

# Makine Öğrenmesi Yöntemleri ile Kan Sayımında Covid-19 Tanısı

## Diagnosis of Covid-19 by Machine Learning from Blood Test

Büşra ÇAKI

Ege Üniversitesi

Bilgisayar Mühendisliği Bölümü

İzmir- Türkiye

govbusra@gmail.com

ORCID:0000-0002-2804-4051

Ahmet EGESÖY

Ege Üniversitesi

Bilgisayar Mühendisliği Bölümü

İzmir -Türkiye

ahmet.egesoy@ege.edu.tr

ORCID:0000-0002-5050-5547

Yasemin TOPALOĞLU

Ege Üniversitesi

Bilgisayar Mühendisliği Bölümü

İzmir - Türkiye

yasemin.topaloglu@ege.edu.tr

ORCID:0000-0003-28161984

### Öz

Sağlık alanında kullanılan yapay zekâ teknolojileri, makine öğrenmesi yöntemleri öncülüğünde; erken tanı, değerlendirme ve karar verme gibi pek çok alanda etkili olmaktadır. Bu teknolojiler küresel bir salgına neden olmuş olan Covid-19 hastalığına tanı koymak ve gelişiminin izlenmesinde önemli başarı göstermiştir. Bu çalışmada hastalığı tanılamak ile birlikte, şiddetini ve bulaşıcılık düzeyinin saptanmasında, makine öğrenme yöntemlerini deneyimleyen çalışmalar taranmıştır. Bu incelemede tüm deneyimlerin dikkate alınması ve anlamlı sonuçlara ulaşılması amacıyla bazı eksik veriler tamamlanmış benzer çalışmalar birleştirilmiş ve sonuçlar karşılaştırılabilir hale getirilmiştir. Ayrıca hata maliyet analizini temel alarak değerlendirme ölçülerine F-β ölçütleri de katılmıştır. Yapay sinir ağı yönteminin hastalığın tanılanması ve şiddetini belirlemede başarılı olduğu görülmektedir. Bulaşma hızının belirlenmesinde ise hangi yöntemin daha iyi olduğunu belirlemede kullanılabilecek güvenilir bir model henüz yoktur.

**Anahtar sözcükler:** Makine öğrenmesi, Covid-19, gerçek-zamanlı ters transkripsiyon polimeraz zincir reaksiyonu, kan sayımı.

### Abstract

Artificial intelligence technologies that are used in healthcare are led by machine learning methods and are effective in many areas such as early diagnosis, assessment and decision making. These technologies have shown significant success in diagnosing and monitoring the course of Covid-19 disease, which has caused a global epidemic. In this paper, studies that experimented with machine learning methods in diagnosing the disease, determining its severity and level of contagion

were reviewed. In the review process, in order to take all of the experiences into account and reach meaningful outcomes, some missing data were completed and similar studies were combined so as to make the results comparable. In addition, F-β criteria were included in the evaluation metrics based on error cost analysis. The artificial neural network method was found to be successful in diagnosing the disease and determining its severity. There is not yet a reliable model that can be used, to determine which method is better in determining the infection rate.

**Keywords:** Machine learning, Covid-19, real-time reverse transcription polymerase chain reaction, blood analysis.

### 1. Giriş

Koronavirüs Hastalığı (Covid-19), 31 Aralık 2019 tarihinde Çin'in Wuhan Eyaletinde ortaya çıkışının ardından kısa sürede tüm dünyaya yayılarak, Dünya Sağlık Örgütü (WHO) tarafından salgın hastalık unvanını kazanmıştır [1]. Aşı ve ilaçların keşfine kadar dünya çapında bu hastalıkla mücadelede en önemli dönüm noktası erken tanılama başarılması olmuştur [1,2]. Hastalığın varlığının ve şiddetinin saptanması, hastalığın tedavisi ve bulaşmanın önlenmesi açısından hayati önem taşımaktadır.

Hastalığın tanılanmasında RT-PCR Sınaması olarak da bilinen Tanı Sınaması ve halk arasında Kan Sınaması olarak anılan Antikor Sınaması kullanılmaktadır. Bir moleküler genetik yöntemi olan RT-PCR sınavında enfeksiyonun varlığı, bizzat antijenin nükleik asit dizileri üzerinden saptanır. RT-PCR sınav verilerinin etkili biçimde yorumlanmasında ne yazık ki hala zorluklar yaşanmaktadır [3]. Diğer bir tanı yöntemi olan Antikor sınavı ise daha önce korona-virüs hastalığına yakalanıp yakalanmadığınızı gösteren bir kan sınavıdır. Bu sınav genellikle virüs vücudu tamamen terk ettikten sonra yapılmaktadır ve kanda virüse karşı antikor gelişip

gelişmediğini saptamaktadır. Yapılan gözlemler, Antikor sinamalarının %100 bulma ve %99 tutturma oranlarına ulaştığını göstermektedir [4]. Hastalık tanısında antikor sinamalarının RT-PCR sinama sonuçlarına göre daha güvenilir ve etkili olduğuna ilişkin kanılar bulunmaktadır [5,6].

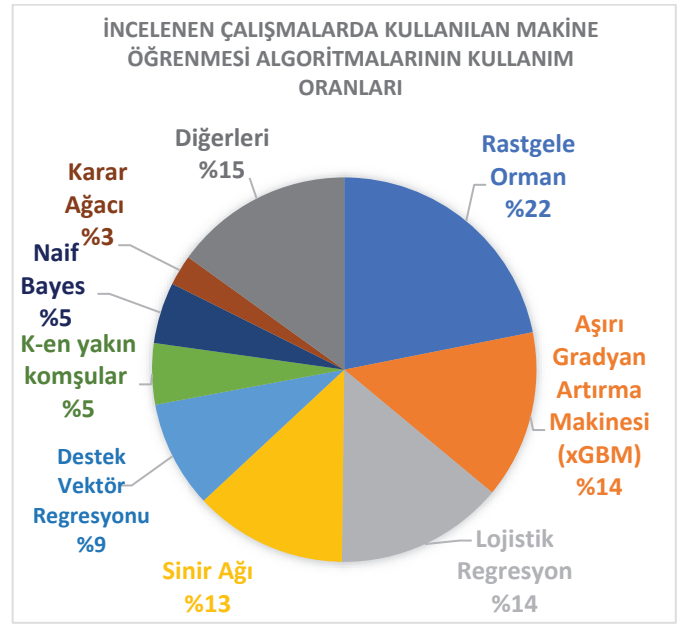
Kan sinamalarının Covid-19 tanılmasında ve hastalığın şiddetini belirlemedeki başarısı, bu tür sinamaların ön plana çıkmasını sağlamaktadır [6,7]. Karar verme aşamasında kan ölçümlerinin yanı sıra, yardımcı veriler olarak ulamsal değişkenlerin de (cinsiyet, yaş ve nefes darlığının varlığı gibi verilerin) hastalığın tanısında etkili olabileceği saptanmıştır [8,9].

Covid-19 tanısı, sadece tıbbi bir sorun olmanın ötesinde, felaketin ekonomik etkileri ve insan hayatının ciddi tehlikeye girmesi nedeniyle, güvenilir veri analizi, eşgüdüm, ölçütleme ve otomasyon gibi bilgi-işlem sorunlarını da beraberinde getirmektedir. Bu karmaşıklık, sadece tıp alanında değil, aynı zamanda bilgi teknolojileri ve veri yönetimi alanlarında da etkili çözümler gerektirmektedir.

Son zamanlarda geliştirilen yapay zekâ temelli makine öğrenmesi teknolojileri, çeşitli ölçütleri, özellikleri ve görüntüleri bir araya getirme yetenekleri sayesinde, hastalık tanısı ve sınıflandırılmasında önemli bir rol oynamıştır [10]. Bunun yanında salgının, ülkelerin mücadele kapasitesini zorlaması ve kontrolden çıkma riskinin oluşması nedeniyle, vaka sayılarındaki zaman içinde oluşan değişimlerin öngörülmesi alanında da çalışmalar yapılmıştır. Bu konuda başarılı bir çalışma [11]'de polinom regresyon yöntemiyle yapılmış olan modellemedir. Bu tür modellerin güvenilirlik kazanması durumunda, gelecekte ortaya çıkabilecek benzer salgınlarda da hastanelerde bulundurulması gereken aşı stoku veya boş yatak kapasitesi gibi kaynakların planlanması için önemli birer rehber olarak kullanılabilirliklerini belirtmek önemlidir.

Salgınla mücadele kapsamında, tanı amaçlı olarak bir dizi algoritmaya dayanan çok sayıda makine öğrenmesi biçimleri önerilmiştir [12,13]. Bu önerilen makine öğrenmesi biçimleri salgının etkilerini anlamak ve daha etkili teşhis yöntemleri geliştirmek amacıyla geniş bir veri kümesini değerlendirmektedir. Makine öğrenmesi yöntemlerinin girdisi olarak, RT-PCR sinama sonuçları ve kan sinaması gibi kritik sağlık verileri üzerine odaklanılmıştır. Bu veri tabanlı yaklaşımlar, hastalığın tanılanması için daha kesin ve hızlı sonuçlar elde etmeyi amaçlamaktadır. Kaynaklarda yer alan çalışmalarda en çok kullanılan makine öğrenme algoritmalarının başında %21,75 ile Rastgele Orman ve %14,10 oranıyla aşırı Gradyan Artırma makinesi olduğu gözlemlenmektedir.

Şekil-1'de hastalığın saptanması, şiddet derecesinin veya bulaşıcılık hızının belirlenmesi konusundaki incelemeye alınmış çalışmalarda farklı makine öğrenmesi algoritmalarının kullanım oranları görülmektedir.



Şekil 1: İncelenen Covid-19 çalışmalarında en çok kullanılan makine öğrenmesi algoritmalarının kullanım oranları.

Bu makalede, Covid-19 ile ilişkili incelemelerde makine öğrenmesinden yararlanan çalışmalar başarımları açısından incelenmiştir. İkinci bölümde hastalığın tanılanmasında kullanılan yöntemler karşılaştırılmıştır. Üçüncü bölümde makine öğrenmesi yöntemleri ile hastalığın şiddetinin belirlenmesine, dördüncü bölümde ise hastalığın bulaşma hızının belirlenmesine yönelik çalışmalar incelenmiştir. Beşinci bölümde elde edilen veriler değerlendirilmiştir. Altıncı bölüm ise sonuç bölümüdür.

## 2. Makine Öğrenmesi ile Covid-19 Tanılama

Makine öğrenmesi yöntemleri, hastalığın otomatik tanılanması için yapılan uygulamalarda hem en kritik hem de en başarılı olarak değerlendirilmektedir. Klinik özellikler, kan değişkenleri ve ölçülen değerlerin, kabul edilen normal aralıklardan yaptıkları sapmalar, tanılama amaçlı algoritmalar için önemli birer veri kaynağıdır [14,15,16]. Algoritma başarımları değerlendirilirken kullanılan veri kümelerindeki özneliklerin seçimi ve uygulanan filtreleme işlemlerinin de dikkate alınması gerekmektedir.

[17] çalışması, makine öğrenme yöntemlerinden rastgele orman, yapay sinir ağı ve eniyileme algoritması kullanarak hastalığın teşhisine odaklanmıştır. Veri kümesi 08 Mart – 05 Nisan 2020 tarihleri arasında Sheba Tıp Merkezi'nden hastaneye başvuran 162 hastanın verileri ile oluşturulmuştur. Değişkenler cinsiyet, yaş, komorbidite, obezite, yoğun sigara içme, diyabet, astım, kalp hızı, solunum hızı, beyaz kan hücre sayısı, belirtilerin gözlemlenmesinden hastaneye yatışa kadar geçen süre ve oksijen doyumu (satürasyonu) olarak belirlenmiştir. Çalışmada kullanılan veri kümesi %70 eğitim ve %30 sinama verisi olmak üzere ikiye ayrılmıştır. Başarım değerlendirmesinde hata matrisi, eğri altında kalan alan ve beklenen ölüm skoru göz önüne alınmıştır. 12 fizyolojik parametre, yaş ve önceki sağlık durumu bilgisine dayalı beklenen ölüm ölçüsü (mulbsta-II): 0,92 bulunmuştur. En başarılı model olarak belirlenen rastgele orman ile alıcı işlem

karakteristik eğrisi (ROC Eğrisi) altında kalan alan: 0,93, bulma: 0,88, tutturma: 0,92 ve doğruluk: 0,92 olmuştur. Bu sonuçlar, makine öğrenmesi yöntemlerinin hastalık teşhisi açısından başarılı ve güvenilir bir araç olabileceğini göstermektedir.

[8] ile, hastalığın teşhisinde rastgele orman, gradyan artırma makinesi, lightboost, catboost ve topluluk öğrenme algoritmalarından torbalama algoritmasını kullanarak bir çalışma gerçekleştirilmiştir. Veri kümesi, Brezilya'nın São Paulo kentindeki Israelta Albert Einstein Hastanesine 28 Mart 2020 ile 3 Nisan 2020 tarihleri arasında gelenler arasında rastgele seçilen 600 kişiye ait kan değerleri ve ulamsal değişkenlerden oluşturulmuştur. Bu verileri kararlı kılmak için sentetik azınlık aşırı örnekleme tekniği (Smote) kullanılmıştır. Veri kümesi üzerinde kan gruplarının etkisini görmek için normallik sınaması (Shapiro) ve parametrik olmayan istatistiksel sınama (Whitney-U) yapılmıştır. Başarım değerlendirmesinde ise doğruluk, bulma, tutturma, kesinlik, f1 puanı ve uyum (kappa) istatistikleri incelenmiştir. Elde edilen en iyi sonuçlar; 0,87 doğruluk, 0,88 bulma, 0,89 kesinlik, 0,87 f1 ölçüsü ve 0,36 uyum ölçütü ile rastgele orman algoritmasında görülmüştür.

[18] çalışmasında, 279 hastadan alınan kan sınamaları kullanılmış ve eksik veri kümesini Zincirleme Denklemle Çoklu Atama (Multivariate Imputation by Chained Equations – MICE) yöntemiyle tamamlanmıştır. Değişken olarak yaş, lökosit, trombosit, CRP, AST, ALT, GGT, LDH, nötrofiller, lenfositler, monositler, eozinofiller ve bazofil değerleri seçilmiştir. Tanı için çeşitli makine öğrenme algoritmaları kullanılmıştır. Bu algoritmalar arasında karar ağacı, aşırı rastgele ağaçlar, k en yakın komşu, lojistik regresyon, naif bayes, rastgele orman, destek vektör makinesi ve sınıflandırma yöntemleri bulunmaktadır. Hastalığı en iyi tahmin eden algoritma olarak rastgele orman belirlenmiştir. Ayrıca, transaminaz değerinin 25'ten düşük olması, Covid-19 pozitifliğinin göstergesi olarak belirlenmiştir. Yapılan başarım değerlendirmesinde, ROC eğrisi altındaki alan 0,84 olarak bulunmuştur. Hata matrisinde 0,82 doğruluk ve 0,65 tutturma elde edilmiştir [18]. Bu sonuçlar, rastgele orman algoritmasının, belirlenen parametreler üzerinden yapılan teşhislerde etkili bir başarım sergilediğini göstermektedir.

[19] çalışması farklı olarak; Covid-19 teşhisinde rastgele orman algoritmasının yaygın olarak kullanıldığı bir alan olan makine öğrenmesi yöntemlerini incelemek yerine, gradyan artırma karar ağaçları üzerinde çalışmıştır. Bu yöntemin rastgele orman algoritmasından farkı, zayıf eğitim verilerini kullanırken bağımsız hareket etmek yerine sıralı olarak değerlendirme yapmasıdır. Kullanılan veri kümesi, İsrail Sağlık Bakanlığı'nın kayıtlarında bulunan Covid-19 sınaması yaptırmış kişilerin sınama sonuçlarından oluşmaktadır. Veri kümesinin %75'i eğitim, %25'i sınama verisi olarak ayrılmıştır. Veri kümesine ulamsal değişkenler, gözlenen hastalıklar ve hastalarla temas hikayeleri dahil edilmiştir. Verileri filtrelemede her tahmin sonucunu özellik değerinin katkısına bölerek modeller arasında ayırım yapmak için özellik alt uzayını kullanan SHAP (SHapley Additive exPlanations) modeli kullanılmıştır. 0,90 olan ROC eğrisi altındaki alan, değişkenler normallik sınaması ile filtreledikten sonra 0,95 olarak ölçülmüştür. Diğer başarım ölçüleri, bulma: 0,87 ve tutturma:

0,71 bulunmuştur. Özellik değişikliklerinin etkisini araştırmak için modelde bulma sabitlenmiştir. Bu durumda ROC eğrisi altındaki alan 0,66 ölçülmüştür.

Covid-19 tanılmasına odaklanan makine öğrenmesi çalışmalarında k-en yakın komşu, destek vektör makinesi, karar ağaçları, rastgele orman, doğrusal regresyon ve lojistik regresyon yöntemlerinin sıkça kullanıldığı gözlemlenmiştir.

Asıl veri kümesinden rastgele seçilen satır ve sütunlarla çok sayıda karar ağacı oluşturup bunların verdiği sonuçları bir arada değerlendiren bir yöntem olan rastgele orman algoritması diğerlerine göre üstün başarı göstermiştir [8,17,18,20,21].

Çizelge-1'de hastalığın tanılması için makine öğrenme yöntemleriyle yapılan çalışmaların karışıklık matrisinden hesaplanan başarı ölçütlerine göre karşılaştırılması gösterilmiştir. Verinin doğru ve yanlış sınıflandırılmış örnek sayısını içeren karışıklık matrisi (Toplam vaka sayısı= DP + DN + YP + YN) ele alınarak Çizelge 1, 3 ve 5' te çalışmalarda verilmeyen eksik değerler, verilen değerler kullanılarak hesaplanmış ve tamamlanmıştır (Çizelgelerde tamamlanan değerler kalın yazı tipi ile altı çizili olarak gösterilmiştir.). Farklı çalışmalarda ön plana çıkan başarım ölçütlerinin çeşitliliği, yöntemlerin karşılaştırılmasında potansiyel sorunları ortaya çıkarmaktadır. Verilerin işlenmiş olmasının tüm algoritmalarda başarıyı artırdığı gözlemlenmiştir. Bu nedenle, incelenen çalışmaların sadece filtrelenmiş ve dengelenmiş veriler kullanarak yapılan denemeleri dikkate alınmıştır.

İstatistik terimlerini hastalık saptama alanına özgü bir biçimde yorumlayacak olursak: Çizelgede yer alan ölçütlerden doğruluk, tüm tespitler içindeki doğru tespitlerin oranıdır (hasta veya sağlıklı). Bulma, tüm gerçek hastalar içindeki tespit edilebilmiş hastaların oranıdır. Kesinlik, hasta teşhisi konmuş bireyler arasındaki gerçek hastaların oranını göstermektedir. Tuturma ise gerçek sağlıklı bireyler arasındaki sağlıklı tescil edilmiş bireylerin oranı için kullanılan terimdir.

Algoritmaların ayırıcı gücünü alandan bağımsız olarak ölçmenin iyi bir yolu ROC eğrisi altında kalan alanı hesaplamaktır. ROC eğrisi (Receiver Operating Characteristic Curve) potansiyel olarak birbiriyle çelişen öncelikleri olan bulma ve tuturma arasındaki dengeyi, değişken bir eşik değeri boyunca modellemek amacıyla kullanılan bir fonksiyondur. Fonksiyon özgüllüğün tam tersi (tümleyen) olan yanlış pozitif oranı ile bulma eksenlerini içeren bir grafik olarak temsil edilir. Bu grafiğin eşik değerinin belirlenmesi konusunda önemli bir işlevi vardır, zira bulma ve tuturma kaygılarının dengelendiği optimum bir nokta tayin etme görevini görselleştirmektedir. Bunun dışında grafiğin altında kalan alan da AUC (Area Under the Curve) olarak bilinen başarım ölçütüdür. Çizelge 1'de bu değer ROC Eğrisi Altında Kalan Alan olarak gösterilmiştir ve teşhis koyan bir algoritmanın toplam ayırıcı gücünü (kaba kuvvetini) ölçmek konusunda kullanışlı bir değerdir. Eşik değeri belirlendiğinde ise bu gücün hangi yöne evrileceği belirlenmiş olmaktadır.

Bir algoritmanın toplam etkinliğini sayısal olarak modellemenin bir başka yolu, uygun bir  $\beta$  değeri için F- $\beta$

oranını hesaplamaktır. Yaygın olarak F puanı veya F1 puanı olarak bilinen ölçüt aslında mümkün olan  $F-\beta$  değerlerinden sadece bir tanesidir ve özel bir durumu ifade etmektedir. Bu ölçütün hesaplanmasında esas olan bulma ve kesinlik değerlerinin ağırlıklı harmonik ortalamasının bulunmasıdır. Beta değeri ise ağırlık çarpanıdır. Dolayısıyla F1 kesinlik ve bulma arasında tam bir denge kurarak ( $\beta=1$ ) her ikisine eşit önem atfetmektedir.

Kesinlik veya bulma farklı bir biçimde ağırlıklandırılmak istendiğinde farklı bir  $F-\beta$  ölçütü uygulanır. Kesinlik değeri, yanlış pozitiflerin ne kadar azaltılabileceğini göstermekte bulma ise yanlış negatiflerin ne kadar azaltılabildiğini ifade etmektedir. Yanlış pozitiflerin daha zararlı olduğu düşünüldüğünde  $\beta < 1$  olacak biçimde bir  $\beta$  parametresine ihtiyaç olacaktır. Tersine yanlış negatiflerin giderilmesinin daha öncelikli olduğunu düşünüyorsak  $\beta > 1$  olacak biçimde bir  $\beta$  belirlenmesi gerekecektir [22].

$\beta$  değerini doğru belirleyebilmek için, bir yanlış negatif vakasının, bir yanlış pozitif değerlendirmeye nazaran kaç kat daha kötü olduğunu subjektif ve kabaca da olsa saptamak ve bunu bir sayı olarak ifade etmek gerekmektedir. Hastalık teşhisi söz konusu olduğunda bu sayının genellikle 1 olmayacağı açıktır.

Hastalık tanılama ve tedavisi kendi önceliklerini dayatan alanlardır ve gereksiz uygulanan tedavilerin göze alınamayan ek riskler yaratabileceği hayali aykırı senaryoları bir kenara bırakırsak, yanlış negatif sınıflandırmaların, yanlış pozitiflere nazaran çok daha kötü bir ulama işaret ettiğini güvenli bir biçimde iddia edebiliriz. Sağlam sanılarak tedaviden mahrum bırakılmış bir hasta, hasta sanılarak yanlışlıkla tedavi edilmiş bir kişiye göre çok daha tehlikeli bir durumdadır. Hele Covid-19 gibi oldukça bulaşıcı hastalıklarda yanlış negatif olarak sınıflandırılan kişilerin hastalığı yaymaya devam edecekleri düşünülürse, durumun asimetrisi (ve dolayısıyla F1 ölçütünün neden iyi bir ölçüt olmadığı) daha iyi anlaşılacaktır.

Hastalık tanılmasında  $\beta$  oranını 1'den büyük seçmek gerekmektedir. Bu alanda bulma, kesinlik karşısında önceliklidir. Çalışmamızda Çizelge 1'de görüldüğü gibi F-2 ve F-10 ölçütleri dikkate alınmıştır (sırasıyla  $\beta=2$  ve  $\beta=10$ ).

İncelenen birçok çalışmada çok sayıda algoritmanın karşılaştırıldığı ve içlerinden birinin en başarılı yöntem olarak işaret edildiği görülmektedir. Çizelge-1'de yer alan değerler her çalışmadaki en başarılı bulunan algoritmaları yansıtmaktadır. Bu sıralamalarda birinci gelme sıklığını eğer bir ölçüt olarak alırsak bu konuda en başarılı olan yöntemin rastgele orman algoritması olduğu görülmektedir. Fakat tüm çalışmalar genelinde F-2 ve F-10 değerlerinin en iyi başarıyı gösteren yöntemlerin yapay sinir ağları ve aşırı gradyan artırma olduğunu görmekteyiz [28,29].

Sonuçta her iki bakış açısı da sadece birinci (veya ikinci) gelen algoritmaları dikkate aldığından algoritmaların tercih edilme sıklığından bağımsız bir biçimde değerlendirmek ve sağlıklı bir sonuca varmak zordur. Söz gelimi rastgele orman algoritmasının birçok karşılaştırmada birincilikle çıkması, kısmen en sık kullanılan algoritma olmasından kaynaklanabilir.

Bu etkiyi önlemek için tüm çalışmalarda yapılan denemelerden elde edilen sonuçlar, kullanılan algoritmalar üzerinden gruplandırılarak geometrik ortalamaları alınmıştır. Bu yöntem, çalışmalardaki tekrar edilemeyen aykırı değerlerin ve istatistiksel gürültünün etkisini en aza indirmek için tercih edilmiştir. Elde edilen sonuçlar, Çizelge-2'de görülmektedir. Tüm denemelerin geometrik ortalaması penceresinden bakıldığında, yapay sinir ağlarının en başarılı algoritma olduğu ikinciliğin ise derin öğrenme algoritmasına ait olduğu görülmektedir. Rastgele orman algoritmasının ise ilginç bir şekilde geride gözlemlenmektedir.

Çizelge-1 üzerinde yapılan değerlendirmede, F-2 ve F-10 değerlerinin en üst düzeyde, özellikle yapay sinir ağı ve aşırı gradyan artırma yöntemlerinde bulunduğu gözlemlenmektedir. Bu yöntemler, diğer başarımların değerlendirme metriklerine bakıldığında bulma değerlerinin diğer yöntemlere kıyasla daha yüksek olduğunu ortaya koymaktadır. Yüksek bulma, yanlış negatiflerin oranının ne kadar düşük olduğunu ifade eder.

Geometrik ortalamalara odaklanıldığında, yapay sinir ağının en üstün başarımlar sergileyen algoritma olduğu görülmektedir. Aşırı gradyan artırma yöntemi ise bu başarı seviyesinden biraz geride kalmaktadır. Tüm değerlendirmelerin göz önüne alındığı bir açıdan, 0,96 doğruluk ve 0,98 kesinlik değerleriyle yapay sinir ağının, hastalığın tanılmasında en etkili algoritma olduğu sonucuna varılmaktadır.

### 3. Hastalık Şiddetinin Belirlenmesi

Literatürdeki çalışmalarda ele alınan diğer bir konu hastalığın şiddetinin belirlenmesi olmuştur [33,34]. Hastalığın şiddetinin sınıflandırılması, hastaneye yatış gerekliliğini belirlemek açısından teşhis kadar kritiktir. Hangi hastaların hastaneye yatacağını doğru bir şekilde belirlemek, hastane kaynaklarının (hastanenin personeli, donanımı, hasta yatak kapasitesi, hastalık sürecinde ihtiyaç duyulacak malzemeler ve yoğun bakım ünitelerinin hazırlanması) etkili bir şekilde kullanılması ve gerekli hazırlıkların yapılması için hayati önem taşımaktadır. Bu nedenle hastalık şiddeti ile ilgili çalışmalar, hastaneye yatmayı gerektiren ve gerektirmeyen vakaların birbirinden ayrılması amacına yöneliktir.

Bazı çalışmalar hastalığın şiddet sınıflandırmasında hangi değişkenlerin (özniteliklerin) belirleyici olduğunu ve bu değişkenlerin önem derecesini belirlemeye odaklanmaktadır [5,6]. Şiddet sınıflandırması yapılırken genel olarak kan değerleri, kategorik değişkenler (cinsiyet, daha önce hastalık geçirip geçirmediği vb.), yaş ve belirtilerin görülme durumu gibi faktörlerin önemli olduğu görülmüştür. Ancak, seçilen öznitelikler konusunda bir standardın bulunmaması, algoritmalar arasında sağlıklı bir karşılaştırma yapmayı zorlaştırmaktadır.

Çizelge 1: Covid-19 tanısı için makine öğrenmesi algoritmalarıyla yapılan çalışmaların karşılaştırılması

Hastalığın Teşhisinde Makine Öğrenme Yöntemlerinin Başarım Değerlendirmesi										
Kaynaklar	İstatiksel Analiz	Önerilen Model	ROC Eğrisi Altında Kalan Alan	Başarım Değerlendirme Ölçütleri						
				Doğruluk	Bulma	Kesinlik	Tutturma	Veri Sayısı	F-2 Oranı	F-10 Oranı
[17]	Student t Mann-Whitney U $\chi^2$	Rastgele Orman	0,79	0,92	0,88	<u>0,87</u>	0,93	162	<u>0,88</u>	<u>0,88</u>
[18]	Zincirli Denklemle Çok Değişkenli Tahmin (MICE)	Rastgele Orman	0,84	0,82	0,92	<u>0,99</u>	0,65	279	<u>0,93</u>	<u>0,92</u>
[20]	Smote	Rastgele Orman	0,99	0,95	0,95	<u>0,97</u>	0,96	287	<u>0,95</u>	<u>0,95</u>
[21]	Smote	Rastgele Orman	0,84	0,95	0,93	0,93	<u>0,93</u>	5644	<u>0,93</u>	<u>0,93</u>
[23]	Karar Ağacı Açıklayıcı (DTX)	Rastgele Orman	0,86	0,88	0,66	<u>0,89</u>	0,91	608	<u>0,70</u>	<u>0,66</u>
[24]	Smote	Rastgele Orman	0,99	0,95	<u>0,95</u>	<u>0,97</u>	0,96	169	<u>0,95</u>	<u>0,95</u>
[25]	Smote Shapley Katkı Açıklamaları (SHAP)	Rastgele Orman	0,92	0,92	<u>0,71</u>	<u>0,90</u>	0,96	5644	<u>0,74</u>	<u>0,71</u>
[26]	Çapraz Doğrulama	Aşırı Gradyan Artırma	0,97	<u>0,97</u>	0,81	<u>0,53</u>	<u>0,97</u>	5333	<u>0,73</u>	<u>0,81</u>
[27]	Normalizasyon Algoritması	Gradyan Artırma Ağaçları	0,85	0,79	0,76	<u>0,74</u>	0,80	3356	<u>0,76</u>	<u>0,76</u>
[28]	Smote	Aşırı Gradyan Artırma	0,99	0,99	0,99	<u>0,99</u>	0,99	5644	<u>0,99</u>	<u>0,99</u>
[19]	Shapley Katkı Açıklamaları (SHAP)	Güçlendirmeye Dayalı Algoritmalar (LightGBM)	0,90	0,95	0,87	<u>0,79</u>	0,72	4769	<u>0,85</u>	<u>0,87</u>
[29]	Korelasyon katsayısı	Yapay Sinir Ağı	0,95	<u>0,96</u>	1	<u>0,98</u>	0,86	151	<u>1</u>	<u>1</u>
[30]	Özellik Seçimi	Lojistik Regresyon	0,97	0,90	0,87	<u>0,86</u>	0,95	912	<u>0,87</u>	<u>0,87</u>
[8]	Normalizasyon	Aşırı Gradyan Artırma	0,67	0,87	0,88	0,89	<u>0,88</u>	600	<u>0,88</u>	<u>0,88</u>
[31]	Çapraz doğrulama	Lojistik Regresyon	0,78	0,84	0,81	0,80	<u>0,82</u>	114	<u>0,81</u>	<u>0,81</u>
[32]	Çapraz doğrulama	Derin Öğrenme	0,90	0,92	0,94	<u>0,92</u>	<u>0,98</u>	1087	<u>0,94</u>	<u>0,94</u>

Çizelge 2: Tanılama makine öğrenme modellerinin başarım değerlerinin geometrik ortalamaları

Tanılama Algoritmalarının Ortalama Başarımları							
Model	ROC Eğrisi Altında Kalan Alan	Ortalama Değerler					
		Doğruluk	Bulma	Kesinlik	Tutturma	F-2 Oranı	F-10 Oranı
Yapay Sinir Ağı	0,95	0,96	1	0,98	0,86	1	0,99
Derin Öğrenme	0,90	0,92	0,94	0,92	0,98	0,94	0,94
Güçlendirmeye Dayalı Algoritmalar (LightGBM)	0,90	0,95	0,87	0,79	0,72	0,85	0,87
Aşırı Gradyan Artırma	0,86	0,90	0,86	0,70	0,91	0,83	0,86
Rastgele Orman	0,88	0,91	0,85	0,93	0,89	0,86	0,85
Lojistik Regresyon	0,78	0,87	0,84	0,83	0,88	0,84	0,84



geometrik ortalama alınarak birleştirilmiştir. Sonuçlar Çizelge-4'te görülmektedir. En başarılı yöntem rastgele orman algoritması görülmektedir. Bu çalışmada 5 farklı istatistiksel analiz ile bu başarı sağlanmıştır. Makine öğrenme yöntemleri farklı istatistiksel analiz yöntemleri ile ayrı ayrı değerlendirilmiş; ek makine öğrenme yöntemleri ile %100 başarıya ulaşmıştır. Fakat değerlendirme 337 hasta verisi üzerinde gerçekleştirilmiştir. Modelin başarısı farklı veri setleri üzerinde değerlendirilmelidir. 337 hastanın cinsiyet dağılımı da orantısızdır ve veri kümesinin değişiklikler yapılarak yeniden değerlendirilmesine ihtiyaç vardır.

Lojistik regresyon, destek vektör makinesi, rastgele orman, k en yakın komşu, güçlendirmeye dayalı algoritma gibi özellikleri keskin bir şekilde ayıran algoritmalarla yapılan çalışmaların hastalığın şiddetini saptama konusunda daha

başarılı sonuçlara ulaşmaları beklenen bir sonuçtur. Bununla birlikte farklı öznelikler kullanan çalışmalar üzerinden algoritmalar ile ilgili genellemeler yapmak doğru olmaz.

Öznelik seçimi yapılmadığında ise algoritmaların başarısı düşmektedir. İki sınıfın daha net ayırım yapmasını sağlayan parametreler ayarlanması ve hastalığın şiddetinde önemli değişkenlerin belirlenmesi ile model geliştirilir. Özellik seçiminin ilk kuralı ise dengeli bir veri kümesidir [30, 41].

Hastalığın şiddetini belirleme odaklı yapılan çalışmaların karşılaştırılmasına yönelik olarak, Çizelge-3'ün incelenmesi sonucunda en üstün başarıyı sergileyen yöntemlerin rastgele orman, yapay sinir ağı ve aşırı gradyan artırma olduğu göze çarpmaktadır. Ancak, rastgele orman algoritmasının tercih edilmesinde çok sayıda filtreleme işlemi uygulanmasının sonuçları üzerinde etkili olduğu düşünülmektedir.

Çizelge 3: Hastalık şiddetinin belirlenmesinde başarımların ölçütleri.

Hastalığın Şiddet Sınıflandırmasında Başarımların Karşılaştırması										
Kaynaklar	İstatistiksel Analiz	Önerilen Model	ROC Eğrisi Altında Kalan Alan	Başarımların Değerlendirme Ölçütleri						
				Doğruluk	Bulma	Keskinlik	Tutturma	Veri Sayısı	F-2 Oranı	F-10 Oranı
[9]	Student t	Destek Vektör Makinesi	0,91	<b>0,81</b>	0,83	<b>0,83</b>	0,79	137	<b>0,83</b>	<b>0,83</b>
[40]	Normalizasyon	Destek Vektör Makinesi	0,85	<b>0,77</b>	<b>0,68</b>	<b>0,78</b>	0,85	235	<b>0,70</b>	<b>0,68</b>
[41]	Karşılıklı bilgi Ki-kare sınamai	Destek Vektör Makinesi	0,96	0,92	0,92	<b>0,71</b>	0,91	10237	<b>0,87</b>	<b>0,92</b>
[33]	Çapraz Doğrulama	Aşırı Gradyan Artırma	0,91	<b>0,92</b>	0,96	<b>0,95</b>	0,42	7842	<b>0,96</b>	<b>0,96</b>
[42]	Çapraz Doğrulama	Aşırı Gradyan Artırma	0,90	<b>0,89</b>	0,56	<b>0,33</b>	0,92	5148	<b>0,49</b>	<b>0,56</b>
[43]	Çapraz Doğrulama Shapley Katkı Açıklamaları (SHAP)	Aşırı Gradyan Artırma	0,92	0,85	<b>0,85</b>	<b>0,65</b>	0,85	3927	<b>0,80</b>	<b>0,85</b>
[9]	Student t	Aşırı Gradyan Artırma	0,91	0,91	0,96	0,97	<b>0,86</b>	137	<b>0,96</b>	<b>0,96</b>
[34]	Smote	Aşırı Rastgele Ağaçlar	0,96	0,96	0,95	<b>0,96</b>	0,95	55676	<b>0,95</b>	<b>0,95</b>
[38]	R normalleştirme	Naif Bayes	0,90	0,95	0,86	<b>0,99</b>	0,75	422	<b>0,88</b>	<b>0,86</b>
[39]	Chi 2 Karşılıklı bilgi F sınamai Anova PCA UMAP	Rastgele Orman	1	1	<b>1</b>	<b>1</b>	1	337	<b>1</b>	<b>1</b>
[44]	Kaplan-Meier Yöntemi	Lojistik Regresyon	0,89	<b>0,86</b>	<b>0,82</b>	<b>0,47</b>	0,81	2307	<b>0,71</b>	<b>0,81</b>
[45]	Eş değişken seçimi Özellik çıkarımı	Lojistik Regresyon	0,881	<b>0,82</b>	0,84	<b>0,80</b>	0,79	183	<b>0,83</b>	<b>0,84</b>
[46]	Özellik seçimi	Yapay Sinir Ağı	0,98	<b>0,96</b>	<b>0,96</b>	<b>0,91</b>	<b>0,96</b>	166	<b>0,95</b>	<b>0,96</b>

Çizelge 4: Hastalık şiddetinin belirlenmesinde başarımların geometrik ortalamaları.

Şiddet Sınıflandırması Algoritmalarının Ortalama Başarımları							
Model	ROC Eğrisi Altında Kalan Alan	Başarımların Değerlendirme Ölçüleri					
		Doğruluk	Bulma	Kesinlik	Tutturma	F-2 Oranı	F-10 Oranı
Rastgele Orman	1	1	1	1	1	1	1
Yapay Sinir Ağı	0,98	0,96	0,96	0,91	0,96	0,95	0,96
Aşırı Rastgele Ağaçlar	0,96	0,96	0,95	0,96	0,95	0,95	0,95
Naif Bayes	0,90	0,95	0,86	0,99	0,75	0,88	0,86
Lojistik Regresyon	0,89	0,84	0,83	0,61	0,80	0,77	0,82
Aşırı Gradyan Artırma	0,91	0,89	0,81	0,67	0,73	0,78	0,81
Destek Vektör Makinesi	0,91	0,83	0,80	0,77	0,85	0,80	0,80

#### 4. Hastalığın Bulaşma Hızının Belirlenmesi

Literatürdeki çalışmalarda makine öğrenmesi kullanılarak hastalığın bulaşma hızını hesaplamak amaçlanmıştır [34,47]. Hastalığın bulaşma hızını etkileyen faktörler arasında hastanın etkileşimde bulunduğu kişi sayısı, hastalığın teşhisi için geçen süre ve hastalığın şiddeti gibi birçok unsurdan bahsedilebilir. Bu bağlamda, farklı çalışmaların bu amaç doğrultusunda nasıl bir yol izlediği detaylı bir şekilde incelenmiştir [48, 49].

[50] çalışması, bulaşma hızı hesaplanmasında sadece kan değerlerinden oluşan veri setlerinin yetersiz kaldığı saptanmıştır. Bulaşma hızını belirlemek için iki kriter belirlemişlerdir. Bunlar, hastalığa sahip kişilerin, ilk olarak Wuhan'da seyahat etme; ikinci olarak, diğer şehirlerde iletişim kurabilecekleri tanıdıkları buluşma olasılıklarıdır. Çalışmada bulaşma hızının belirlenmesinde bu verilerin yetersiz olduğu sonucuna ulaşılmıştır.

[51] çalışması ise Hong Kong Üniversitesi-Shenzhen Hastanesine ateş, solunum semptomları ve akciğer grafilerinde pulmoner infiltrat bulguları/şikayetleri ile başvuran iki hastada hastalığın bulaşma oranını belirlemek istemiştir. Bu ailenin diğer beş üyesi de 11 Ocak 2020 ile 15 Ocak 2020 tarihleri arasında incelenmiştir. Değerlendirilen özellikler, semptomların başlangıcı ile Wuhan'a varış arasındaki süre, hastaneye yatış ile semptomların başlangıcı arasındaki süre ve semptom ve bulgular (ateş, öksürük). Hastalarda incelenen değişkenler hastalık belirtileri ve kan değerleridir.

Hastalara ishal enfeksiyon sınaması, gerçek zamanlı RT-PCR sınaması yapılmış; solunumları, biyolojik istatistiksel çalışmalarla incelenmiştir. Hasta yakınlarının hastalığı nerden kapıldığı ve hasta kişinin kaç kişiye hastalığı bulaştırdığıyla ilgili bir veri elde edilememiştir. Bu yüzden bulaşıcılık hızı belirlenememiştir [51].

Hastalığın bulaşma hızını etkileyebilecek faktörler arasında hastalık şiddeti ve hastalığın saptanmasında belirleyici özelliklerin etkisi de bulunmaktadır. Bu değişkenlerin önem

derecesinin belirlenmesi, hastalık tanısının ve şiddetinin belirlenmesinde daha hızlı ve etkili bir yol açacaktır [48, 49]. Bu bağlamda, değişkenlerin önem derecelerini belirlemeye yönelik yapılan çalışmalar Çizelge-5'te karşılaştırılmıştır.

Kullanılan veri kümesinde önemli değişkenlerin belirlenebilmesi için modellerin başarımlarına bakıldığında lojistik regresyon ve yapay sinir ağı ROC eğrisi altında kalan alan değeri ile en başarılı algoritma görülmüştür. Geometrik ortalaması alınması da sonucu değiştirmemiştir (Çizelge-6). Yapay sinir ağı hastalığın teşhisi, şiddeti ve hastalıkta önemli değişkenlerin belirlenmesinde en başarılı yöntem olmuştur.

#### 5. İncelenen Çalışmaların Değerlendirilmesi

İncelenen çalışmalarda en yaygın kullanılan makine öğrenmesi yöntemlerinin sırasıyla rastgele orman, aşırı gradyan artırma makinesi, regresyon karar ağacı ve yapay sinir ağları, olduğu görülmektedir.

Hastalığın teşhisine yönelik yapılan çalışmalarda bazı başarımların verilmemesi gözlemlenmiştir. İncelenen çalışmalar arasında doğruluk, bulma, kesinlik ve tutturma olarak sayabileceğimiz dört temel başarımların tamamını veren bir çalışmaya rastlanmamıştır. Seçilen ölçütler arasında bir benzerlik de gözlemlenmemektedir. Örneğin [24] ve [25]'de tıbbi teşhis alanındaki en önemli ölçütlerden biri olan bulma değeri verilmemiştir. Bu ve bu durumda olan diğer çalışmalarda eksik sonuçlar elde edilen veriler kullanılarak hesaplanmış ve böylece tamamlanmış bir tablo üzerinde tüm çalışmaların başarımlarının gösterilmesi sağlanmıştır.



Çizelge 5: İdeal bir modelde önemli değişkenlerin belirlenmesi için oluşturulan modellerin kıyaslanması.

Önemli Değişkenlerin Belirlenmesinde Başarım Karşılaştırması										
Kaynaklar	İstatistiksel Analiz	Önerilen Model	ROC Eğrisi Altında Kalan Alan	Başarım Değerlendirme Ölçüleri						
				Doğruluk	Bulma	Kesinlik	Tutturma	Veri Sayısı	F-2 Oranı	F-10 Oranı
[49]	Smote	Yapay Sinir Ağı	1	0,95	0,95	<u>0,95</u>	0,95	580	<u>0,95</u>	<u>0,95</u>
[48]	Minimum fazlalık (mRMR)	Lojistik Regresyon	1	0,95	0,98	0,95	<u>0,91</u>	110	<u>0,97</u>	<u>0,98</u>
[47]	Student-t Ki-kare sına	Lojistik Regresyon	0,93	0,95	<u>0,95</u>	<u>0,82</u>	<u>0,95</u>	536	<u>0,92</u>	<u>0,95</u>
[34]	Ki-kare sına Çapraz doğrulama	Karar Ağacı	0,92	0,89	0,84	0,94	<u>0,89</u>	55676	<u>0,86</u>	<u>0,84</u>

Çizelge 6: Önemli değişkenlerin belirlenmesi için oluşturulan modellerin geometrik ortalaması hesaplamıştır.

Değişken Belirleme Yöntemlerinin Ortalama Başarımları							
Model	ROC Eğrisi Altında Kalan Alan	Başarım Değerlendirme Ölçüleri					
		Doğruluk	Bulma	Kesinlik	Tutturma	F-2 Oranı	F-10 Oranı
Lojistik Regresyon	0,96	0,95	0,96	0,88	0,93	0,94	0,96
Yapay Sinir Ağı	1	0,95	0,95	0,95	0,95	0,95	0,95
Karar Ağacı	0,92	0,89	0,84	0,94	0,89	0,86	0,84

Yapay zekânın tıbbi teşhis amaçlı kullanımında bulma ölçütü özellikle önemli olduğu görülmektedir. Duyarlılığın ağırlığının istenildiği gibi artırılabilmesi F-β değerlerinin hesaplanması yöntemlerinin başarısını ölçmek için tercih edilmesi gereken gerçekçi bir kıstas sunmaktadır. Bir algoritmanın genel olarak başarılı gibi görünmesi, özellikle verilerin dengesiz olduğu tıp gibi alanlarda yanıltıcı olabilir. Beta değerinin 1'den büyük olduğu bir F-β ölçütü daha uygundur, çünkü yanlış negatif vakalar genellikle yanlış pozitif vakalara göre daha tehlikelidir. Bu iki hususa simetrik yaklaşan ölçütler, yanıltıcı olabilir.

Bu çalışmada F-2 ve F-10 değerleri hesaplanmıştır. Teşhis çalışmaları için yapılan karşılaştırmada ilk bakışta fark edilmesi zor olan bir başarı sıralaması ortaya çıkmış ve yapay sinir ağları diğer yöntemlerin önüne geçmiştir. ROC eğrisi altında kalan alan açısından yüksek puan aldığı halde F-2 skorundan düşük not alan çalışmaların varlığı, eşik değeri belirleme işlemlerinde hata yapılmış olabileceğini düşündürmektedir.

Verilerin ön işleme tabi tutulması (istatistiksel analiz/filtreleme işlemleri) tüm çalışmalarda ROC eğrisi altında kalan alan değerini kayda değer biçimde artırmıştır. Bu durum filtreleme ve analiz işlemlerinin yararını göstermektedir [51].

Tüm değerlendirmelerde veri kümesinin büyüklüğü ve içeriği

çok büyük değişkenlik göstermektedir. Çok sayıda değişkene rağmen algoritmalara yönelik genel bir yargıya varabilmek amacıyla aynı algoritmayı kullanan tüm denemelerin sonuçlarının geometrik ortalamasını içeren ayrı bir karşılaştırma yapılmıştır. Bu karşılaştırmada yapay sinir ağları en ön sıradaki yerini korumuş, rastgele orman algoritması ise biraz daha gerilere düşmüştür (Çizelge 2). Genelleme yapmanın çok daha zor olduğu hastalık şiddetinin ve bulaşıcılığının belirlenmesi alanlarında da aynı karşılaştırma yöntemi tekrar edilmiştir (Çizelge 4-6).

Makine öğrenmesi yöntemleri ile hastalığın şiddetinin belirlenmesi çalışmalarının değerlendirmede iki temel sorun bulunmaktadır". Bunlardan birincisi çalışmalarda seçilen (veya elde edilebilen) öznelitliklerin birbirinden farklı olması; ikincisi ise, yapılan tahminlerin doğruluğunu sına etmek için kullanılacak güvenilir bir çıktı verisinin olmamasıdır. Sonuçların doğruluğu, ölen veya yoğun bakıma yatan hastalar üzerinden değerlendirilmiştir. Gerçekte kimin ne kadar risk yaşadığını doğrudan gösteren bir ölçüm mümkün değildir.

Girdi verisi olarak ise beklenen ölüm oranı ölçütü [6]'da lenfopeni, bakteriyel koenfeksiyon, sigara içme, hipertansiyon ve yaş değişkenlerini kullanarak yapılan çalışmada başarıyla kullanılmıştır. İncelenen hastalardan ölen ve yoğun bakıma yatan hastaların değerlerindeki aralık

sapması ve ortak özellikler hastalığın şiddeti ile ilişkili bulunmuştur. Yaş, büyüme hormonu (GHS), CD3, çalışma durumu, diyabet, kanser varlığı, protein seviyesi, sitokin, oksijen seviyesi ve lenfosit sayısı hastalık şiddetinde belirleyici rol oynamıştır [15,30,41,44,45].

Metin etiketleme için kullanılan naif bayes ve özelliklerin birbirinden daha keskin ayrıldığı algoritmalar, yani lojistik regresyon, destek vektör makinesi, rastgele orman, k en yakın komşu ve yetkilendirme tabanlı algoritma, hastalığın şiddetini belirlemede oldukça başarılı bulunmuştur. Ölçütlere bakıldığında ise rastgele orman yönteminin 337 veri içeren bir çalışmada istatistiksel analiz işlemleri ile birlikte %100 doğruluğa ulaşması dikkat çekicidir [39]. Bunun yanında yapay sinir ağı yaklaşımı burada da başarısını sürdürmektedir [46].

Hastalığın bulaşma hızının belirlenmesi sürecinde, sinama öncesi ve sonrasında hastanın insanlarla olan etkileşiminin veri kümesinde temsil edilmesi gerektiği değerlendirilmektedir. Bu insani veriler olmadan, bulaşma oranının başarısı, ölçülemeyecektir [51].

Tüm deneylere ilişkin başarımların verilerinin geometrik ortalamalarına bakıldığında hastalığın teşhisinde, şiddetinin belirlenmesinde ve değişkenlerin önem derecesinde yapay sinir ağı yönteminin tutarlı bir biçimde başarılı olduğu görülmektedir. Hastalık şiddetinin belirlenmesi alanında rastgele orman algoritmasının önemli bir çıkış yaptığı söylenebilir. Bununla birlikte bu sonuç tek bir çalışmanın ölçütlerinden kaynaklanmaktadır.

## 6. Sonuç

Bu çalışmada literatürde yer alan makine öğrenmesi algoritmaları ile yapılan Covid-19 çalışmaları hem sonuçları açısından hem de değerlendirme yöntemleri açısından sistematik bir biçimde incelenmiştir. İnceleme Covid-19'un tanısı, hastalığın şiddeti ve bulaşıcılık hızının hesaplanması konularını kapsamaktadır.

Çalışmalar veri büyüklüğü, öznitelik seçimi ve hasta popülasyonu açısından büyük farklılıklar göstermektedir. Ayrıca karşılaştırmaların dayandırıldığı başarı ölçütleri ile ilgili standartlar oluşturulmamıştır. Değişkenlerin fazlalığı, algoritmalar üzerinde güvenilir bir karşılaştırma yapmak konusunda bir engeldir.

Bu çalışmada karşılaştırmaları anlamlı kılmak amacıyla çeşitli yöntemler kullanılmıştır. Eksik olan başarımlar ölçütleri hesaplanmış ve alana özgü nedenlerle bu ölçütlerin tümünden daha önemli olan F-β değerleri (F-2 ve F-10 olmak üzere) hesaplanmıştır. Ayrıca aynı algoritmayı kullanan denemeler gruplanarak başarımlar ölçütlerinin geometrik ortalaması alınmıştır. Tüm bu hesaplamalar sonucunda yapay sinir ağı tabanlı yaklaşımların (derin öğrenme dahil edilebilir) hastalığın kan tahlilinden teşhisi konusunda en başarılı yöntemler olduğu görülmektedir.

Covid-19 şiddetinin belirlenmesinde farklı özniteliklerin kullanılması nedeniyle yöntemler arası karşılaştırma sonuçları yeterince güvenilir olmasa da makine öğrenmesi yöntemlerinin genel olarak hastalık şiddetinin belirlenmesinde de teşhis alanında olduğu kadar etkili

oldukları gözlemlenmektedir.

Hastalığın bulaşma hızının hesaplanmasında ise elimizde güvenilir bir yöntem yoktur. Eldeki yöntemlerle bazı etkili değişkenler belirlenebilmektedir; fakat bulaşma hızının tespiti için hem bulaşma olaylarının detaylı bir dökümü hem de fiziksel temas sıklığının temsil edilmesine ilişkin insan ilişkilerinin yüksek sadakatli bir modeline ihtiyaç vardır. Bu derece detaylı verilerin elde edilmesi günümüzdeki sosyal, kültürel ve yasal çerçevelerde çok zordur.

Çalışmanın odak noktası, başarımlar değerlerinin analizi üzerinden salgın hastalıklarda hangi metriklerin önemli olduğunu belirlemek ve Covid-19 gibi günümüzde mutasyona uğrayarak ortaya çıkabilecek yeni hastalıkların değerlendirmesini sadece başarımlar metrikleri üzerinden değil, aynı zamanda yanlış kararların performansa öznel bir etki yapmadığını vurgulamaktadır. Bu bağlamda, makine öğrenmesi yöntemleriyle incelenen bir hastalığın değerlendirilmesinde, farklı bakış açılarının da göz ardı edilmemesi önemlidir. Yani, başarımın ötesinde, kararların objektif bir şekilde değerlendirilmesi gerekliliği vurgulanarak, salgın hastalıkların etkili bir şekilde analiz edilebilmesi için çeşitli perspektiflere odaklanılmalıdır.

Sonuç olarak eksiksiz ve dengeli bir veri kümesinin (eğitim ve sinama kümesi) en doğru sonuçları vereceği araştırmalarla gösterilmiştir. Ancak gerçek hayat veri kümelerinin dengesiz olması, daha iddialı bir model elde etmek için daha fazla sayıda örneklem gerektirmektedir. Modellerin başarımını tam bir doğrulukla değerlendirebilmek için aynı veri kümesi üzerinde yöntemlerin tamamını denemek ve daha sonra bu deneyi farklı veri setleri ile tekrarlamak gerekmektedir.

Tüm eksikliklere ve belirsizliklere rağmen, makine öğrenmesi yöntemleri hastalığın teşhisi ve takibi konusunda etkili birer yardımcıdır. Bu teknolojileri kullanan yazılımlar Covid-19 ve benzer salgınlarda risk ve maliyetlerin azaltılması ve kaynakların daha etkili yönetilmesi gibi yararlarıyla hekimlerin güvenilir destekçileri olacaktır.

## Kaynakça

- [1] Lewnard JA, Lo NC. The scientific and ethical basis for social-distancing interventions against COVID-19. *Lancet Infect Dis.* 2020 Jun;20(6):631-633.
- [2] Koo JR, Cook AR, Park M, Sun Y, Sun H, Lim JT, Tam C, Dickens BL. Interventions to mitigate the early spread of SARS-CoV-2 in Singapore: a modeling study. *Lancet Infect Dis.* 3-2020 Jun;20(6):678-688.
- [3] Pandemi ilan edilmesi-Dünya Sağlık Örgütü. DSÖ Genel Direktörü'nün 11 Şubat 2020 tarihinde 2019-nCoV medya brifinginde yaptığı açıklamalar. <https://www.who.int/dg/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-ncov-on-11-february-2020> (2020).
- [4] Rasheed J, Jamil A, Hameed AA, Al-Turjman F, Rasheed A. COVID-19 in the Age of Artificial Intelligence: A Comprehensive Review. *Interdiscip Sci Comput Life Sci.* Published online April 22, 2021.
- [5] Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, Liu L, Shan H, Lei CL, Hui DSC, Du B, Li LJ, Zeng G, Yuen KY, Chen RC, Tang CL, Wang T, Chen PY, Xiang J, Li SY, Wang JL, Liang ZJ, Peng YX, Wei L, Liu Y, Hu YH, Peng P, Wang JM,

- Liu JY, Chen Z, Li G, Zheng ZJ, Qiu SQ, Luo J, Ye CJ, Zhu SY, Zhong NS, China Medical Treatment Expert Group for Covid-19. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med*. 2020 Apr 30;382(18):1708-1720.
- [6] Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, Qiu Y, Wang J, Liu Y, Wei Y, Xia J, Yu T, Zhang X, Zhang L. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020 Feb 15;395(10223):507-513.
- [7] Rodriguez-Morales A, Cardona-Ospina J, Gutierrez-Ocampo E, Villamizar-Pe~na R, Holguin-Rivera Y, Escalera-Antezana J, Alvarado-Arnez L, Bonilla-Aldana D, Franco-Paredes C, Henao-Martinez A. Clinical, laboratory and imaging features of COVID-19: a systematic review and meta-analysis. *Trav Med Infect Dis* 2020;34: 101623.
- [8] Demirarslan M., Suner A. Rutin kan sınamalarıyla Covid-19 tanı tahmininde makine öğrenmesi yöntemleriyle mobil uygulama geliştirilmesi. *Ege Tıp Dergisi*. 2021, 60(4): 384-393.
- [9] Yao H, Zhang N, Zhang R, Duan M, Xie T, Pan J, Peng E, Huang J, Zhang Y, Xu X, Xu H, Zhou F, Wang G. Severity Detection for the Coronavirus Disease 2019 (COVID-19) Patients Using a Machine Learning Model Based on the Blood and Urine Sinamas. *Front Cell Dev Biol*. 2020 Jul 31;8:683.
- [10] Booth, A.L, Abels, E, McCaffrey, P, Development of a prognostic model for mortality in COVID-19 infection using machine learning. *Mod. Pathol*. 2021, 34, 522–531.
- [11] Özen N. S., Saraç S. and Koyuncu M., "COVID-19 Vakalarının Makine Öğrenmesi Algoritmaları ile Tahmini: Amerika Birleşik Devletleri Örneği", *Avrupa Bilim ve Teknoloji Dergisi*, no. 22, pp. 134-139, Jan. 2021.
- [12] WHO Gender and COVID-19, World Health Organization: Geneva, Switzerland, 2020.
- [13] Malom, Z., Rahman, M.M.S., Nasrin, S., Taha, T.M., Asari, V.K. COVID\_MTNNet: COVID-19 detection with multi-task deep learning approaches. *arXiv* 2020, arXiv:2004.03747.
- [14] Michael J. Loeffelholz & Yi-Wei Tang (2020) Laboratory diagnosis of emerging human coronavirus infections – the state of the art, *Emerging Microbes & Infections*, 9:1, 747-756.
- [15] Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF, Tan W, China Novel Coronavirus Investigating and Research Team. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med*. 2020 Feb 20;382(8):727-733.
- [16] Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat Microbiol* 5, 536–544 (2020).
- [17] Assaf D, Gutman Y, Neuman Y, Segal G, Amit S, Gefen-Halevi S, Shilo N, Epstein A, Mor-Cohen R, Biber A, Rahav G, Levy I, Tirosch A. Utilization of machine-learning models to accurately predict the risk for critical COVID-19. *Intern Emerg Med*. 2020 Nov;15(8):1435-1443.
- [18] Brinati D, Campaigner A, Ferrari D, Locatelli M, Banfi G, Cabitza F. Detection of COVID-19 Infection from Routine Blood Exams with Machine Learning: A Feasibility Study. *J Med Syst*. 2020 Jul 1;44(8):135.
- [19] Zoabi Y, Deri-Rozov S, Shomron N. Machine learning-based prediction of COVID-19 diagnosis based on symptoms. *NPJ Digit Med*. 2021 Jan 4;4(1):3.
- [20] Sumayh S. Aljameel, Irfan Ullah Khan, Nida Aslam, Malak Aljabri, Eman S. Alsulmi, "Machine Learning-Based Model to Predict the Disease Severity and Outcome in COVID-19 Patients", *Scientific Programming*, vol. 2021, Article ID 5587188, 10 pages, 2021.
- [21] Freitas Barbosa, V.A., Gomes, J.C., de Santana, M.A. et al. Heg.IA: an intelligent system to support diagnosis of Covid-19 based on blood sinamas. *Res. Biomed. Eng*. 38, 99–116 (2022).
- [22] Scikit-Learn, F-Bta Skoru, Erişim adresi: [https://scikit-learn.org/stable/modules/generated/sklearn.metrics.fbeta\\_score.html](https://scikit-learn.org/stable/modules/generated/sklearn.metrics.fbeta_score.html), Erişim Tarihi:02.02.2023.
- [23] Alves MA, Castro GZ, Oliveira BAS, Ferreira LA, Ramirez JA, Silva R, Guimaraes FG. Explaining machine learning based diagnosis of COVID-19 from routine blood sinamas with decision trees and criteria graphs. *Comput Biol Med*. 2021 May;132:104335.
- [24] Jiangpeng Wu, Pengyi Zhang, Liting Zhang, Wenbo Meng, Junfeng Li, Chongxiang Tong, Yonghong Li, Jing Cai, Zengwei Yang, Jinhong Zhu, Meie Zhao, Huirong Huang, Xiaodong Xie, Shuyan Li, Rapid and accurate identification of COVID-19 infection through machine learning based on clinical available blood sinama results *medRxiv* 2020.04.02.20051136.
- [25] Chadaga K, Chakraborty C, Prabhu S, Umakanth S, Bhat V, Sampathila N. Clinical and Laboratory Approach to Diagnose COVID-19 Using Machine Learning. *Interdiscip Sci*. 2022 Jun;14(2):452-470.
- [26] Kukar M, Gunčar G, Vovko T, Podnar S, Černelč P, Brvar M, Zalaznik M, Notar M, Moškon S, Notar M (2021) COVID-19 diagnosis by routine blood sinamas using machine learning. *Sci Rep* 11(1):1–9.
- [27] Yang HS, Hou Y, Vasovic LV, Steel PAD, Chadburn A, Racine-Brzostek SE, Velu P, Cushing MM, Loda M, Kaushal R, Zhao Z, Wang F. Routine Laboratory Blood Sinamas Predict SARS-CoV-2 Infection Using Machine Learning. *Clin Chem*. 2020 Nov 1;66(11):1396-1404.
- [28] AlJame M, Ahmad I, Intiaz A, Mohammed A. Ensemble learning model for diagnosing COVID-19 from routine blood sinamas. *Inform Med Unlocked*. 2020;21:100449
- [29] Kang J, Chen T, Luo H, Luo Y, Du G, Jiming-Yang M. Machine learning predictive model for severe COVID-19. *Infect Genet Evol*. 2021 Jun;90:104737.
- [30] Sun N, Yang Y, Tang L, Li Z, Dai Y, Xu W, et al. (2021) A Prediction Model Based on Machine Learning for Diagnosing the Early COVID-19 patients. *J Antivir Antiretrovir*. S18:002.
- [31] Arpacı I, Huang S, Al-Emran M, Al-Kabi MN, Peng M (2021) Predicting the COVID-19 infection with fourteen clinical features using machine learning classification algorithms. *Multimedia Tools Appl* 80(8):11943–11957.
- [32] Zhang RK, Xiao Q, Zhu SL, Lin HY, Tang M. Using different machine learning models to classify patients into mild and severe cases of COVID-19 based on multivariate blood sinamaing. *J Med Virol*. 2022 Jan;94(1):357-365.
- [33] Plante TB, Blau AM, Berg AN, Weinberg AS, Jun IC, Tapson VF, Kanigan TS, Adib AB (2020) Development and external validation of a machine learning tool to rule out COVID-19 among adults in the emergency department using routine blood sinamas: a large, multicenter, real-world study. *J Med Internet Res* 22(12): e24048.
- [34] dos Santos Santana ÍV, da Silveira AC, Sobrinho Á, Silva LC, da Silva LD, Santos DF, Gurjão EC, Perkusich A (2021) Classification models for COVID-19 sinama prioritization in Brazil: machine learning approach. *J Med Internet Res* 23(4): e27293.
- [35] Luo C. L., Rong Y., Chen H., Zhang W., Wu L., Wei D., et al. (2019). A logistic regression model for noninvasive prediction of AFP-negative hepatocellular carcinoma. *Technol. Cancer Res. Treat*. 18:1533033819846632.

- [36] Heijnen B. J., Bohringer S., Speyer R. (2020). Prediction of aspiration in dysphagia using logistic regression: oral intake and self-evaluation. *Eur. Arch. Otorhinolaryngol.* 277 197–205.
- [37] Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF, Tan W, China Novel Coronavirus Investigating and Research Team. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med.* 2020 Feb 20,382(8):727-733.
- [38] Zhang RK, Xiao Q, Zhu SL, Lin HY, Tang M. Using different machine learning models to classify patients into mild and severe cases of COVID-19 based on multivariate blood snamaing. *J Med Virol.* 2022 Jan,94(1):357-365.
- [39] Laatifi M, Douzi S, Bouklouz A, Ezzine H, Jaafari J, Zaid Y, El Ouahidi B, Naciri M. Machine learning approaches in Covid-19 severity risk prediction in Morocco. *J Big Data.* 2022,9(1):5.
- [40] A.F.de M. Batista, J.L. Miraglia, T.H.R. Donato, A.D.P. Chia vegatto Filho, COVID-19 diagnosis prediction in emergency care patients: a machine learning approach, 2020.
- [41] An C, Lim H, Kim DW, Chang JH, Choi YJ, Kim SW. Machine learning prediction for mortality of patients diagnosed with COVID-19: a nationwide Korean cohort study. *Sci Rep.* 2020 Oct 30,10(1):18716.
- [42] Du R, Tsougenis ED, Ho JWK, Chan JKY, Chiu KWH, Fang BXH, Ng MY, Leung ST, Lo CSY, Wong HF, Lam HS, Chiu LJ, So TY, Wong KT, Wong YCI, Yu K, Yeung YC, Chik T, Pang JWK, Wai AK, Kuo MD, Lam TPW, Khong PL, Cheung NT, Vardhanabhuti V. Machine learning application for the prediction of SARS-CoV-2 infection using blood snamas and chest radiograph. *Sci Rep.* 2021 Jul 9,11(1):14250
- [43] Bertsimas D, Lukin G, Mingardi L, Nohadani O, Orfanoudaki A, Stellato B, Wiberg H, Gonzalez-Garcia S, Parra-Calderón CL, Robinson K, Schneider M, Stein B, Estirado A, A Beccara L, Canino R, Dal Bello M, Pezzetti F, Pan A, Hellenic COVID-19 Study Group. COVID-19 mortality risk assessment: An international multi-center study. *PLoS One.* 2020 Dec 9,15(12): e0243262.
- [44] Sánchez-Montañés M, Rodríguez-Belenguer P, Serrano-López AJ, Soria-Olivas E, Alakhdar-Mohmara Y. Machine Learning for Mortality Analysis in Patients with COVID-19. *Int J Environ Res Public Health.* 2020 Nov 12,17(22):8386.
- [45] Hu C, Liu Z, Jiang Y, Shi O, Zhang X, Xu K, Suo C, Wang Q, Song Y, Yu K, Mao X, Wu X, Wu M, Shi T, Jiang W, Mu L, Tully DC, Xu L, Jin L, Li S, Tao X, Zhang T, Chen X. Early prediction of mortality risk among patients with severe COVID-19, using machine learning. *Int J Epidemiol.* 2021 Jan 23,49(6):1918-1929.
- [46] Kocadagli O, Baygul A, Gokmen N, Incir S, Aktan C. Clinical prognosis evaluation of COVID-19 patients: An interpretable hybrid machine learning approach. *Curr Res Transl Med.* 2022 Jan,70(1):103319.
- [47] Gangloff C, Rafi S, Bouzillé G, Soulat L, Cuggia M (2021) Machine learning is the key to diagnose COVID-19: a proof-of-concept study. *Sci Rep* 11(1):1–1.
- [48] Wu G, Zhou S, Wang Y, Lv W, Wang S, Wang T, Li X. A prediction model of outcome of SARS-CoV-2 pneumonia based on laboratory findings. *Sci Rep.* 2020 Aug 20,10(1):14042.
- [49] Göreke V, Sarı V, Kockanat S. A novel classifier architecture based on deep neural network for COVID-19 detection using laboratory findings. *Appl Soft Comput.* 2021 Jul,106:107329.
- [50] Sanche, S., Lin, Y., Xu, C., Romero-Severson, E., Hengartner, N., & Ke, R. (2020). Şiddetli Akut Solunum Sendromunun Yüksek Bulaşıcılığı ve Hızlı Yayılması Koronavirüs 2. Ortaya Çıkan Bulaşıcı Hastalıklar, 26(7), 1470-1477.
- [51] Chan JF, Yuan S, Kok KH, To KK, Chu H, Yang J, Xing F, Liu J, Yip CC, Poon RW, Tsoi HW, Lo SK, Chan KH, Poon VK, Chan WM, Ip JD, Cai JP, Cheng VC, Chen H, Hui CK, Yuen KY. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet.* 2020 Feb 15,395(10223):514-523.