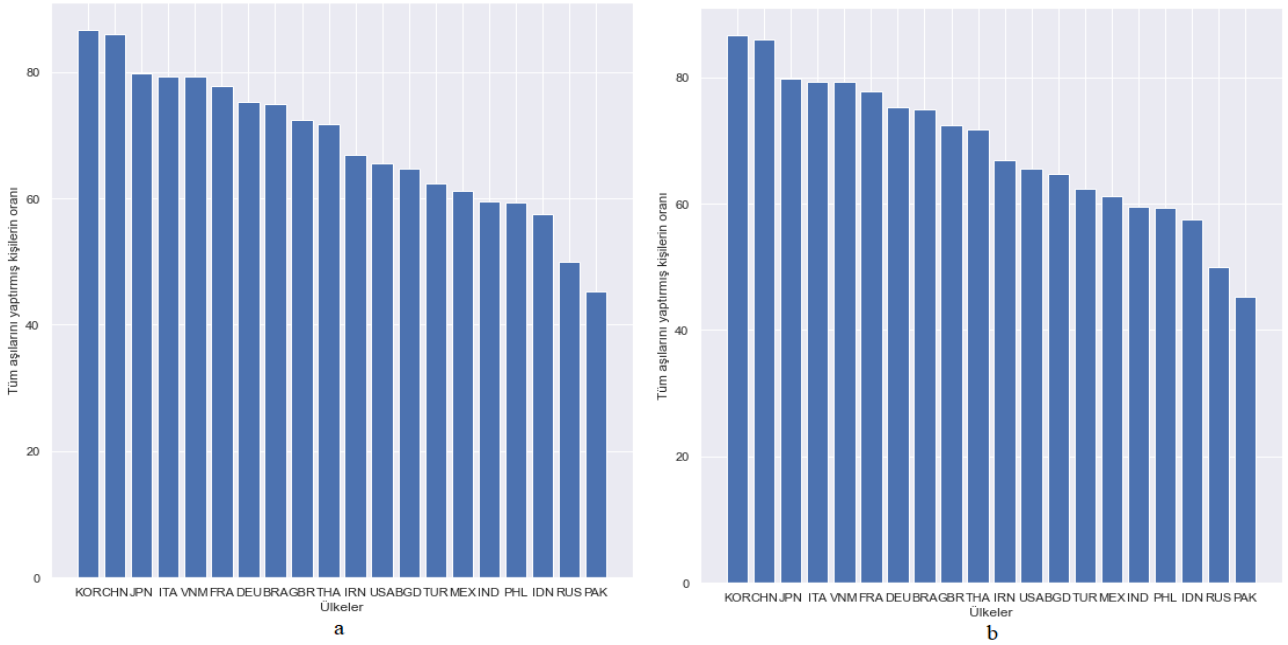
Total vaccinations per hundred	Belirli bir güne kadar ülkede yapılan aşı sayısı ile toplam nüfus arasındaki yüzdelik oranı
Total number of people fully vaccinated per hundred	Ülkede tüm aşılarını yaptırmış nüfus ile toplam nüfus arasındaki yüzdelik oranı
Number of vaccinations per day	Belirli bir gün ve ülke için günlük aşı sayısı
Daily vaccinations per million	Ülkedeki belirli bir tarih için aşı sayısı ile toplam nüfus arasındaki oranı
Source name	Verinin kaynağı

Kullanılan veriseti Mart 2022 yılına kadar ülkelerin aşılama verilerini içermektedir. Şekil 1'de uygulanan aşı dozları en yüksek olan ilk 20 ülke görülmektedir.

Şekil 1'de görüldüğü gibi Güney Kore, Çin İtalya, Fransa ve Vietnam uygulanan aşı dozu en yüksek olan ilk 5 ülkedir.

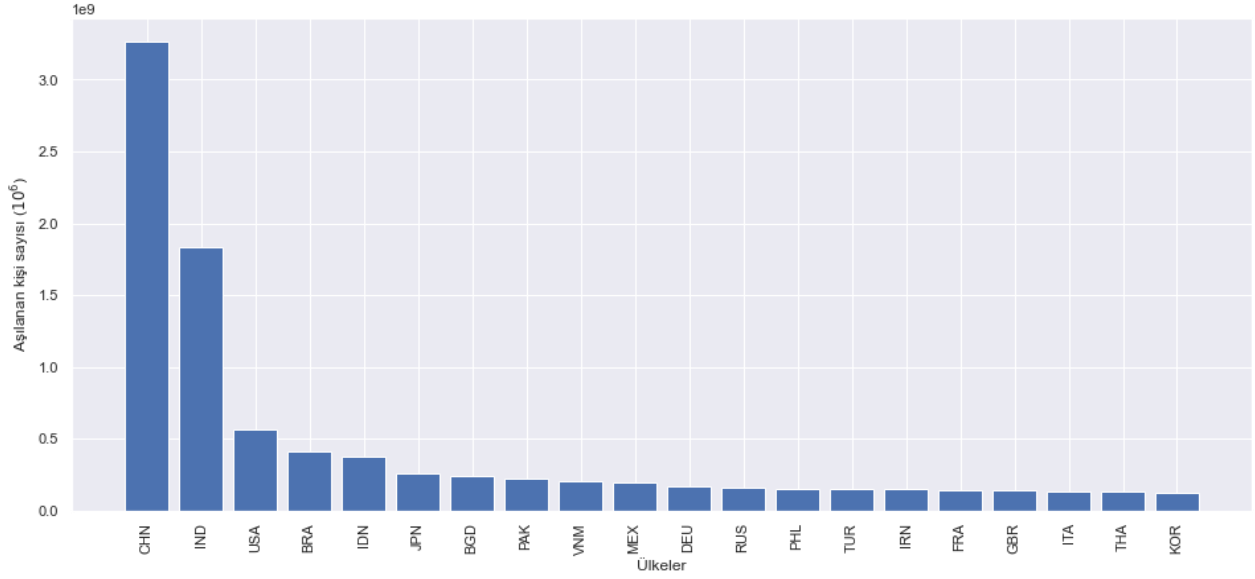
Şekil 2'de en fazla aşı yapılan ülkeler görülmektedir.

Şekil 2'de görüldüğü gibi Çin, Hindistan, ABD ve Brezilya gibi nüfus bakımından yoğun olan ülkeler, aşılama kişi sayısına göre önde gelmektedir.

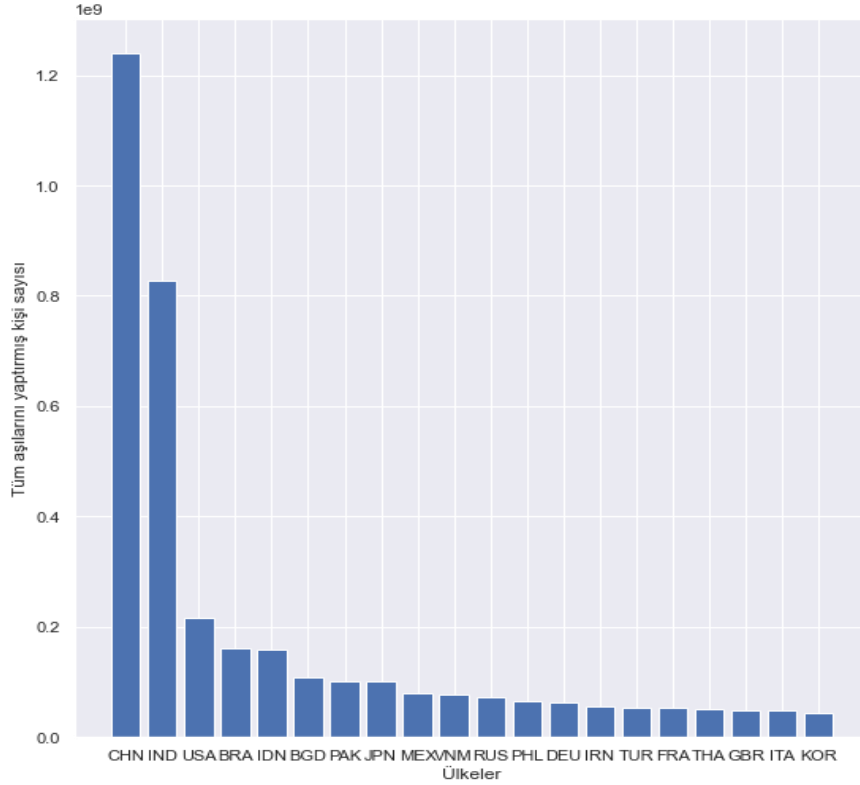


Şekil 1. Uygulanan aşı dozları ve tüm aşılarını yaptırmış kişilerin oranı en yüksek olan ilk 20 ülke

Research article/Araştırma makalesi
DOI:10.29132/ijpas.1125729



Şekil 2. En fazla aşı yapılan ülkeler



Şekil 3. Ülkelere göre tüm aşılarını yaptırmış kişi sayıları

Şekil 3'te ülkelere göre tüm aşılarını yaptırmış kişi sayıları görülmektedir.

Şekil 3'te görüldüğü gibi Çin, Hindistan, ABD, Brezilya ve Endonezya nüfus yoğunlukları itibarıyla tüm aşılarını yaptırmış kişi sayılarına göre öne çıkmaktadır.

Veri Ön-işleme

Günlük yapılan aşı sayısını tahmin edebilmek için öncelikle verisetindeki 'Daily vaccinations' sütunu seçilmiştir. Günlük yapılan aşı sütunundaki eksik ve hatalı olan alanlar kontrol edilmiştir. Bu

çalışmada kullanılan aşılama verileri, zaman serisi verileridir. Zaman serisi verileri, belirli bir zaman indeksine bağlı, kronolojik olarak sıralanmış verilerdir. Zaman serisi verileri kullanılarak tahmin modellerinin geliştirilebilmesi için bu verilerin denetimli öğrenme problemi yapısına dönüştürülmesi gereklidir. Denetimli öğrenme problemleri, girdi değişkenlerinden çıktı değişkenlerinin nasıl tahmin edileceğinin belirlenmesini sağlayan problemlerdir. Girdi değişkeni olarak düzenlenen zaman indekslerinden bir sonraki zaman indeksindeki çıktı değişkeninin belirlenmesi yoluyla zaman serisi verileri denetimli öğrenme problemine dönüştürülebilir.

Zaman serisi verileri kayan pencere kullanılarak denetimli öğrenme problemine dönüştürülebilir. Kayan pencere yöntemi, pencerenin içine alınan zaman indekslerindeki değerler kullanılarak gelecekteki bir değer tahmin edilmesi temeline dayanmaktadır. Kayan pencerenin boyutu, önceki zaman adımlarının sayısına göre belirlenir. Bu çalışmada, kayan pencerenin boyutu 3 olarak seçilerek t_1 , t_2 ve t_3 zaman adımlarındaki veriler girdi, t_4 zaman adımıdaki veri ise çıktı olacak şekilde yapılandırılmıştır.

Veriler denetimli öğrenme yapısına dönüştürüldükten sonra 0-1 aralığında normalize edilmiştir. Normalizasyonun amacı, verisetindeki sayısal sütunların değerlerini, değer aralığındaki farklılıkları bozmadan ortak bir ölçeğe değiştirmektir. Normalizasyon, geliştirilen modelin performansını ve eğitimin kararlılığını etkiler. Normalizasyondan sonra veriler %70 eğitim ve %30 test olarak ayrılmıştır. Eğitim verilerinin %10'u model doğrulaması için ayrılmıştır. Doğrulama verileri, model parametrelerinin optimizasyonu için kullanılmıştır. Karşılaştırılan modellerin en iyi sonuçları vermesi için GridSearchCV kütüphanesi kullanılarak parametre optimizasyonu yapılmıştır. GridSearch yönteminde, modelde test edilecek hiperparametreler ve değerleri için tüm kombinasyonlarla ayrı ayrı model kurulur ve belirlenen metriğe göre en başarılı hiperparametre seti belirlenir. Parametre optimizasyonu ile RMSE değerleri en düşük olan parametreler belirlenmiş ve modeller oluşturulmuştur. Modeller belirlenen hiperparametrelere göre oluşturulduktan sonra 10 kez çalıştırılarak elde edilen sonuçların tüm metrikler için ortalaması alınmıştır.

Geliştirilen LSTM Tabanlı Tahmin Modeli

LSTM, uzun vadeli bağımlılıkları hatırlamadaki başarısı nedeniyle diğer derin öğrenme modellerine göre öne çıkmaktadır. Bu çalışmada, önerilen LSTM tabanlı derin öğrenme modeli, DT, kNN, LR, RF, SVM, MLP, CNN ve RNN ile kapsamlı bir şekilde test edilmiştir. Geliştirilen LSTM tabanlı modelin mimarisi Şekil 4'te görülmektedir.

Bu çalışmada, LSTM modelini geliştirmek için ağ yapısı ve model parametrelerinin iyileştirilmesi amaçlanmıştır. Burada, uyarlanabilir optimize edicilerden biri olan Adam, tekrarlayan sinir ağlarındaki karmaşık eğitim dinamiklerini daha iyi ele almak için kullanılmıştır. Sonuçlar, LSTM katmanları eklenerek analiz edilmiştir. Ek olarak, gizli katman sayısı, eğitim sayısı, nöron sayısı ve yığın boyutu gibi hiperparametreler optimize edilerek LSTM'in en yüksek tahmin doğruluğunu elde etmesi sağlanmıştır.

Hata Metrikleri

Regresyon problemlerinde üretilen tahminlerin doğruluğunu ölçmek için temel olarak RMSE, MAE ve R^2 metrikleri kullanılmaktadır.

RMSE, hataların standart sapmasını ölçer. Veri kümesindeki gerçek değerler ile tahmin edilen değerler arasındaki farkın karelerinin ortalamasının karekökünü ifade eder. RMSE, Eşitlik 1. kullanılarak hesaplanır. Burada y gerçek değerler, \hat{y} tahmin edilen değerler ve n ise örnek sayısıdır.

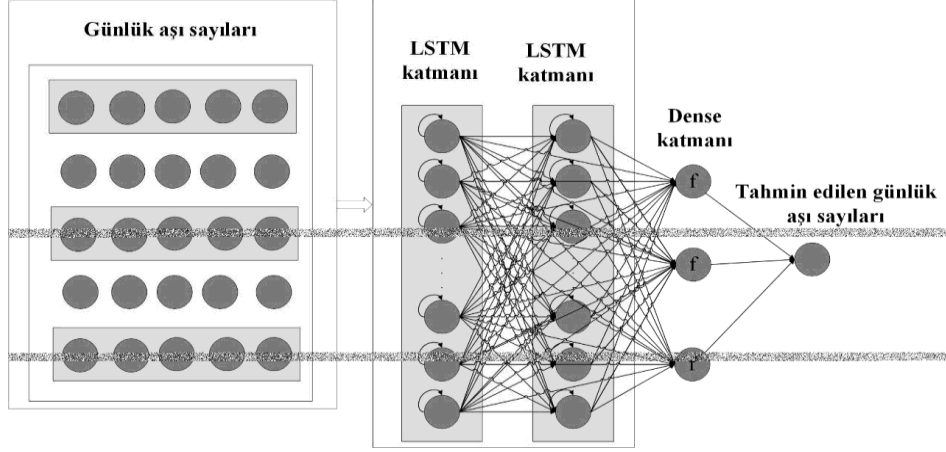
$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2} \quad (1)$$

MAE, veri kümesindeki gerçek değerler ile tahmin edilen değerler arasındaki farkların mutlak değerlerinin ortalamasını ifade eder. Veri kümesindeki hataların ortalamasını ölçer. MAE, Eşitlik 2. kullanılarak hesaplanır.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y - \hat{y}| \quad (2)$$

R^2 , modellerin verisetine uyumunun bir ölçüsüdür. R^2 , veri noktalarının regresyon çizgisi etrafındaki dağılımını değerlendirir. Aynı veriseti için, daha yüksek R^2 değerleri, gerçek değerler ile tahmin edilen değerler arasındaki farkın az olmasını ifade eder. R^2 , Eşitlik 3. kullanılarak hesaplanır.

$$R^2 = \frac{\sum(y-\hat{y})^2}{(y-\bar{y})^2} \quad (3)$$



Şekil 4. Geliştirilen LSTM tabanlı tahmin modeli

DENEYSEL SONUÇLAR

Uygulanan aşı dozları en yüksek olan ilk 20 ülke için DT, kNN, LR, RF, SVM, MLP, CNN, RNN ve geliştirilen LSTM tabanlı derin öğrenme modelinin sonuçları RMSE, MAE ve R^2 metriklerine göre karşılaştırmalı olarak analiz edilmiştir.

RMSE metriğine göre deneysel sonuçlar Tablo 1'de görülmektedir.

Tablo 2'de görüldüğü gibi geliştirilen LSTM tabanlı model, RMSE metriğine göre karşılaştırılan diğer modellere göre daha iyi bir tahmin performansına sahiptir. LSTM'den sonra, RNN, MLP, LR ve SVM karşılaştırılan diğer modellere göre daha başarılı sonuçlara sahiptir.

Tablo 3'te, MAE metriğine göre deneysel sonuçlar görülmektedir.

Tablo 2'de görüldüğü gibi geliştirilen LSTM tabanlı model, MAE metriğine göre karşılaştırılan diğer modellere göre daha iyi bir tahmin performansına sahiptir. LSTM'den sonra RNN, MLP, LR ve SVM karşılaştırılan diğer modellere göre daha başarılı sonuçlara sahiptir.

Şekil 5'te görüldüğü gibi geliştirilen LSTM tabanlı model, R^2 metriğine göre karşılaştırılan diğer modellere göre daha iyi bir tahmin performansına sahiptir. LSTM'den sonra, RNN, MLP, LR ve SVM karşılaştırılan diğer modellere göre daha başarılı sonuçlara sahiptir.

Deneysel sonuçlar geliştirilen LSTM tabanlı modelin karşılaştırılan diğer modellere göre daha başarılı bir tahmin performansına sahip olduğunu göstermiştir. LSTM'in bu çalışmadaki başarısı uzun vadeli gözlem dizilerini öğrenme yeteneği ile açıklanabilir. LSTM, karmaşık diziler ve büyük veri kümeleri üzerinde çalışıldığında parametrik modellerden ve standart tekrarlı sinir ağı mimarilerinden daha iyi sonuçlar elde etmektedir. LSTM, geleneksel doğrusal zaman serisi verilerini genelleştirebilen ve çalışmak için durağan veriler gerektirmeyen bir yapıya sahiptir. Ancak RNN gibi modeller durağanlığı sağlamak için seçilen aktivasyon fonksiyonunun doğrusal olmamasını gerektirir.

Tablo 2. RMSE metriğine göre deneysel sonuçlar

Kod	DT	kNN	LR	RF	SVM	MLP	CNN	RNN	LSTM
KOR	32716.06	35631.06	18623.11	28803.43	20209.69	18505.36	25439.16	19263.49	18485.28
CHN	450080.02	503102.81	170654.77	324949.15	212437.13	462700.09	325629.52	270616.81	195612.17
ITA	24769.14	26391.40	14537.24	25191.80	15766.86	15005.34	26250.57	14540.88	14446.04
FRA	25984.85	33645.18	17178.33	26318.19	21185.53	21184.77	38952.74	19032.37	16247.14
VNM	84172.93	81269.64	50513.93	72303.36	50144.90	53189.31	73908.29	50012.82	49422.34

GBR	107506.29	86833.72	18156.09	89542.13	21089.03	22686.68	37388.45	16208.88	14607.07
DEU	67268.11	71289.79	22274.98	67283.19	26745.11	44195.72	52457.44	35127.59	20909.15
JPN	34564.14	32851.16	22212.79	25477.86	21351.57	28592.70	32080.30	24138.26	21700.59
BRA	104567.47	115347.15	85150.98	101022.49	162591.72	89995.36	93320.41	178001.10	85195.16
THA	49340.64	46850.47	26483.48	42571.53	28169.41	27557.58	41354.06	26627.13	25179.84
TUR	37005.50	35363.01	13952.58	27835.75	14650.58	12654.27	18335.50	13024.72	12304.40
IRN	45888.12	44967.91	17554.94	44593.27	23040.69	23272.66	25355.04	20108.02	17147.39
USA	66543.84	78935.33	41326.81	63692.66	45941.18	45855.59	71050.58	43769.42	28120.82
MEX	159379.62	152098.55	124891.97	127156.98	125241.09	125395.87	152759.30	123138.90	121500.48
BGD	509859.52	584496.25	163273.27	486471.08	158162.20	167281.30	214600.76	195180.77	156870.88
IDN	97047.42	95981.80	73281.79	82650.02	73867.32	73334.53	83414.04	73262.34	73099.47
PHL	123711.39	121844.83	55528.21	115484.04	57171.11	63111.17	94922.06	53736.16	50238.99
IND	623631.15	610306.82	554253.86	529799.70	569677.28	535742.57	514448.54	561915.63	508728.67
RUS	84295.82	85485.78	65057.34	70346.80	64994.38	66540.62	75341.53	65553.75	64866.79
PAK	247491.03	252911.98	57973.66	245667.79	57551.48	69503.16	86859.86	62024.58	57217.20

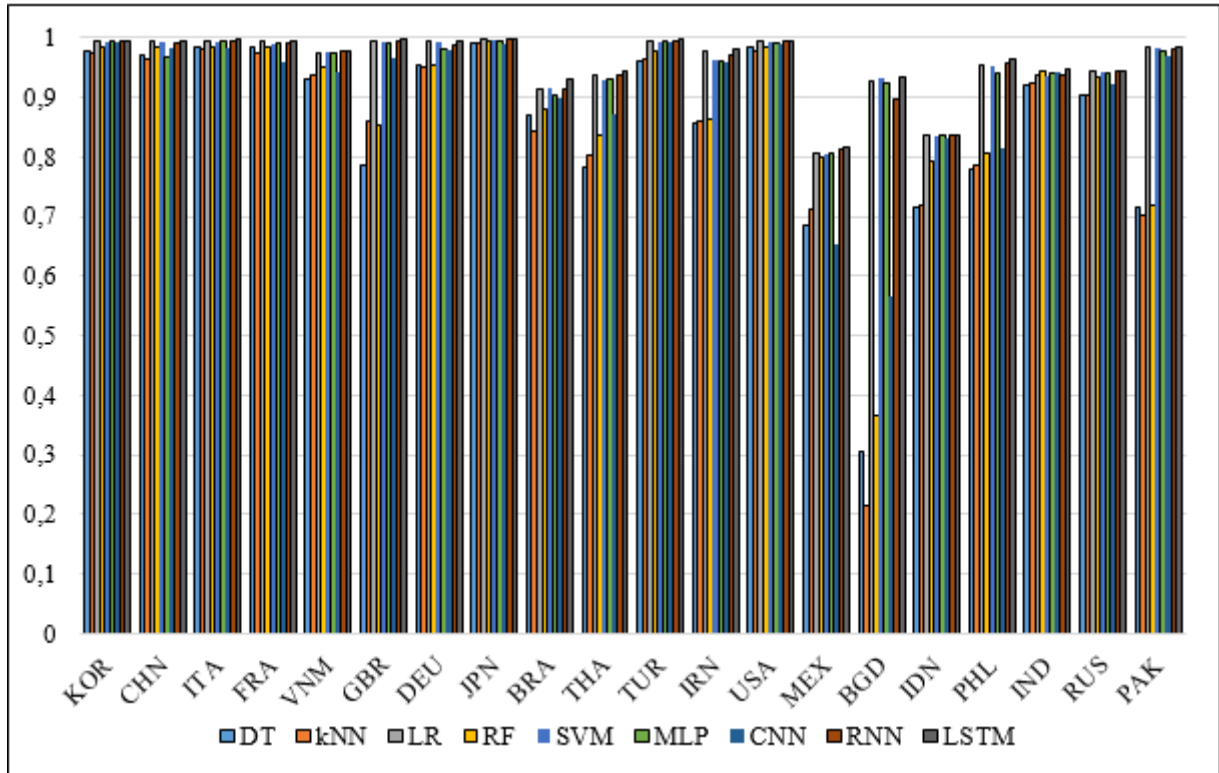
Tablo 3. MAE metriğine göre deneysel sonuçlar

Kod	DT	kNN	LR	RF	SVM	MLP	CNN	RNN	LSTM
KOR	23579.56	25827.37	11958.73	19858.37	15335.59	11105.30	15721.49	13948.36	11092.08
CHN	343043.61	388971.22	133312.66	234790.50	172925.04	319618.66	249176.99	201784.64	107526.44
ITA	16556.21	16853.95	6829.78	16854.92	8313.29	7254.60	18618.64	7032.90	6367.47
FRA	19113.30	22553.13	9116.49	19118.21	13435.99	11314.72	24683.58	12929.40	9029.13
VNM	64181.78	64922.54	37936.15	55881.56	38086.29	40431.99	58983.68	38752.46	37050.06
GBR	66541.66	49198.72	10235.49	50456.70	16170.27	10352.84	17485.12	11422.70	7445.50
DEU	38963.98	41995.77	14408.82	38191.54	22111.39	25941.69	31873.87	22687.94	12555.98
JPN	26966.72	26203.39	15614.56	19673.52	14833.56	19456.76	22737.11	16544.97	15488.15
BRA	78851.61	85085.14	90597.49	71725.93	56875.33	57768.31	61976.82	97515.39	56373.88
THA	35533.20	35364.95	14930.16	31575.67	18508.40	15113.96	24843.54	15087.12	14551.94
TUR	26522.38	24979.73	10314.23	19633.65	11092.75	7403.02	11312.76	6197.19	6107.36
IRN	37537.47	37182.01	13137.78	36349.47	19147.96	18136.93	19657.82	15824.17	12180.75
USA	46290.12	54836.69	20625.20	45971.43	26580.70	21574.68	38865.73	22328.01	17225.55
MEX	67385.68	64382.56	47791.49	50574.70	48056.13	47748.41	70425.18	46744.54	45974.23
BGD	252355.74	321327.06	94209.95	279065.03	85822.08	98555.78	143700.57	137393.55	84436.40
IDN	79386.72	75963.67	49081.34	57279.27	50083.40	48992.88	61213.51	50241.59	48865.01
PHL	69793.40	67008.25	33294.72	58285.73	35294.57	36717.05	54477.95	34110.93	32966.75
IND	430140.15	461875.30	388758.21	388929.15	404829.82	376615.94	371275.09	388043.45	364736.02
RUS	59106.03	58806.66	44625.14	48371.04	44242.26	46180.47	51481.88	44359.37	43963.45
PAK	129612.87	141745.53	43563.15	123780.21	44363.40	48769.28	66505.71	47077.08	43331.93

Tablo 4. R² metriğine göre deneysel sonuçlar

Kod	DT	kNN	LR	RF	SVM	MLP	CNN	RNN	LSTM
KOR	0.978	0.974	0.993	0.983	0.991	0.993	0.991	0.993	0.994
CHN	0.971	0.964	0.994	0.985	0.993	0.969	0.981	0.990	0.995
ITA	0.983	0.981	0.994	0.983	0.993	0.993	0.981	0.994	0.996
FRA	0.985	0.975	0.994	0.985	0.990	0.990	0.960	0.992	0.995
VNM	0.932	0.937	0.975	0.950	0.976	0.973	0.943	0.976	0.978
GBR	0.787	0.861	0.993	0.852	0.991	0.990	0.965	0.995	0.996

DEU	0.955	0.949	0.994	0.955	0.992	0.980	0.978	0.987	0.995
JPN	0.991	0.992	0.996	0.995	0.996	0.994	0.990	0.996	0.997
BRA	0.871	0.843	0.914	0.880	0.914	0.904	0.897	0.915	0.929
THA	0.783	0.804	0.937	0.838	0.929	0.932	0.870	0.937	0.944
TUR	0.962	0.965	0.994	0.978	0.994	0.995	0.991	0.995	0.996
IRN	0.855	0.860	0.979	0.863	0.963	0.962	0.960	0.972	0.982
USA	0.984	0.977	0.993	0.985	0.992	0.992	0.988	0.993	0.994
MEX	0.685	0.713	0.807	0.799	0.805	0.805	0.654	0.812	0.817
BGD	0.304	0.215	0.928	0.366	0.933	0.925	0.565	0.898	0.934
IDN	0.714	0.720	0.837	0.792	0.834	0.836	0.831	0.837	0.838
PHL	0.779	0.785	0.955	0.807	0.952	0.942	0.816	0.958	0.964
IND	0.921	0.925	0.938	0.943	0.934	0.942	0.942	0.937	0.947
RUS	0.905	0.902	0.943	0.934	0.943	0.940	0.923	0.943	0.944
PAK	0.714	0.702	0.984	0.719	0.984	0.977	0.968	0.982	0.985



Şekil 5. R² metriğine göre deneysel sonuçlar

SONUÇLAR

Covid-19 salgını neredeyse dünyanın her ülkesindeki insanların yaşamlarını ciddi bir şekilde etkilemiştir. Hastalığın kesin bir tedavisi olmaması sebebiyle pandeminin gelecek seyrini tahmin etmek

zordur. Ancak hastalığın seyrini hafifletmek amacıyla aşılarda geliştirilmiştir. Bu çalışmada, Mart 2022 yılına kadar ülkelerin aşılama verilerini içeren bir veri seti kullanılarak günlük yapılan aşı sayısını tahmin etmek için makine öğrenmesi ve derin öğrenme

Research article/Araştırma makalesi
 DOI:10.29132/ijpas.1125729

modellerinin uygulamalı bir analizi sunulmuştur. Yapılan toplam aşı sayısı ve nüfusa göre hesaplanan aşılama oranları en yüksek ilk 20 ülke belirlenerek bu ülkelerde günlük olarak yapılan aşı sayılarının tahmin edilmesi amaçlanmıştır. DT, kNN, LR, RF, SVM, MLP, CNN, RNN ve geliştirilen LSTM tabanlı derin öğrenme modeli RMSE, MAE ve R² metrikleri kullanılarak karşılaştırılmıştır. Geliştirilen LSTM tabanlı modelin karşılaştırılan modellere göre günlük yapılan aşı sayısı tahmininde en başarılı tahmin performansına sahip olduğu görülmüştür. LSTM'in ardından RNN, MLP, LR ve SVM karşılaştırılan diğer modellere göre daha başarılı sonuçlara sahip olmuştur. Geliştirilen model uygulanan ülkelerin tamamına yakınında 0.90'ın üzerinde R² değerine sahip olmuştur. Bu çalışmada elde edilen sonuçlar aşı ve enjektör gibi malzemelerin tedariği ile sağlık personellerinin çalışma saatlerinin planlanmasında ve aşı yaptıracak kişilerin gruplar halinde aşılmasında stratejiler geliştirmek için günlük hayata uyarlanabilir. Bu çalışma ile gelecekte günlük aşı miktarının tahmin edilmesi ile etkili bir model geliştirilmiş ve gelecekte bu alanda çalışacak araştırmacılar için bir referans kaynağı olmak amaçlanmıştır. Bu çalışmada elde edilen başarı oranı hibrit yöntemler kullanılarak artırılabilir.

ÇIKAR ÇATIŞMASI BEYANI

Yazarlar bu makale ile ilgili herhangi bir çıkar çatışması bildirmemektedir.

ARAŞTIRMA VE YAYIN ETİĞİ BEYANI

Yazarlar bu çalışmanın araştırma ve yayın etiğine uygun olduğunu beyan eder.

KAYNAKLAR

- Abbasimehr, H. ve Paki, R. (2021). Prediction of COVID-19 confirmed cases combining deep learning methods and Bayesian optimization. *Chaos, Solitons & Fractals*, 142, 110511.
- Alassafi, M. O. Jarrah, M. ve Alotaibi, R. (2022). Time series predicting of COVID-19 based on deep learning. *Neurocomputing*, 468, 335-344.
- Alazab, M. Awajan, A. Mesleh, A. Abraham, A. Jatana, V. ve Alhyari, S. (2020). COVID-19 prediction and detection using deep learning. *International Journal of Computer Information Systems and Industrial Management Applications*, 12 (June), 168-181.
- Arora, P. Kumar, H. ve Panigrahi, B. K. (2020). Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. *Chaos, Solitons & Fractals*, 139, 110017.

- Bisgin, A. Sanlioglu, A. D. Eksi, Y. E. Griffith, T. S. ve Sanlioglu, S. (2021). Current update on severe acute respiratory syndrome coronavirus 2 vaccine development with a special emphasis on gene therapy viral vector design and construction for vaccination. *Human Gene Therapy*, 32(11-12), 541-562.
- Bodapati, J. D. ve Veeranjaneyulu, N. (2019). Feature extraction and classification using deep convolutional neural networks. *Journal of Cyber Security and Mobility*, 261-276.
- Che Azemin, M. Z. Hassan, R. Mohd Tamrin M. I. ve Md Ali, M. A. (2020). COVID-19 deep learning prediction model using publicly available radiologist-adjudicated chest X-ray images as training data: preliminary findings. *International Journal of Biomedical Imaging*, 2020.
- Cucinotta, D. ve Vanelli, M., (2020). "WHO declares COVID-19 a pandemic." *Acta bio-medica: Atenei Parmensis*, vol. 91, no. 1, pp. 157-160.
- Daily and Total Vaccination for COVID-19 in the World from Our World in Data, <https://www.kaggle.com/datasets/gpreda/covid-world-vaccination-progress> (Erişim Tarihi: 20 Mart 2021)
- Franco-Lopez, H. Ek, A. R. ve Bauer, M. E. (2001). Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. *Remote sensing of Environment*, 77(3), 251-274.
- Ghritlahre, H. K. ve Prasad, R. K. (2018). Exergetic performance prediction of solar air heater using MLP, GRNN and RBF models of artificial neural network technique. *Journal of environmental management*, 223, 566-575.
- Gupta, A. Sardar, P. Cash, M. E. Milani, R. V. ve Lavie, C. J. (2021). Covid-19 vaccine-induced thrombosis and thrombocytopenia-a commentary on an important and practical clinical dilemma. *Progress in cardiovascular diseases*.
- Hazra, A. ve Gogtay, N. (2016). Biostatistics series module 6: correlation and linear regression. *Indian journal of dermatology*, 61(6), 593.
- Hou, R. Huang, C. R. Zhou, M. ve Jiang, M. (2019). Distance between Chinese registers based on the Menzerath-Altmann law and regression analysis. *Glottometrics*, 45, 24-57.
- Ibrahim, I. ve Abdulazeez, A. (2021). The role of machine learning algorithms for diagnosing diseases. *Journal of Applied Science and Technology Trends*, 2(01), 10-19.
- Jamal, M. Shah, M. Almarzooqi, S. H. Aber, H. Khawaja, S. El Abed, R. ... ve Samaranayake, L. P. (2021). Overview of transnational recommendations for COVID-19 transmission control in dental care settings. *Oral diseases*, 27, 655-664.

Research article/Araştırma makalesi
 DOI:10.29132/ijpas.1125729

- Kızılluluk, S. ve Can, U. (2021). Kalite Test Fonksiyonları Kullanılarak Güncel Metasezgisel Optimizasyon Algoritmalarının Karşılaştırılması. *International Journal of Pure and Applied Sciences*, 7(1), 100-112.
- Lalmuanawma, S. Hussain, J. ve Chhakchuak, L. (2020). Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. *Chaos, Solitons & Fractals*, 139, 110059.
- Laroui, S. Omara, H. LAZAAR, M. ve MAHBOUB, O. (2019). Comparative study of performing features applied in CNN architectures. In *ICCWCS 2019: Third International Conference on Computing and Wireless Communication Systems*, April 24-25, Faculty of Sciences, Ibn Tofail University-Kénitra-Morocco (p. 313).
- Liu, G. ve Guo, J. (2019). Bidirectional LSTM with attention mechanism and convolutional layer for text classification. *Neurocomputing*, 337, 325-338.
- Liu, Y. Gong, C. Yang, L. ve Chen, Y. (2020). DSTP-RNN: A dual-stage two-phase attention-based recurrent neural network for long-term and multivariate time series prediction. *Expert Systems with Applications*, 143, 113082.
- Liu, Y. Liu, S. Wang, Y. Lombardi, F. ve Han, J. (2018). A stochastic computational multi-layer perceptron with backward propagation. *IEEE Transactions on Computers*, 67(9), 1273-1286.
- Nurcahyanto, H. Prihatno, A. T. Alam, M. M. Rahman, M. H. Jahan, I. Shahjalal, M. ve Jang, Y. M. (2022). Multilevel RNN-Based PM10 Air Quality Prediction for Industrial Internet of Things Applications in Cleanroom Environment. *Wireless Communications and Mobile Computing*, 2022.
- Pinter, G. Felde, I. Mosavi, A. Ghamisi, P. ve Gloaguen, R. (2020). COVID-19 pandemic prediction for Hungary; a hybrid machine learning approach. *Mathematics*, 8(6), 890.
- Prasad, A. M. Iverson, L. R. ve Liaw, A. (2006). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems*, 9(2), 181-199.
- Rahman, A. S. Shamrat, F. J. M. Tasnim, Z. Roy, J. ve Hossain, S. A. (2019). A comparative study on liver disease prediction using supervised machine learning algorithms. *International Journal of Scientific & Technology Research*, 8(11), 419-422.
- Royer, H. D. ve Reinherz, E. L. (1987). T lymphocytes: ontogeny, function, and relevance to clinical disorders. *New England Journal of Medicine*, 317(18), 1136-1142.
- Rubin, R. (2021). COVID-19 vaccines vs variants—determining how much immunity is enough. *Jama*, 325(13), 1241-1243.
- Samaranayake, L. P. Seneviratne, C. J. ve Fakhrudin, K. S. (2021). Coronavirus disease 2019 (COVID-19) vaccines: A concise review. *Oral diseases*.
- Sen, P. C. Hajra, M. ve Ghosh, M. (2020). Supervised classification algorithms in machine learning: A survey and review. In *Emerging technology in modelling and graphics* (pp. 99-111). Springer, Singapore.
- Shivanna, A. ve Agrawal, D. P. (2020). Prediction of defaulters using machine learning on Azure ML. In *2020 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)* (pp. 0320-0325).
- Song, Y. Y. ve Ying, L. U. (2015). Decision tree methods: applications for classification and prediction. *Shanghai archives of psychiatry*, 27(2), 130.
- Staudemeyer, R. C. ve Morris, E. R. (2019). Understanding LSTM—a tutorial into long short-term memory recurrent neural networks. *arXiv preprint arXiv:1909.09586*.
- Şengür, D. (2021). KOVİD-19 Salgını Sırasında Öğrencilerin Öğrenme Alışkanlıklarının Schur Ayırıştırma Tabanlı Dalgacık Aşırı Öğrenme Makineleri ile Tahmini. *International Journal of Pure and Applied Sciences*, 7(1), 13-18.
- Tomppo, E. ve Halme, M. (2004). Using coarse scale forest variables as ancillary information and weighting of variables in k-NN estimation: a genetic algorithm approach. *Remote Sensing of Environment*, 92(1), 1-20.
- Wang, N. Shang, J. Jiang, S. ve Du, L. (2020). Subunit vaccines against emerging pathogenic human coronaviruses. *Frontiers in microbiology*, 11, 298.
- Wang, P. Zheng, X. Ai, G. Liu, D. ve Zhu, B. (2020). Time series prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: Case studies in Russia, Peru and Iran. *Chaos, Solitons & Fractals*, 140, 110214.
- Younis, M. C. (2021). Evaluation of deep learning approaches for identification of different coronavirus species and time series prediction. *Computerized Medical Imaging and Graphics*, 90, 101921.
- Zhang, L. ve Yan, W. Q. (2020). Deep learning methods for virus identification from digital images. In *2020 35th International Conference on Image and Vision Computing New Zealand (IVCNZ)* (pp. 1-6). IEEE 39.
- Zhang, N. Cai, Y. X. Wang, Y. Y. Tian, Y. T. Wang, X. L. ve Badami, B. (2020). Skin cancer diagnosis based on optimized convolutional neural network. *Artificial intelligence in medicine*, 102, 101756.