

Covid-19 Tespitinde Makine Öğrenmesi: Kan Gazı Analizine Dayalı Bir Yaklaşım

Faruk Ayata

Van Yüzüncü Yıl Üniversitesi, Başkale Meslek Yüksek Okulu, Van, TÜRKİYE

e-mail: farukayata@yyu.edu.tr

DOI: 10.57244/dfbd.1492816

Geliş tarihi/Received:30/05/2024

Kabul tarihi/Accepted:04/08/2024

Özet

COVID-19'un ortaya çıkmasıyla birlikte, dünya genelinde bilim insanları ve uluslar bu salgınla mücadeleye odaklandı. Hala süregelen araştırmalar, hastalığın yayılma ve korunma yöntemleri, güvenlik için geliştirilen aşılarda ve ilaçlar gibi konuları dünya gündeminin merkezine yerleştiriyor. Bu noktada, salgının kontrolünde kullanılan testlerin doğruluğu kritik bir önem taşımaktadır. Yeni varyantların ortaya çıktığı bu dönemde, yapay zekâ ve hastalık teşhis süreçlerinin entegrasyonu, müdahale ve önlem alma hızını artıracaktır. Bu çalışmada, salgın yönetimine katkı sağlamak amacıyla makine öğrenmesi yöntemleri kullanılarak bireylerin kan gazı değerlerinden COVID-19 test sonuçlarını tahmin etmek hedeflenmiştir. Van Yüzüncü Yıl Üniversitesi Dursun Odabaş Tıp Merkezi'nden elde edilen veri seti, bireylerden alınan kan gazı analiz örneklerinden (109 pozitif, 1146 negatif) oluşmaktadır. Bu veriler kullanılarak, bireylerin test sonuçlarını belirlemede etkili bir yöntemin geliştirilmesi amaçlanmıştır. Bu hedef doğrultusunda, COVID-19 hastalığının tahmini için Rastgele Orman (RO), Destek Vektör Makineleri (DVM), Yapay Sinir Ağları (YSA), Karar Ağaçları (KA), Aşırı Gradyan Artırma (AGA), K-en Yakın Komşu (KNN) ve Naive Bayes (NB) gibi makine öğrenmesi algoritmaları kullanılmıştır. Analizler sonucunda, KNN yönteminin %97.61 ile diğer yöntemlere kıyasla daha yüksek bir başarı sağladığı görülmüştür. Makine öğrenmesi tekniklerinin hastalık tespit ve erken müdahale gibi durumlarda önemli bir rol oynayabileceği ve bu tür yaklaşımların salgınla mücadelede etkili araçlar sunabileceği anlaşılmaktadır.

Anahtar kelimeler: Covid-19, Makine öğrenmesi yöntemleri, Kan gazları, Yapay zekâ

Machine Learning In Covid-19 Detection: An Approach Based On Blood Gas Analysis

Abstract

With the emergence of COVID-19, scientists and nations worldwide have focused on combating this pandemic. Ongoing research places topics such as the spread and prevention methods of the disease, vaccines, and drugs at the center of the global agenda. The accuracy of tests used in outbreak management has become critically important. During this period, when new variants are emerging, the integration of artificial intelligence and disease diagnosis processes will enhance the speed of intervention and prevention. This study aims to predict COVID-19 test results from individuals' blood gas values using machine learning methods to contribute to outbreak management. The dataset, obtained from Van Yüzüncü Yıl University Dursun Odabaş Medical Center, consists of blood gas analysis samples from individuals (109 positive, 1146 negative). Using this data, the study aims to develop an effective method for determining test results. For this purpose, machine learning algorithms such as Random Forest (RF), Support Vector Machines (SVM), Artificial Neural Networks (ANN), Decision Trees (DT), Extreme Gradient Boosting (XGBoost), K-Nearest Neighbors (KNN), and Naive Bayes (NB) were used. The analyses showed that the KNN method achieved a higher success rate of 97.61% compared to other methods. It is understood that machine learning techniques can play a significant role in disease detection and early intervention and that such approaches can provide effective tools in combating the pandemic.

Keywords: Covid-19, Machine learning methods, Blood gases, Artificial intelligence

GİRİŞ

COVID-19 pandemisinin başlamasıyla birlikte, hastalığın hızlı ve doğru bir şekilde tespit edilmesi, salgının kontrol altına alınmasında hayati önem taşımaktadır. Bu bağlamda, çeşitli biyokimyasal ve hematolojik parametrelerin kullanımıyla COVID-19'un teşhis edilmesi üzerine yoğun araştırmalar yapılmaktadır. Hemogram (kan sayımı) ve kan gazı analizleri, bu tür tanısal yaklaşımlar arasında öne çıkmaktadır.

Hastalık tespiti ve makine öğrenmesinin bu alandaki önemi giderek artmaktadır. Makine öğrenmesi, büyük veri kümeleri üzerinde çalışarak belirli kalıpları ve anomalileri tespit etme yeteneği sayesinde, tıbbi teşhislerde önemli bir araç haline gelmiştir. Farklı algoritmalar ve modeller, çeşitli hastalıkların erken teşhisinde kullanılmaktadır. Bir çalışmada, makine öğrenmesi tekniklerinin diyabet, böbrek hastalığı ve kalp hastalığı gibi kronik hastalıkları tespit etmede yüksek doğruluk sağladığı gösterilmiştir (Ahsan ve ark., 2022). Yapay Zeka (YZ), sağlık sistemlerinde sadece hastalık teşhisi için değil, aynı zamanda hasta yönetimi, ilaç keşfi ve risk değerlendirmesi gibi alanlarda da kullanılır. Bu teknolojiler, sağlık hizmetlerinin kişiselleştirilmesi ve hasta bakımının iyileştirilmesi için yenilikçi çözümler sunar.

COVID-19 gibi hızlı yayılan ve küresel bir tehdit oluşturan hastalıkların erken teşhis edilmesi, tedavi sürecinin etkinliğini artırmak ve sağlık sistemlerinin yükünü hafifletmek için kritik öneme sahiptir. YZ ve makine öğrenmesi teknikleri, hemogram ve kan gazı analizleri gibi veriler üzerinden hastalığın teşhisini daha hızlı ve güvenilir bir şekilde yapmayı mümkün kılmaktadır (Kucukcan ve ark., 2020; Gamsızkan ve ark., 2022).

Yapılan literatür taramaları sonucunda, hemogram (kan sayımı) ve kan gazı değerleri kullanılarak COVID-19 tespiti üzerine çeşitli çalışmaların bulunduğu tespit edilmiştir. Bu çalışmalar, kan testlerindeki belirli parametrelerin analiz edilmesiyle COVID-19 teşhisinde makine öğrenmesi ve diğer analitik tekniklerin uygulanmasını içermektedir.

Dicle Tıp Dergisi'nde yayımlanan bir çalışmada, COVID-19 hastalarının klinik ilerleyişi ile hemogram parametreleri arasındaki ilişki incelenmiştir. Bu çalışmada, 70 COVID-19 hastasının demografik verileri ve hemogram parametreleri retrospektif olarak değerlendirilmiş ve hastanede kalış süreleri ile eosinofil sayısı ve trombosit dağılım genişliği (PDW) arasında anlamlı farklar bulunmuştur. Bununla birlikte, klinik ilerleme ile kan değerleri arasında genel bir fark bulunmamıştır (Erdogmuş Kucukcan ve Kucukcan, 2020).

Bir başka çalışma, COVID-19 hastalarının kan gruplarına göre hemogram parametrelerini incelemiştir. Bu çalışmada, hemoglobin ve beyaz kan hücresi (WBC) değerlerinde COVID-19 pozitif ve negatif gruplar arasında anlamlı farklar bulunmuştur. Ancak, COVID-19 pozitif hastaların kan grupları ile laboratuvar parametreleri arasında anlamlı bir ilişki tespit edilmemiştir (Gamsızkan ve ark., 2022).

Konuralp Tıp Dergisi'nde yayımlanan başka bir çalışmada ise, COVID-19 teşhisinde platelet kütle indeksi (PMI), PLT/MPV oranı ve diğer hemogram parametrelerinin tanısal değeri araştırılmıştır. Bu çalışmada, COVID-19 pozitif ve negatif gruplar arasında lökosit, lenfosit, monosit, nötrofil sayıları, LMR, PLR, PMI ve PLT/MPV değerlerinde anlamlı farklar bulunmuştur (Şeyhanlı ve Yasak, 2021).

Kan gazı analizi, COVID-19'un şiddetinin belirlenmesi ve hipoksinin erken tespiti için de kullanılmıştır. San Raffaele Hastanesi veri seti kullanılarak yapılan bir

çalışmada, rutin laboratuvar kan test sonuçları üzerinden COVID-19'un tespiti için ensemble öğrenme modeli uygulanmış ve etkili bir yöntem geliştirilmiştir (Abayomi-Alli ve ark., 2022). Ayrıca, COVID-19'un teşhisinde venöz kan gazı analizinin yararını değerlendiren çalışmalar, hipoksinin erken tespitinin tedavi başarısı ve hayatta kalma oranı üzerinde önemli bir etkisi olduğunu göstermiştir (De Gruyter, 2021).

Ayata ve Seyyarer (2024), Van Yüzüncü Yıl Üniversitesi Dursun Odabaş Tıp Merkezi'nden aldıkları kan gazı veri seti üzerinde bir çalışma gerçekleştirmişlerdir. Amaç, bireylerin kan gazı değerlerinden pozitif veya negatif COVID-19 test sonucunu tahmin eden bir fonksiyon elde etmektir. Bu amaçla, Çok Değişkenli Doğrusal Regresyon (MLR) modeli kullanmışlardır. Çalışmada, RMSProp optimizasyon algoritması kullanılarak tüm ölçüm yöntemlerinde %58-91.23 arasında başarı oranları elde etmişlerdir (Ayata ve Seyyarer, 2024).

Gökçe ve ark. (2022) yaptıkları çalışmada COVID-19 tanısı konulan ve yoğun bakım ünitesinde (YBÜ) yatarak tedavi gören hastaların kan gazı değerlerini retrospektif olarak incelemişlerdir. 2000 hasta ve 300 sağlıklı gönüllünün verilerini değerlendirmiş. Bulgulara göre, COVID-19 pozitif hastaların pH seviyeleri negatif olanlara göre daha düşüktür. Ayrıca, pCO₂ ve pO₂ değerlerinde önemli farklılıklar olduğunu tespit etmişlerdir. Kadın ve erkek hastalar arasında oksihemoglobin seviyelerinde anlamlı farklar bulmuşlardır. Çalışma, COVID-19'un kan gazı parametreleri üzerindeki etkilerini ve CA enzim aktivitelerinin enfeksiyonlarda önemli bir rol oynayabileceğini göstermişlerdir.

Babayiğit ve ark. (2021) çalışmalarında COVID-19 nedeniyle YBÜ tedavi gören hastalarda kan transfüzyonu uygulamalarının özellikleri, ilişkili faktörleri ve sonuçlarını retrospektif olarak incelemişlerdir. 21 Mart 2020 ile 1 Haziran 2020 tarihleri arasında YBÜ'de tedavi gören ve 48 saatten fazla kalan hastaların verilerini analiz edilmişlerdir. Kan transfüzyonu yapılan hastaların yoğun bakımda kalış süreleri, yapılmayanlara kıyasla anlamlı şekilde daha uzun bulunmuştur. Transfüzyon yapılan hastalarda KOAH oranlarının yüksek olduğu ve bu hastaların mortalite oranlarının daha yüksek olduğunu tespit etmişlerdir. Çalışma, aneminin COVID-19 hastalarında yaygın olduğunu ve ağır klinik seyir, entübasyon, mekanik ventilasyon gereksinimi olan hastalarda kan transfüzyonunun sık uygulandığını ortaya koymuştur. Bu sonuçlar, COVID-19 hastalarında kan transfüzyonunun yönetimi ve ilişkili faktörler hakkında önemli bilgiler sunmaktadır.

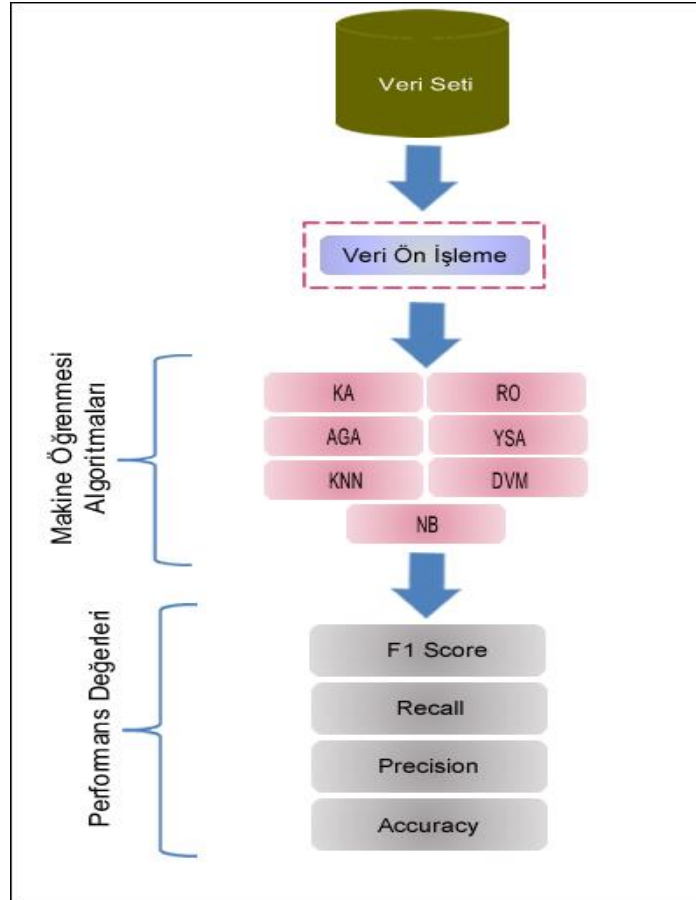
COVID-19 salgını, küresel sağlık sistemlerini ciddi şekilde zorlayan ve milyonlarca insanın hayatını etkileyen bir kriz olarak dünya genelinde büyük bir mücadele gerektirmiştir. Salgınla mücadelede, hastalığın hızlı ve doğru bir şekilde teşhis edilmesi, enfeksiyonun yayılmasını kontrol altına almak ve uygun tedavi yöntemlerini belirlemek açısından hayati önem taşımaktadır. Bu bağlamda, yapay zekâ ve makine öğrenmesi tekniklerinin sağlık alanına entegrasyonu, teşhis süreçlerinin hızını ve doğruluğunu artırarak büyük bir potansiyel sunmaktadır. Van Yüzüncü Yıl Üniversitesi Dursun Odabaş Tıp Merkezi'nden elde edilen kan gazı analiz verileri kullanılarak gerçekleştirilen bu çalışma, COVID-19 test sonuçlarının tahmin edilmesinde yüksek başarı sağlayan yöntemler geliştirmeyi hedeflemiştir. Bu doğrultuda, çalışmada laboratuvar testlerinde sıklıkla kullanılan kan gazı verilerinden COVID-19 virüsünün tespiti ve bulaş sonrasında hastalığın takibi için makine öğrenmesi yöntemlerinin performans değerlendirilmesinin yapılması amaçlanmaktadır. Model olarak RO, KA, DVM, KNN, NB, YSA, AGA, veri seti olarak da Türkiye'nin Van ilinden alınan COVID-19 veri seti kullanılmaktadır.

MATERYAL VE YÖNTEM

Bu bölümünde, çalışmada kullanılan makine öğrenmesi algoritmaları, performans ölçüm yöntemleri ve veri seti kısaca özetlenmektedir. Ayrıca sistemin çalışmasını belirten diyagram bu bölümde sunulup açıklanmaktadır.

Veri toplama ve mimari çerçeve

Çalışmada kullanılan veri seti Van Yüzüncü Yıl Üniversitesi Dursun Odabaş Tıp Merkezinden alınan kan gazı inceleme örneklerinden (109 pozitif, 1146 negatif) oluşmaktadır. Sistemin çalışma mimarisi Şekil 1’de gösterilmektedir. Elde edilen veri setine ilk önce veri ön işleme adımları uygulanmıştır. Veri ön işleme adımlarında öncelikle veri seti bir Excel dosyasından yüklenir. Kategorik değişkenler tespit edilerek, bu değişkenler Label Encoder kullanılarak sayısal değerlere dönüştürülür. Ardından, hedef değişken ('Sonuc') ve özellikler (bağımsız değişkenler) belirlenir. Veri seti eğitim ve test setlerine (%70 eğitim, %30 test) ayrılır, böylece modelin performansı bağımsız bir veri seti üzerinde değerlendirilebilir. Bu adımlar, veri setinin makine öğrenmesi algoritmaları için uygun hale getirilmesini sağlar ve modelin doğruluğunu artırır. Veri setine ayırma (Test – Eğitim) işlemi yapıldıktan sonra sekiz ML modeli ile sistem eğitilmektedir. Son olarak test veri seti ile sınıflandırmada modelin başarısı tespit edilmektedir.



Şekil 1. Mimari çerçeve

Model Seçimi

- **Rastgele Orman**
RO, birçok karar ağacının bir araya getirilmesiyle oluşan bir topluluk yöntemidir. Bu algoritma, verilerin bootstrap örneklerinden oluşturulan karar ağaçlarını kullanarak hem sınıflandırma hem de regresyon problemlerini çözer. Rastgele Orman, yüksek doğruluk oranı, genelleme hatasını azaltma ve önemli değişkenleri belirleme gibi avantajlar sağlar. Algoritmanın temel prensibi, her bir ağacın ayrı ayrı karar vermesi ve bu kararların çoğunluk oyu ile nihai sonucu belirlemesidir. Bu özellik, özellikle büyük veri setleri ve karmaşık modeller için idealdir (Probst ve ark., 2019; Özkan, 2023)
- **Karar Ağaçları**
KA, veriyi belirli kriterlere göre bölerek sınıflandırma veya regresyon problemlerini çözmeye yarayan bir makine öğrenimi algoritmasıdır. Her bir dal, bir karar kuralını veya ayrıştırıcıyı temsil ederken, yapraklar sınıflandırma veya tahmin sonuçlarını temsil eder. KA, anlaşılması ve yorumlanması kolay modeller sunar, bu da onları özellikle veri keşfi ve öngörü analizi için popüler kılar (Kudyba, 2014; Gür, 2023).
- **Destek Vektör Makineleri**
DVM, sınıflandırma ve regresyon analizinde kullanılan güçlü bir denetimli öğrenme algoritmasıdır. DVM, veriyi yüksek boyutlu bir uzaya projeler ve veriyi en iyi şekilde ayıran hiperdüzlemi bulmaya çalışır. Bu yöntem, özellikle sınıflandırma problemlerinde yüksek doğruluk oranları ile dikkat çeker ve genellikle karmaşık veri kümeleri üzerinde etkili sonuçlar üretir (Özdemir Güler, 2023).
- **K-en Yakın Komşu**
KNN, en basit ve en sezgisel makine öğrenimi algoritmalarından biridir. Bu algoritma, sınıflandırma veya regresyon problemlerinde bir veri noktasını, en yakın K komşusuna göre sınıflandırır veya tahmin eder. KNN, genellikle parametrik olmayan bir yöntem olarak kabul edilir ve verinin doğrudan örneklenmesi ile çalışır. Bu yöntem, özellikle verinin dağılımı hakkında herhangi bir varsayım gerektirmemesi nedeniyle esneklik sağlar (Hacıbeyoğlu, 2023).
- **Naive Bayes**
NB, olasılık teorisine dayanan basit ancak güçlü bir sınıflandırma algoritmasıdır. Bu algoritma, her bir özelliğin sınıftan bağımsız olduğunu varsayar ve bu varsayım altında Bayes teoremi kullanarak sınıflandırma yapar. NB, özellikle metin sınıflandırma ve spam filtreleme gibi uygulamalarda yaygın olarak kullanılır (Ayan, 2024).
- **Yapay Sinir Ağları**
YSA, biyolojik sinir ağlarından esinlenerek geliştirilmiş makine öğrenimi algoritmalarıdır. Bu ağlar, katmanlar halinde düzenlenmiş yapay nöronlardan oluşur ve veri işleme, sınıflandırma, regresyon gibi çeşitli görevleri

gerçekleştirebilir. YSA'lar, derin öğrenme tekniklerinin temelini oluşturur ve özellikle büyük veri setleri ve karmaşık görevler için güçlüdür (Karaatlı ve ark., 2012; Odabaşı, 2023).

- **Aşırı Gradyan Artırma**

AGA, gradyan artırma algoritmasının bir genişletmesi olup, performansı ve hesaplama hızını artırmak için optimize edilmiştir. AGA, veri setindeki hataları azaltmak için ardışık ağaçlar oluşturur ve her bir yeni ağaç, önceki ağaçların hatalarını düzeltmeye çalışır. Bu yöntem, özellikle yapılandırılmış verilerle çalışırken yüksek doğruluk oranları ile bilinir (Chen ve Guestrin, 2016; Şimşek, 2024).

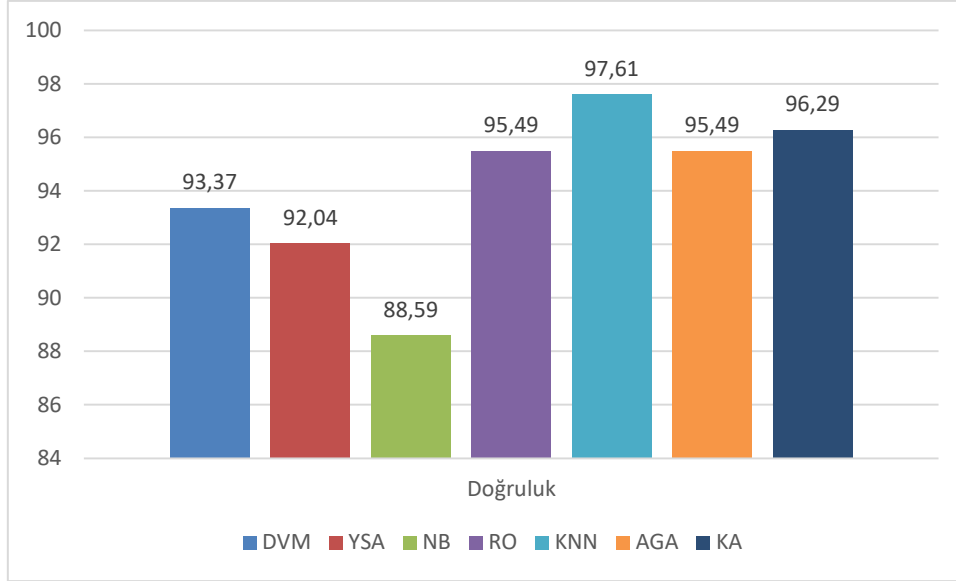
BULGULAR VE TARTIŞMA

Tablo 1 incelendiğinde, KNN algoritması, en yüksek doğruluk (%97.61), kesinlik (%98.74), ve F1 skoru (%90.54) ile genel olarak en iyi performansı gösterirken, NB algoritması en yüksek duyarlılık (%81.62) ile öne çıkmaktadır. DVM algoritması yüksek kesinlik (%96.64) sağlasa da düşük duyarlılık (%58.33) ve F1 skoru (%62.55) ile dikkat çekmektedir. YSA algoritması, özellikle düşük kesinlik (%46.02) ve F1 skoru (%47.93) ile genel olarak en düşük performansı sergilemektedir.

Tablo 1. Makine öğrenmesi algoritmalarının performans değerleri

| Performans Metrikleri | | | | |
|-----------------------|----------|----------|------------|---------|
| ML Alg. | Doğruluk | Kesinlik | Duyarlılık | F1 Skor |
| DVM | 93.37 | 96.64 | 58.33 | 62.55 |
| YSA | 92.04 | 46.02 | 50.00 | 47.93 |
| NB | 88.59 | 68.05 | 81.62 | 72.06 |
| RO | 95.49 | 86.44 | 80.80 | 83.33 |
| KNN | 97.61 | 98.74 | 85.00 | 90.54 |
| AGA | 95.49 | 84.33 | 85.37 | 84.84 |
| KA | 96.29 | 88.14 | 85.80 | 86.93 |

Bu grafik, farklı makine öğrenimi algoritmalarının doğruluk performansını karşılaştırıyor. KNN algoritması %97.61 ile en yüksek doğruluk oranına sahipken, KA algoritması %96.29, RO ve AGA algoritmaları ise %95.49 ile onu takip etmektedir. DVM ve YSA algoritmaları ise sırasıyla %93.37 ve %92.04 doğruluk oranlarına sahiptir. NB algoritması ise %88.59 doğruluk oranıyla en düşük performansı sergilemektedir.



Şekil 2. Makine öğrenmesi algoritmalarının doğruluk değerleri

Şekil 2’de farklı makine öğrenimi algoritmalarının doğruluk performansları karşılaştırılmaktadır.

Tablo 2’de AGA, RO ve KA modelleri, az sayıda yanlış pozitif (FP) ve yanlış negatif (FN) değerlerle dengeli bir performans sergileyerek yüksek doğru negatif (TN) ve iyi düzeyde doğru pozitif (TP) sonuçlar gösterdiği görülmektedir. Naive Bayes modeli, yüksek FP değeriyle dikkat çekerken, doğru pozitif (TP) değerleri de bu modeller arasında en yükseklerden biri. KNN modeli ise sıfır FP değeriyle ve yüksek TN ile pozitif sınıfı tanıma yeteneğini göstermektedir. Buna karşın, DVM ve YSA modelleri hiç TP değeri elde edememişler, bu da pozitif sınıfı tamamen yanlış sınıflandırdıklarını göstermektedir. Bu durum, AGA, RO, KA ve KNN modellerinin daha etkili olabileceğini işaret etmektedir.

Tablo 2. Karmaşıklık matris değerleri

| Model | TN | FP | FN | TP |
|-------|-----|----|----|----|
| AGA | 338 | 9 | 8 | 22 |
| RO | 342 | 5 | 10 | 20 |
| DVM | 347 | 0 | 30 | 0 |
| KNN | 347 | 0 | 9 | 21 |
| KA | 341 | 6 | 8 | 22 |
| NB | 312 | 35 | 8 | 22 |
| YSA | 347 | 0 | 30 | 0 |

SONUÇLAR

Bu çalışma, makine öğrenmesi yöntemlerini kullanarak bireylerin kan gazı değerlerinden COVID-19 test sonuçlarını tahmin etmeyi amaçlamaktadır. Bu doğrultuda çalışmada incelenen çeşitli makine öğrenimi algoritmaları arasında, KNN algoritması %97.61 doğruluk, %98.74 kesinlik ve %90.54 F1 skoru ile en iyi genel

performansı sergilerken, NB algoritması %81.62 ile en yüksek duyarlılık değerine sahiptir ancak %88.59 doğrulukla diğerlerinden daha düşük performans gösterir. DVM ve YSA algoritmaları, sırasıyla düşük duyarlılık (%58.33) ve F1 skoru (%62.55) ile ve düşük kesinlik (%46.02) ve F1 skoru (%47.93) ile dikkat çekerek, bu algoritmaların bu veri seti için yeterince etkili olmadığını ortaya koymaktadır.

Karmaşıklık matrisine göre, AGA, RO ve KA modelleri dengeli performanslar sergilemiş, az sayıda yanlış pozitif ve yanlış negatif değerler ile yüksek doğru negatif (TN) ve iyi düzeyde doğru pozitif (TP) sonuçlar göstermişlerdir. Ancak, DVM ve YSA gibi modeller hiç TP değeri elde edememişlerdir, bu da bu modellerin pozitif sınıfı doğru bir şekilde sınıflandırmada ciddi zorluklar yaşadıklarını gösterir.

Sonuç olarak, veri setinin özelliklerine ve ihtiyaç duyulan performans metriklerine göre model seçimi yapılmalıdır. Özellikle yüksek doğruluk ve kesinlik gerektiren uygulamalar için KNN, dengeli bir performans arayışında ise AGA, RO ve KA modelleri tercih edilebilir. NB, yüksek duyarlılık gerektiren durumlar için uygun olabilir, ancak genel doğruluk düşük olduğundan dikkatli kullanılmalıdır.

Yapay zekâ tabanlı teşhis yöntemlerinin, hastalık teşhisindeki doğruluğu artırarak ve yanıt sürelerini kısaltarak daha geniş kapsamlı kullanımını desteklemektir. Bu yöntemlerin geniş çaplı entegrasyonu, pandemiye yanıt verme kapasitemizi önemli ölçüde güçlendirecektir. Mevcut yapılan çalışmaya ek olarak aşağıda belirtilen öneriler verilebilir:

- Farklı demografik gruplardan, coğrafi bölgelerden ve hastalık şiddet seviyelerinden daha fazla veri toplayarak modelin genelleme yeteneği artırılabilir.

- XGBoost, LightGBM gibi ileri modelleri veya Deep Learning teknikleri araştırılıp ve ensemble yöntemlerle birden fazla model birleştirilerek performans artırılabilir.

- Kan gazı analizlerine ek olarak, diğer biyomarkerler ve klinik verileri modele dahil edilebilir ve özellik seçimi algoritmaları kullanarak en etkili özellikler belirlenebilir.

Çıkar Çatışması

Makale yazarları aralarında herhangi bir çıkar çatışması olmadığını beyan ederler.

Yazar Katkısı

Yazarlar makaleye eşit oranda katkı sağlamış olduklarını beyan ederler.

KAYNAKLAR

Abayomi-Alli, O. O., Damaševičius, R., Maskeliūnas, R., & Misra, S. (2022). An Ensemble Learning Model for COVID-19 Detection from Blood Test Samples. *Sensors*, 22(6), 2224. <https://doi.org/10.3390/s22062224>.

Ahsan, M. M., Luna, S. A., & Siddique, Z. (2022). Machine-Learning-Based Disease Diagnosis: A Comprehensive Review. *Healthcare*, 10(3), 541. <https://doi.org/10.3390/healthcare10030541>

Ayan, S., & Bilgin, T. T. (2024). Uyku Sağlığı ile Yaşam Tarzı Arasındaki İlişkinin PCA, Naive Bayes ve Rastgele Orman Ağaçları Yöntemleri ile İncelenmesi ve Karşılaştırılması. *Uluslararası Yönetim Bilişim Sistemleri Ve Bilgisayar Bilimleri Dergisi*, 8(1), 41-56. <https://doi.org/10.33461/uybisbbd.1415925>

- Ayata F, Seyyarer E. (2024). COVID-19 Diagnosis from Blood Gas Using Multivariate Linear Regression. *Hittite J Sci Eng.* 11(1),15-23.
- Babayiğit, M., Yıldız, G., Türksal, E., Babayiğit, M. A., Fırat, P., Selvi, A., & Kılınçarslan, A. (2021). Yoğun Bakımdaki COVID-19 Hastalarında Kan Transfüzyonu ve İlişkili Faktörlerin Analizi: Tek Merkez Örneği. *Batı Karadeniz Tıp Dergisi*, 5(3), 366-373.
- Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining*,785-794. <https://doi.org/10.1145/2939672.2939785>
- De Gruyter. (2021). Venous blood gas analysis in patients with COVID-19 symptoms in the emergency department. De Gruyter. Retrieved from <https://www.degruyter.com/document/doi/10.1515/cclm-2020-0975/html>.
- Erdogmuş Kucukcan, N., Kucukcan, A. (2020). The relationship between hemogram parameters with clinical progress in COVID-19 patients. *Dicle Tıp Dergisi*, 47(4), 763-769. <https://doi.org/10.5798/dicletip.850158>.
- Gamsızkan, Z., Kaya, A., Davran, F., Tuncel, M. C., Işık, A., Şahin, İ. (2022). COVID-19 Hastalarının Kan Gruplarına Göre Hemogram Parametrelerinin Değerlendirilmesi. *Value in Health Sciences*, 12(2), 208-214. <https://doi.org/10.33631/sabd.1108566>.
- Gökçe, H. Ş., Güler, Ö. Ö., Şimşek, E., Karagülleoğlu, Z. Y., & Çarhan, A. (2022). COVID-19 Hastalarında Yeni Bir Yaklaşım Olarak Oksihemoglobin Karboksihemoglobin, Kan Gazı Değerlerinin İncelenmesi: Longitudinal Bir Çalışma. *Yoğun Bakım Hemşireliği Dergisi*, 26(3), 92-99.
- Gür, Ö. (2023). Karar Ağacı Destekli Hile Tespiti ve Bir Uygulama. *Alanya Akademik Bakış*, 7(1), 511-528. <https://doi.org/10.29023/alanyaakademik.1196078>.
- Hacıbeyoglu, M., Çelik, M., & Erdaş Çiçek, Ö. (2023). K En Yakın Komşu Algoritması ile Binalarda Enerji Verimliliği Tahmini. *Necmettin Erbakan Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 5(2), 65-74. <https://doi.org/10.47112/neufmbd.2023.10>
- Karaatlı, M., Helvacıoğlu, Ö., Ömürberk, N., ve Tokgöz, G., 2012. Yapay sinir ağırları yöntemi ile otomobil satış tahmini. *Uluslararası Yönetim İktisat ve İşletme Dergisi*, 8(17), 87-100.
- Kudyba, S. (2014). Big Data, Mining, and Analytics: Components of Strategic Decision Making. *CRC Press, Taylor & Francis Group, Florida*.
- Odabaşı, M. B., & Cengiz Toklu, M. (2023). Yapay Sinir Ağları ve Derin Öğrenme Algoritmalarının Kripto Para Fiyat Tahmininde Karşılaştırmalı Analizi. *Journal of Intelligent Systems: Theory and Applications*, 6(2), 96-107. <https://doi.org/10.38016/jista.1228629>
- Özdemir Güler, Z., Bakır, M. A., & Kardiye, F. (2023). Destek vektör makinesi ile elde edilen olasılık çıktılarına dayalı yeni bir istatistiksel süreç izleme yöntemi. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 39(2), 1099-1112. <https://doi.org/10.17341/gazimmfd.1192354>
- Özkan, B., Parim, C., & Çene, E. (2023). Ülkelerin Gelişmişlik Düzeylerinin Karar Ağacı ve Rastgele Orman Yöntemleriyle Tahmin Edilmesi. *EKOIST Journal of Econometrics and Statistics* (38), 87-104. <https://doi.org/10.26650/ekoist.2023.38.1172190>.

- Probst P, Wright MN, ve Boulesteix A. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: data mining and knowledge discovery*, 9(3), e1301.
- řeyhanlı, E. S., & Yasak, İ. H. (2021). Diagnostic Value of Platelet Mass Index, Plt/Mpv Ratio and Other Hemogram Parameters in Covid-19 Patients Who Presented to Emergency Department. *Konuralp Tıp Dergisi*, 13(1), 101-107. <https://doi.org/10.18521/ktd.826613>.
- řimřek, F. F. (2024). Optik ve radar görüntüleri ile aşırı gradyan artırma algoritması kullanılarak tarımsal ürün desen tespiti. *Geomatik*, 9(1), 54-68. <https://doi.org/10.29128/geomatik.1332997>