

PAPER DETAILS

TITLE: Saglik Hizmetlerinde Güncel Makine Öğrenmesi Algoritmaları

AUTHORS: Lütviye Özge POLATLI,Melis Almula KARADAYI

PAGES: 117-143

ORIGINAL PDF URL: <https://dergipark.org.tr/tr/download/article-file/2432612>

Sağlık Hizmetlerinde Güncel Makine Öğrenmesi Algoritmaları

Lütfiye Özge POLATLI¹, Melis Almula KARADAYI²

Öz

Son yıllarda araştırmacılar tarafından makine öğrenmesi algoritmalarını kullanarak sağlık süreçlerinin iyileştirilmesi konusu büyük bir trend haline gelmiştir. Makine öğrenmesi, sağlık hizmetlerinde kaliteyi yükseltmek, hastalık yayılımlarını önlemek, hastalıkları erken teşhis etmek, hastane operasyon maliyetlerini azaltmak, hükümete sağlık hizmetleri politikalarında yardımcı olmak ve sağlık hizmetinin verimliliğini artırmak için kullanılan popüler ve etkili bir yöntem haline gelmiştir. Bu derleme çalışmasında, sağlık alanında gerçekleştirilen makine öğrenmesi çalışmaları özetlenmiş ve sınıflandırılmıştır. Özellikle halk sağlığını tehdit eden ve dünyada ölüm nedenleri listesinde ilk sıralarda yer alan, bulaşıcı olmayan hastalık çalışmalarına odaklanılmıştır. Ayrıca dünyanın en büyük ölümcül hastalıklar listesinde yer alan ve son yıllarda halk sağlığı için acil durum ilan edilen COVID-19 hastalığına da yer verilmiştir. Bu çalışmanın amacı, sağlık alanında çalışma yapan araştırmacılara uygun algoritmalarını seçmesinde yardımcı olmaktır. Derleme çalışmasının sonucunda sağlık hizmetlerinde en iyi performans gösteren sınıflandırma algoritması ortalama %100 doğruluk başarıyla Decision Tree (DT), Random Forest (RF), Gaussian Naive Bayes (GNB) olmuştur.

Anahtar kelimeler: Makine öğrenmesi algoritmaları, sağlık hizmetleri, bulaşıcı hastalıklar, bulaşıcı olmayan hastalıklar

1. Arş. Gör., Altınbaş Üniversitesi, Veri Analitiği Bölümü, lutfiye.polatli@std.medipol.edu.tr, <https://orcid.org/0000-0001-7118-5370>
2. Dr. Öğretim Üyesi, İstanbul Medipol Üniversitesi, Endüstri Mühendisliği Bölümü, makaradayi@medipol.edu.tr, <https://orcid.org/0000-0002-6959-9168>

Gönderim Tarihi : 17.05.2022

Kabul Tarihi : 20.07.2022

Atıfta Bulunmak İçin:

Polatlı, L.Ö. ve Karadayı, M.A. (2022). Sağlık Hizmetlerinde Güncel Makine Öğrenmesi Algoritmaları, *Eurasian Journal Of Health Technology Assessment*,6(2):117-143.

A Review on Machine Learning Algorithms in Healthcare

Lütfiye Özge POLATLI¹, Melis Almula KARADAYI²

Abstract

In recent years, the issue of improving health processes by using machine learning algorithms by researchers has become a big trend. Machine learning has become a popular and effective method used to improve the quality of healthcare services, prevent disease outbreaks, diagnose diseases early, reduce hospital operating costs, assist the government in healthcare policies, and increase healthcare efficiency. In this review, machine learning studies carried out in the field of health are summarized and classified. In particular, the focus is on studies of non-communicable diseases, which threaten public health and are at the top of the list of causes of death in the world. In addition, the COVID-19 disease, which is on the list of the world's largest deadly diseases and has been declared a public health emergency in recent years, is also included. The purpose of this study is to assist researchers working in the field of health in choosing appropriate algorithms. As a result of the compilation studies, the best performing classification algorithm in healthcare services was Decision Tree(DT), Random Forest (RF), Gaussian Naive Bayes (GNB) with an average accuracy of 100%.

Keywords: Machine learning algorithms, healthcare, communicable diseases, non-communicable diseases

3. Research Assistant, Altınbaş University, Department of Data Analytics, lutfiye.polatli@std.medipol.edu.tr , <https://orcid.org/0000-0001-7118-5370>
4. Assist Prof., Istanbul Medipol University, Department of Industrial Engineering, makaradayi@medipol.edu.tr, <https://orcid.org/0000-0002-6959-9168>

Received : 17.05.2022
Accepted : 20.07.2022

Cite This Paper:

Polatli, L.Ö. ve Karadayı, M.A. (2022). A Review on Machine Learning Algorithms in Healthcare, Eurasian Journal Of Health Technology Assessment,6(2):117-143.

1. Giriş

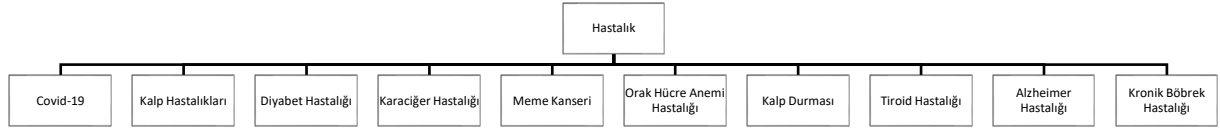
Son yıllarda gelişen teknolojiyle beraber sağlık sektöründe de önemli değişimler gözlemlenmektedir. Teknolojik ilerleme, mevcut sağlık teknolojilerini geliştirmiş ve yeni sağlık teknolojilerini de sektöre kazandırmıştır (Karadayı vd, 2020). Veri setleri ile çalışılan ve olaylara gerçekçi bir yaklaşım sunan makine öğrenmesi algoritmaları, araştırmacılar tarafından çok fazla ilgi görmeye başlamıştır.

Sağlık hizmeti sunucuları, iyi ve kabul edilebilir bir bakım kalitesine yönelik artan talebe yanıt vermek zorundadır (Kalender vd., 2020). Sağlık kurumlarında yapılan süreç iyileştirmeleri sayesinde kurumların maliyetlerinde düşüş sağlanırken hasta memnuniyeti ve hizmet kalitesi artırılmaktadır. Bu doğrultuda, literatürde sağlık süreçlerinin iyileştirilmesi konusuna odaklanan birçok uluslararası çalışma mevcuttur (Karadayı vd., 2019). Hastalıkları sınıflandırma, öngörme, teşhis etme gibi çalışmalar sağlık sektöründe araştırmacıların ilgi odağı olmuştur. Hastalıkların uyarılarını erken fark etmek, erken tanı ve hastalık sınıflandırması hastalığın yönetimi için en önemli unsur olmasıyla beraber hastalık bulgularını araştırmak ve ileri tetkik için hastaların sevk edilmesi hastaların ölüm riskini azaltmaktadır. Ayrıca tetkiklere bağlı olarak hastalığı sınıflandırmak ve hızlı bir şekilde teşhis koymak hastanın yaşam kalitesini artırmaktadır.

Sağlık hizmetlerinde karar verme, çeşitli yönleri bir araya getirmesi ve farklı paydaşları içermesi nedeniyle karmaşık bir süreçtir (Öztürk vd., 2020). Hastalık tahmin çalışmalarında en çok kullanılan yöntemlerden biri olan makine öğrenmesi algoritmaları ise çok hızlı sonuç vermektedir. Bununla birlikte birçok algoritmaya sahip olan makine öğrenmesi elde edilen çıktının sonunda algoritmanın performans yüzdesini de göstermektedir. Diğer yandan araştırmacılar veri için hangi algoritmayı uygulama konusunda karışıklıklar yaşarken bazen de seçtikleri algoritmada zaman kaybettikleri de görülmektedir.

Dünya Sağlık Örgütü 2018 istatistiklerine göre; trakea, bronş ve akciğer kanserlerinden ölümler 1,2 milyondan 1,8 milyona yükselmiş ve şu anda önde gelen ölüm nedenleri arasında 6. sırada yer almıştır. 2019'da Alzheimer hastalığı ve diğer bunama türleri 7. önde gelen ölüm nedeni sıralamasında yerini almıştır. Diyabet, 2000 yılından bu yana %70'lik önemli bir artışın ardından, ilk 10 ölüm nedeni arasına girmiş olup aynı zamanda diyabet, 2000 yılından bu yana %80'lik bir artışla ilk 10 neden arasında erkek ölümlerindeki en büyük artıştan da sorumlu olmuştur. Böbrek hastalıkları tüm dünyada, önde gelen ölüm nedenleri arasında 13. sıradan 10. sıraya yükselmiş ve ölüm sayısı 2000'de 813.000'den 2019'da 1,3 milyona çıkmıştır (WHO, 2018).

Bulaşıcı olmayan hastalıklar, makine öğrenmesi araştırmaları alanında önemli oranda yer almakla birlikte araştırmacılar makine öğrenmesi algoritmalarını seçerken bazı karmaşıklıklar yaşadıkları gözlemlenmektedir. Bu amaçla, bu çalışma ile genel olarak halk sağlığını tehdit eden ve dünyada ölüm nedenleri listesinde ilk sıralarda yer alan bulaşıcı olmayan hastalıklardan oluşan çalışmalarda en iyi performans gösteren (doğruluk yüzdelik oranları yüksek olan) algoritmalar özetlenmiştir. Bu kapsamda erken teşhis, hastalık sınıflandırması, ilaç sınıflandırması, hastalık tahmin çalışmaları gibi birçok makine öğrenmesi çalışmaları ayrıntılı bir şekilde sınıflandırılmıştır. Ek olarak bulaşıcı olan ve son yıllarda halk sağlığı için acil durum ilan edilen COVID-19 hastalığına da yer verilmiştir. Sağlık alanında araştırılan makine öğrenmesi çalışmaları hastalık bazında incelenmiştir. Önerilen makine öğrenmesi sınıflandırma şemada gösterilmektedir (Şekil 1).



Şekil 1: Makine Öğrenmesi Çalışmalarının Sınıflandırılması.

Bu literatür çalışmasının temel amacı ise sağlık alanında çalışma yapan araştırmacılara, uygun algoritmaları seçmesinde yardımcı olmaktır. Bu çalışmada, algoritma seçme de yaşanan karmaşıklık ve zaman kaybının önlenmesi adına sağlık hizmetlerinde uygulanan makine öğrenmesi algoritma çalışmaları sınıflandırılmıştır. Ayrıca uygulanan algoritma sonucunda algoritma performans oranları gibi birçok bilginin yer aldığı güncel çalışmaların toplandığı bir derleme çalışmasıdır. Ayrıca halk sağlığını tehdit eden ve dünya da ölüm nedenleri listesinde ilk sıralarda yer alan hastalıklar için ayrı tablolar oluşturulmuş ve algoritmanın performans sonuçları hakkında bilgilere de yer verilmiştir. Çalışmada toplam 10 hastalık ve 24 farklı algoritma yer almaktadır.

2. Sağlık Sektöründe Makine Öğrenimi Çalışmaları

2.1 COVID-19 Hastalığına Ait Çalışmalar

Dünya Sağlık Örgütü, Şubat 2020'de SARS-CoV-2 virüsünün yol açtığı bulaşıcı hastalığa Covid-19 adını vermiştir. Dünya Sağlık Örgütü, bu salgını 30 Ocak 2020'de Uluslararası Öne Sahip Halk Sağlığı Acil Durumu olarak ilan etti (Sun vd.,2020). Başlangıçta, ciddi bir vaka olarak görülmemiş ancak, enfeksiyonunun yayılma hızı ve yüksek ölüm oranı kısa sürede tüm dünyada sağlık sistemlerinde birçok sorun yaşanmasına neden olmuştur.

Ortaya konulan derleme çalışması kapsamında sağlık hizmetlerinde 2020-2021 yılları arasında gerçekleştirilmiş olan makine öğrenmesi çalışmalarının analizini gerçekleştirmek amacıyla, araştırmaların sistematik bir incelemesi yapılmıştır. Türkçe ve İngilizce olarak "sağlık", "Covid-19" ve "bulaşıcı hastalıklar" ve "sağlık hizmetleri", veya "makine öğrenmesi" veya "algoritma performans analizi" veya "algoritma" anahtar kelimeleri kullanılarak DergiPark, Google Akademik, Science Direct ve Web of Science veri tabanları taranmıştır. Makaleler gözden geçirilirken, uygun literatür seçimi için salgınla ilgili farklı bakış açıları ve yöntemler temel alınarak anahtar kelimelerle kapsam daraltılmıştır.

Bu çalışma kapsamında taranan, COVID-19 hastalığı için farklı algoritmaların bir araya geldiği makine öğrenmesi çalışmaları, doğruluk oranı en yüksek olan algoritma, ele alınan problem, karşılaştırma yapılan algoritmalar ve sonuçlar bilgileriyle birlikte Tablo 1'de özetlenmiştir. Tarama sonucunda, RF (Random Forest), SVM (Support Vector Machine), DT (Decision Tree), ANN (Artificial Neural Network), LR (Logistic Regression), K Nearest Neighbor (K-NN), PR (Polynomial Regression) algoritmaları ile yapılan çalışmalara ulaşılmıştır.

Gambhir vd., (2020) Hindistan'daki COVID-19 hastalığının bulaşma eğilimini analiz etmişlerdir. Hindistan Sağlık ve Aile Refahı Bakanlığına ait veri seti ve toplam 154 gün olan 22 Ocak 2020-24 Haziran 2020 arası elde edilen veriler kullanılmıştır. Polynomial Regression ve Support Vector Machine (SVM) algoritmalarını karşılaştırmışlardır. Temmuz ve ağustos ayları için vakalardaki artış tahmin edilmiş ve yaklaşık %93 doğruluk göstermiştir.

Bayat vd., (2020) SARS-CoV-2 hastalığının teşhisi için test sayısı yeterli olmadığı durumda hastalığı tanımlamak için başka yöntemler denenmektedir. SARS-CoV-2 test sonuçları ile SARS-CoV-2 test tarihi civarında 2 günlük süre içinde toplanan 20 rutin laboratuvar testi arasındaki ilişkiyi analiz etmek için bir makine öğrenmesi modeli kullanmışlardır. Toplam yatarak ve ayakta tedavi gören 75 991 hastadan oluşan bir veri seti kullanmışlardır. Sonucu pozitif ve negatif olan hastaları karşılaştırmışlardır. Çalışma sonucunda (CI %95) %82,4 duyarlılık ve %86,4 genel doğruluk [%86,0-%86,9] çıkmıştır.

Saha vd., (2021) hastaların göğüs röntgen görüntülerini inceleyerek COVID-19'u teşhis etmeyi amaçlamışlardır. EMCNet adlı otomatik tespit şeması önermişler ve evrişimli sinir ağı kullanarak modeli geliştirmişlerdir. Veri seti, eğitim ve test analizi olmak üzere ikiye ayırmışlardır. Veri seti toplam da 4600 hastadan oluşmaktadır. Veriye ön işlem ve veri bölme aşamalarından sonra evrişimli sinir ağı modelini tamamlamışlardır. Elde edilen çıktıları sınıflandırmak için makine öğrenmesi algoritmalarını kullanmışlardır. Evrişimsel sinir ağı algoritması %96,52 doğruluk oranına ulaşmıştır. Çalışmada Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT) ve AdaBoost makine öğrenmesi algoritmalarını kullanmışlardır. Bu algoritmaların çıktıları iyi sonuçlar sağlamıştır. Bu sınıflandırmada en iyi performans sağlayan makine öğrenmesi algoritması Support Vector Machine (SVM) olmuş ve %96,96 doğruluk oranında çıkmıştır.

Pourhomayoun vd., (2021) çalışmalarında, 307.382 hasta örneği ve dünyanın da 146 ülkesinden 2.670.000'den fazla laboratuvar onaylı COVID-19 hastasından oluşan bir veri seti kullanmışlardır. Çalışmada, hastaneler ve tıbbi tesisler için kimin önce ilgilenilmesi gerektiği, kimin hastaneye kaldırılmada daha yüksek önceliğe sahip olduğunun tanımlanması amaçlanmıştır. Aynı zaman da sistemde aşırı kalabalığın önlenmesi ve gerekli bakımın sağlanmasındaki gecikmelerin ortadan kaldırılmasına yardımcı olacak bir yapay zeka modeli önermişlerdir. Sonuçlar, ölüm oranını tahmin etmede %89,98 oranında genel doğruluk göstermektedir. COVID-19 hastalarında ölüm oranını tahmin etmek için Support Vector Machine (SVM), Artificial Neural Networks (ANN), Random Forest (RF), Decision Tree (DT), Logistic Regression (LR) ve K Nearest Neighbor (K-NN) dahil olmak üzere birkaç makine öğrenmesi algoritması kullanmışlardır. En iyi performans gösteren algoritma Artificial Neural Networks (ANN) olmuş ve %89,98 doğruluk oranında çıkmıştır.

Alves vd., (2021) genel kullanıma açık anonim bir veri seti kullanılarak, rutin kan testleri sonuçları üzerinden COVID-19 tanısının makine öğrenmesi algoritmaları yardımıyla tahmin edilmesine yönelik model geliştirilmesi amaçlanmaktadır. Ham veri seti, 5644 örnek ve 111 özellik içermektedir. Kriter grafiği kullanarak modele karar vermek için Decision Tree (DT) ve Random Forest (RF) algoritmalarını kullanmışlardır. Çalışma sonucunda en yüksek doğruluk oranına sahip algoritma Random Forest (RF) algoritması olmuş ve %88 doğruluğuna sahip çıkmıştır.

Tablo 1. COVID-19 Hastalığı Alanında Yapılan Makine Öğrenmesi Çalışmaları.

Yazar ve Yayın Yılı	Algoritma	Ele Alınan Problem	Karşılaştırma Yapılan Algoritmalar	Sonuçlar	Doğruluk (%)
Saha vd.(2021)	SVM	Hindistan'daki COVID-19 hastalığının bulaşma eğilimini analiz etmek	PR, RF	Temmuz ve Ağustos ayları için vakalardaki artışı tahmin edilmiştir.	%93
Pourhomayoun vd (2021)	XGBoost	SARS-CoV-2 pozitif ve negatif hastaları teşhis etmek	-	COVID-19 enfeksiyon durumunu belirlemeye yönelik tamamlayıcı bir yöntem sağlayabilir.	
Gambhir vd. (2020)	SVM	Göğüs röntgeni görüntülerini inceleyerek COVID-19 teşhisi tanımlamak	RF, SVM, DT, AdaBoost	SVM en iyi tahmincidir.	%96,96
Bayat vd. (2020)	ANN	COVID-19 hastalığının ölüm riskini belirlemek	SVM, ANN, RF,DT, LR, K-NN	Ölüm oranını tahmin etmede ANN algoritması en iyi tahmincidir.	%89,98
Alves vd. (2021)	RF	Brezilya hastane verisine bakarak COVID-19 hastalığını teşhis etmek	DT, RF	Random Forest (RF) algoritması en iyi tahmincidir.	%88

2.2. Bulaşıcı Olmayan Hastalıklarla İlgili Çalışmalar

2.2.1. Kalp Hastalığına İlişkin Çalışmalar

Ortaya konulan derleme çalışması kapsamında sağlık hizmetlerinde 2016-2021 yılları arasında gerçekleştirilmiş olan makine öğrenmesi çalışmalarının analizini gerçekleştirmek amacıyla, araştırmaların sistematik bir incelemesi yapılmıştır. Türkçe ve İngilizce olarak "sağlık", "bulaşıcı olmaya hastalıklar" ve "sağlık hizmetleri", veya "makine öğrenmesi" veya "performans analizi" veya "Algoritma" anahtar kelimeleri kullanılarak DergiPark, Google Akademik, Science Direct ve Web of Science veri tabanları taranmıştır.

Ölümcül hastalıklar listesinin başında kalp hastalıkları gelmektedir. Küresel olarak önemli bir ölüm nedeni olarak kabul edilen kalp hastalığı, Dünya Sağlık Örgütü (2016) istatistiklerine göre, 2016 yılında yaklaşık 17,9 milyon kişinin ölümüne yol açmıştır. Bu nedenle kalp hastalığının sınıflandırılması, erken teşhisi, kalp hastalığı riski gibi çalışmalar önemli yer tutmaktadır. Derleme çalışması kapsamında taranan, kalp hastalığı için farklı algoritmaların bir araya geldiği makine öğrenmesi çalışmaları, doğruluk oranı en yüksek olan algoritma, ele alınan problem, karşılaştırma yapılan algoritmalar ve sonuçlar bilgileriyle birlikte Tablo 2’de özetlenmiştir. Tarama sonucunda NB (Naive Bayes), RF (Random Forest), SVM (Support Vector Machine), DT (Decision Tree), ANN (Artificial Neural Network), LR (Logistic Regression), K-NN (K Nearest Neighbor), PR (Polynomial Regression), Hybrid Neural Network (HNN), RDNN (Recurrent Deep Neural Network) algoritmalarına ulaşılmıştır.

Mung vd., (2020) akciğer kanseri, kalp hastalığı, diyabet hastalığı, rahim ağzı kanseri hastalıklarından oluşan veri setlerine ayrı ayrı 4 algoritma uygulamışlardır. UCI makine öğrenmesi veri kümelerini kullanmışlardır. Veri setini incelemek için, Naive Bayes (NB), Decision Tree (DT), K Nearest Neighbor (K-NN), Ensemble Learning (EL) algoritmaları ile karşılaştırma yapmışlardır. Veri setini test ve eğitim seti olarak incelemişlerdir. Sınıf etiketleri, sınıflandırıcı için tahmin edilen olasılıklar p 'ye dayalı olarak tahmin edilebilir bir model oluşturmuşlardır. Kalp hastalığı verisinde 370 hasta, 14 öznitelik kaydı oluşturmuşlardır. Kalp hastalığıyla birlikte diğer hastalıklarda da doğruluk sonucu en yüksek Ensemble Learning algoritması vermiştir.

Ahmed vd., (2020) 2016 yılı için Cleveland’a ait kalp hastalarının yaş, cinsiyet, kan şekeri değerleri,

kolesterol değerleri, maksimum kalp atışı gibi birçok veriden oluşan bir veri seti kullanmışlardır. Veri setine 4 algoritma uygulamışlardır. Tweet akışlarından “hrtdis” başlıklı kelimeleri ayırt ederek Kafka’da depolanmasını sağlamışlardır. Çevrimiçi tahmin, eğitim veri kümesindeki sıra ile aynı sırada bir özellik vektörü olarak öznitelikleri almışlardır. Bir tweetin kalp hastalığı belirtileri içerip içermediğini tahmin etmeyi hedeflemişlerdir. Random Forest (RF) sınıflandırıcısının %94,9 ile en yüksek doğruluğu elde ederek diğer modellerden daha iyi performans gösterdiğini sunmuşlardır.

Li vd., (2020) kalp hastalığı hastalarını teşhis etmek için 297 hasta ve 13 öznitelik verilerine sahip kalp hastalarına ait bir veri seti kullanmışlardır. Kalp hastalığı tahmini için birçok algoritma denemişlerdir. Makine öğrenmesi sınıflandırıcıları olarak Logistic Regression (LR), K Nearest Neighbor (K-NN) Artificial Neural Networks (ANN), Support Vector Machine (SVM), Naive Bayes (NB) ve Decision Tree (DT) algoritmalarını kullanmışlardır. Model değerlendirmesinin en iyi uygulamalarını öğrenmek ve hiperparametre ayarlaması için çapraz doğrulama yöntemi kullanmışlardır. Sınıflandırıcıların performansları, öznitelik seçim algoritmaları ile seçilen öznitelikler üzerinde kontrol etmişlerdir. FCMIM, MRMR, LASSO, LLBFS'yi içeren dört standart özellik seçim algoritmasını veri setine uygulamışlardır. Doğruluk oranı en yüksek olan algoritmayı tanımlamışlardır. Features Selected By Proposed Fs algoritmasıyla (FCMIM) birlikte Support Vector Machine (SVM) algoritmasının doğruluk oranlarını yüksek sonuç elde ettiğini gözlemlemişlerdir.

Balaji vd., (2020) kardiyovasküler hastalık olarak bilinen kalp hastalığı teşhisini tahmin etmek için en iyi sınıflandırıcıyı bulmayı amaçlamışlardır. UCI veri tabanından alınan kalp hastalığı verisine; Naive Bayes (NB), Decision Tree (DT), K Nearest Neighbor (K-NN) ve Random Forest (RF) algoritması gibi denetimli öğrenme algoritmalarını uygulamışlardır. Kalp hastalığı hastalarının veri setinde 303 hasta ve 76 öznitelik içerir. Bu 76 öznitelikten yalnızca 14 özniteliği test için dikkate almışlardır. Çalışmada, makine öğrenmesi algoritmalarında en yüksek doğruluk oranına sahip algoritmayı tanımladıktan sonra hastalarda kalp hastalığına yakalanma olasılığını tahmin etmişlerdir. Kardiyovasküler hastalık olarak bilinen kalp hastalığı teşhisini tahmin etmek için en iyi sınıflandırıcıyı sonuç olarak ele almışlardır. K Nearest Neighbor (K-NN) algoritması başarılı bir sonuç vermiştir.

Sridhar vd., (2021) kalp hastalığı hastalarının teşhisini tahmin etmek için en iyi sınıflandırıcıyı bulmak amacıyla Kaliforniya Üniversitesine ait (UCI veri tabanı) Cleveland kalp hastaları veri kümesi üzerine çalışmışlardır. 297 hastalar arasından 252 tanesini kullanarak analiz etmişlerdir. Recurrent Deep Neural Network (RDNN) ve Hybrid Neural Network (HNN) algoritmalarını kalp hastalığı verisine uygulamışlardır. RDNN algoritmasının doğruluk sonucu ise %97,78 oranında çıkmıştır.

Jothi vd., (2021) kalp hastalığının risk düzeyini tahmin etmek için K Nearest Neighbor (K-NN) ve Decision Tree (DT) algoritması kullanmışlardır. Makine öğrenmesi kullanılarak bir kalp hastalığı riskini belirlenmiş ve kalp hastalığına sahip olabilecek bireyler tahmin edilmiştir. Kalp hastalığı verisi yaş, cinsiyet, hızlı kan şekeri, göğüs ağrısı vb. 13 tıbbi parametre içermektedir. Sonuçlar Python programlama kullanılarak üretilmiştir. İki algoritma sonucunda doğruluk yüzdesi en yüksek çıkan algoritma, Decision Tree (DT) modelinin verdiğini gözlemlemişlerdir. Bu algoritmanın doğruluk oranı %81 çıkmıştır.

Waris vd., (2021) çalışmalarında kalp hastalığını erken teşhis/ tahmin etmeyi hedeflemişlerdir. Kalp hastaları verisinde sigara içme, beslenme alışkanlıkları, diyabet, tansiyon gibi birçok bilgiler yer almaktadır. Bu bilgiler doğrultusunda senaryolarla doğruluk, hassasiyet ve özgünlük testleri

oluşturmuşlardır. Beş adımdan oluşan bu çalışmada ilk adım, veri ön işleme, ikinci adım verinin hesaplanması, üçüncü adım K-NN algoritmasının uygulanması, dördüncü adım sınıf tahmini aşamasının tamamlanması ve son adım ise algoritma doğruluğunu oluşturmaktır. Girdiler için diyagramlar çizmişlerdir. En iyi sonucu Novel K Nearest Neighbor (Novel K-NN) algoritmasının verdiğini gözlemlemişler ve bu algoritmanın doğruluk oranı %93 çıkmıştır.

Tablo 2. Kalp Hastalığı Alanında Yapılan Makine Öğrenmesi Çalışmaları.

Yazar ve Yayın Yılı	Algoritma	Ele Alınan Problem	Sonuçlar	Karşılaştırma Yapılan Algoritmalar	Doğruluk (%)
Mung vd.,(2020)	Ensemble Learning	Akciğer kanseri, Kalp hastalığı, diyabet hastalığı, Rahim ağzı kanseri hastalıklarının analizi için eniyi algoritmayı bulmak	Tüm veriler için en iyi algoritmanın Ensemble Learning algoritması olduğu kararı verilmiştir.	NB, DT, K-NN, Ensemble Learning	99,93
Ahmed ve diğ.(2019)	RF	Kalp hastalığı tahmini için yüksek doğruluk sağlayan optimum makine öğrenme algoritmasını bulmak	RF algoritmasının tam özellikli doğruluğunun, rekabetçi çalışmaya kıyasla doğruluğu %11 artırdığını kanıtlamışlardır.	DT, RF, SVM, LR	94,9
Shah ve diğ.(2021)	K-NN	Kalp hastalığı teşhisini tahmin etmek için en iyisınıflandırıcıyı bulmak	En iyi sonucun K-NN algoritması olduğu kararı verilmiştir.	NB, DT, K-NN, RF	90,8
G.M. Sridhar ve diğ.(2021)	RDNN	Kalp hastalarının teşhisini tahmin etmek için en iyi sınıflandırıcıyı bulmak	Kalp hastalığı tanısında farklı sınıflandırıcı teknikler uygulanmışlardır.	RDNN, HNN	97,78
Jothi ve diğ.(2021)	DT	Kalp hastalığının risk düzeyini tahmin etmek	Hastanın kalp hastalığına yakalanma ihtimalini hesaplayan bir sistem kurmak hedeflenmiştir.	K-NN, DT	81
Waris ve diğ.(2021)	Novel KNN	Kalp ile ilgili hastalıkların teşhisinde yardımcı olmak ve çeşitli makine öğrenme tekniklerini test etmek	Kalp hastalığı tanısında farklı sınıflandırıcı teknikler uygulanmışlardır.	Novel K-NN, K-NN	93
Li ve diğ.(2020)	SVM	Kalp hastalığını zamanında ve verimli bir şekilde tanımlamak	Kalp hastalığını teşhis etmek için verimli ve doğru bir sistem önermek hedeflenmektedir.	LR, K-NN, ANN, SVM, NB, DT	92,37

2.2.2. Diyabet Hastalığı İçin Ulaşılan Çalışmalar

Diyabet hastalığı olan hasta sayısı 2015 yılında 415 milyon hastaya ulaşmış ve 2040 yılında bu sayının yaklaşık 642 milyon hastaya çıkması beklenmektedir (Kafes vd.,2018). Diyabet tedavisi yetersiz kalmasıyla beraber çok maliyetlidir. Bu nedenle diyabet hastalık sebeplerini bulmak, hastalığı erken teşhis etmek ve önlem almak çok önemlidir. Derleme çalışması kapsamında taranan, diyabet hastalığı için farklı algoritmaların bir araya geldiği makine öğrenmesi çalışmaları, doğruluk oranı en yüksek olan algoritma, ele alınan problem, karşılaştırma yapılan algoritmalar ve sonuçlar bilgileriyle birlikte Tablo 3’ te özetlenmiştir. Ekler kısmında algoritma doğruluk oranları ayrıca gösterilmiştir. Tarama sonucunda NB (Naive Bayes), RF (Random Forest), SVM (Support Vector Machine), DT (Decision Tree), LR (Logistic Regression), K-NN (K Nearest Neighbor), LDA (Linear Discriminant Analysis) algoritmalarına ulaşılmıştır.

Lukmanto vd., (2019) diyabet hastalığının sınıflandırılması ve tespitinde F-Skor özellik seçimi ve bulanık Support Vector Machine ‘den (SVM) yararlanmışlardır. Pima Indian Diabetes (PID) veri setine uygulama yapmışlardır. Veri seti içerisinde glukoz değerleri, kan basıncı, insülin değeri, vücut kitle endeksi gibi birçok veri yer almaktadır. Sonuç olarak Support Vector Machine (SVM) algoritma doğruluğu %89,02 çıkmıştır.

Mujumdar vd., (2019) Pima Indian Veri setini kullanarak glukoz, VKİ, yaş, insülin vb. faktörleri içeren veri setin de diyabet sınıflandırması için bir diyabet tahmin modeli önermişlerdir. Diyabet hastalığını sınıflandırmayı hedeflemişlerdir. Logistic Regression (LR) %96'lık en yüksek doğruluğu verdiğini ve ek olarak ardışık düzen sonuçları (Pipelining Results) %98,8 doğruluk oranıyla en iyi model olarak AdaBoost sınıflandırıcısı olduğunu bulmuşlardır. AdaBoost, Gradient Boosting, Random Forest (RF), Logistic Regression, Extra Trees Classifier, Linear Discriminant Analysis, algoritmaları kullanılmıştır.

Nguyen vd., (2019) çalışmalarında, tip 2 diyabet hastalığı başlangıcının tahminini hedeflemişlerdir. Amerika Birleşik Devletleri nüfusu için devlet hastanesi kayıt verilerini kullanarak diyabet başlangıcı olan hastaları tahmin etmeyi amaçlamışlardır. Veri seti, 2009-2011 yılları arasında, 1904'ü tip 2 diyabet olan 9948 hasta verisinden oluşmuştur. %80 eğitim seti olarak kullanılan veri kümesi örnekleri ve verilerin %20'si test seti olarak ayarlamışlardır. SMOTE (Synthetic Minority Oversampling Technique) algoritması kullanarak %82,12 doğruluk sonucuna ulaşmışlardır.

Tigga vd., (2020) Hindistan’da, diyabet hastalığı riskini öğrenmek için 952 örnek ve 18 soru içeren anket ile verileri online sistem üzerinden toplamışlardır. Veri seti içerisinde; yaş, cinsiyet, sigara içme durumu, vücut kitle endeksi, alkol alışkanlığı, uyku düzeni gibi birçok öznelik bulunmaktadır. Logistic Regression (LR), Support Vector Machine (SVM), K Nearest Neighbor (K-NN), Naive Bayes (NB), Decision Tree (DT), Random Forest (RF) algoritmalarını kullanarak doğruluk oranı en yüksek olan algoritmayı bulmayı hedeflemişlerdir. Doğruluk oranı en yüksek olan algoritmanın Random Forest (RF) algoritması olduğunu ve doğruluğunun %94,10 olduğunu belirlemişlerdir.

Viloria vd., (2020) Kolombiya hasta veri seti kullanmışlardır. Hastaların diyabet hastası olma ihtimalini Support Vector Machine (SVM) algoritmasını kullanarak tahmin etmeye çalışmışlardır. Elde edilen sonuçlarda diyabet hastasıdır, diyabet hastalığına yatkınlığı var, diyabet hastalığı yok gibi sonuçlar elde etmişler. Support Vector Machine (SVM) algoritması %99,2 gibi iyi bir doğruluk yüzdeliğine sahip çıkmıştır.

Kumar vd., (2020) Pima Indian Diabetes veri kümesini (PID), UCI deposundan almışlardır. Çalışma ile diyabet hastalarına doğru ilacın verilmesi amaçlanmıştır. Veri setinde toplam 9 öznelik, 768 hasta bulunmaktadır. Diyabet tip 2'yi tahmin etmek için, diyabet hastalığı verisinde Random Forest (RF) ve Decision Tree (DT) ve Deep Neural Network (DNN) algoritmaları kullanmışlardır. Veriyi 60 öğrenme, 40 test şeklinde ayırarak analiz ettiklerinde doğruluğu %96,77, 70 öğrenme 30 test şeklinde ayırarak analiz ettiklerinde doğruluğu %97,54, 80 öğrenme 20 test şeklinde ayırarak analiz ettiklerinde doğruluğu %98,16 elde etmişlerdir. Bu çalışma sonucunda Deep Neural Network (DNN) ile %98,16 doğruluk hesaplamışlardır.

Reddy vd., (2020) çalışmada Support Vector Machine (SVM), K Nearest Neighbor (K-NN), Logistic Regression (LR) Naive Bayes (NB) gibi birçok makine öğrenmesi algoritması kullanmışlardır. Pima Indian Veri setini kullanmışlar. Glukoz, VKİ, yaş, insülin vb. faktörleri

içeren veri setine diyabet sınıflandırması için bir diyabet tahmin modeli önermişlerdir. Bu çalışmanın temel amacı, insandaki diyabeti maksimum doğrulukla öngören bir model tasarlamaktır. En iyi yüksek performansa sahip olan algoritmanın %98,48 doğruluk yüzdesi ile Random Forest (RF) algoritması olduğunu bulmuşlardır.

Daniyal Baig vd., (2020) diyabet hastalığını erken teşhis etmeyi amaçlamışlardır. Önce veriler toplanmış, veri analizi, veri ön işleme, analiz gibi çalışmalardan sonra makine öğrenmesi algoritmalarını uygulamışlardır. Çalışma, Linear Regression, Decision Tree (DT), K Nearest Neighbor (K-NN) ve Random Forest (RF) gibi algoritmaları ve bunların hastalık tahmini ve erken tedavide erken teşhis için faydalarını tartışmışlardır. En iyi performans gösteren algoritma ise %99 doğruluk oranı ile Random Forest (RF) olmuştur.

Tablo 3. Diyabet hastalığı Alanında Yapılan Makine Öğrenmesi Çalışmaları.

Yazar ve Yayın Yılı	Algoritma	Ele Alınan Problem	Sonuçlar	Karşılaştırma Yapılan Algoritmalar	Doğruluk (%)
Nguyen vd. (2019)	SMOTE	Diyabet başlangıcı olan hastaları tahmin etmek	Makine öğrenmesi modelleriyle iyi sonuçlar elde etmişlerdir.	-	82,12
Tiggaa vd. (2020)	RF	Diyabet riskini tahmin etmek	Random Forest (RF) en iyi tahmincidir.	LR, K-NN, SVM, NB, DT, RF	94,1
Lukmanto vd. (2019)	SVM	Diyabet hastalığını sınıflandırmak	Diyabet hastaları tahmin etmede %89,02 doğruluk göstermektedir.	-	89,02
Viloria vd. (2020)	SVM	Hastaların diyabet hastası olma ihtimali tahmin etmek	En iyi sonucu Support Vector Machine (SVM) algoritması vermiştir.	-	99,2
Kumar vd. (2020)	DNN	Diyabet hastalarına doğru ilacı önermek	DNN hastalara doğru ilacı teşhis etmeye yardımcı olmaktadır.	K-NN, DT	98,16
Reddy vd. (2020)	RF	İnsandaki diyabeti maksimum doğrulukla öngören bir model tasarlamak	Random Forest diğer sınıflandırıcılara kıyasla en yüksek performansı göstermiştir	SVM, K-NN, LR, NB	98,48
Mujumdar vd. (2019)	AdaBoost	Diyabet sınıflandırması için bir diyabet tahmin modeli önermek	En iyi çözüm veren algoritmanın AdaBoost algoritması olduğunu bulmuşlardır.	AdaBoost, Gradient Boost, RF, LR, DT, LDA	98,8
Baig vd. (2020)	RF	Makine öğrenmesi algoritmaları kullanarak diyabet hastalığını erken teşhis etmek	En iyi çözüm veren algoritmanın Random Forest algoritması olduğunu bulmuşlardır.	Linear Regression, DT, K-NN, RF	99

2.2.3. Karaciğer Hastalığı İçin Ulaşılan Çalışmalar

Vücutta en önemli organlardan biri olan karaciğer karın boşluğunun sol üst köşesin de yer alır. Yaygın karaciğer hastalık problemleri, karaciğer yağlanması, karaciğer kanser, hepatit gibi birçok hastalık vardır. Kronik hastalıkların yaygınlığının artması, yaşam süresinin uzaması, bilim ve mühendislikteki gelişmeler sağlık teknolojisindeki yenilikleri hızlandırmaktadır (Öztürk vd., 2016). Derleme çalışması kapsamında taranan, karaciğer hastalığı için farklı algoritmaların bir araya geldiği makine öğrenmesi çalışmaları doğruluk oranı en yüksek olan algoritma, ele alınan problem, karşılaştırma yapılan algoritmalar ve sonuçlar Tablo 4'te özetlenmiştir. Çalışmalarda, NB

(Naive Bayes), RF (Random Forest), SVM (Support Vector Machine), DT (Decision Tree), LR (Logistic Regression), K Nearest Neighbor (K-NN) LDA (Linear Discriminant Analysis), ANN (Artificial Neural Network) algoritmalarının uygulandığı görülmektedir.

Gatos vd., (2017) makine öğrenme algoritması ile ultrason kayma dalgası elastografi (ultrasound shear wave elastography/SWE) görüntülemesini kullanarak kronik karaciğer hastalığını sınıflandırmışlardır. Veri seti için, 126 kişiden alınan karaciğer görüntüsünü kullanmışlardır. 126 hastadan oluşan bir klinik veri setine, ilk olarak, bir RGB-to-stiffness ters eşleme tekniği kullanıldıktan sonra karşılık gelen farklı renk bölgelerini, SWE üreticisi tarafından sağlanan renk çubuğundan elde edilen belirli sertlik değeri aralıklarıyla ilişkilendiren beş kümeli bir ayırım yapmışlardır. Support Vector Machine (SVM) modeli ile kronik karaciğer hasta denek ayrımcılığına göre sınıflandırılmasında en yüksek doğruluk ile (%93,5) en iyi değer elde edilmiştir.

Gogi vd., (2018) karaciğer hastalarının sınıflandırması/tahminlemesi amacı ile sınıflandırma için makine öğrenmesi algoritmalarından yararlanmışlardır. Karışıklık matrisi, sınıflandırıcının gerçek pozitif ve yanlış pozitif oranlarını vermiştir. Eğri Altında Kalan Alan (AUC) neredeyse 1'e eşit bulunmuştur. Test veri kümesindeki sınıflandırma modeli performansını gösteren karışıklık matrisi 572 hasta test etmişlerdir. Bunlardan 394'ü "Evet" ve 178'i "Hayır" ile tahmin sonucu bulmuşlardır. Decision Tree (DT), Support Vector Machine (SVM), Doğrusal Ayrımcı ve Logistic Regression (LR) algoritmalarını uygulamışlardır. Veri seti hastaların laboratuvar parametrelerini içermektedir. Matlab yazılımını kullanarak en iyi algoritmanın %95,8 doğruluk oranı ile Logistic Regression olduğunu kanıtlamışlardır.

Thaiparnit vd., (2018) karaciğer hastalığını sınıflandırmak amacıyla 7 tip karaciğer hastalığı için sınıflandırma yapmışlar ve 359 hasta verisi kullanmışlardır. Sınıflandırmayı değerlendirmek için, testin hipotezi, kuralları kullanarak 5 katlı çapraz doğrulama ayarlamışlardır. Veri Kurallar Bölümü, OneR kuralı, Tree Decision Stump technique, Tree REPTree ve Random Forest (RF), tekniklerini kullanmışlardır. Random Forest (RF) algoritması, diğer sınıflandırma modellerinden daha yüksek performans göstermiştir.

Wu vd., (2019) yağlı karaciğer hastalığını (FLD) tahmin etmek ve erken teşhisi sağlamak amacı ile Random Forest (RF), Naive Bayes (NB), Artificial Neural Networks (ANN), Logistic Regression (LR) kullanarak karaciğer hastalığını erken teşhis etmek için model oluşturmuşlardır. Random Forest (RF), modeli, %87,47 doğruluk oranı ile diğer sınıflandırma modellerinden daha yüksek performans göstermiştir. Klinik ortamda Random Forest (RF), modelinin uygulanması, doktorların yağlı karaciğer hastalarını birincil koruma, gözetim, erken tedavi ve yönetim için sınıflandırmasına yardımcı olabileceği belirtilmiştir.

Srivenkatesh (2019) farklı makine öğrenmesi ile karaciğer enfeksiyonunu öngörmek ve en verimli algoritmayı seçmek amacıyla karaciğer enfeksiyon tahmin sınıflaması yapmıştır. Karaciğer hastalığına sahip, on öznitelikten oluşan 583 hastanın verilerini kaggle veri setinden almışlardır. Bu araştırmada, karaciğer hastalığının incelenmesi için Naive Bayes, Logistic Regression Support Vector Machine (SVM), Random Forest (RF), K Nearest Neighbor (K-NN) olmak üzere beş çeşit sınıflandırıcı uygulamışlardır. Çalışmanın sonucunda Logistic Regression (LR) modeli, diğer sınıflandırma modellerinden daha yüksek performans göstermiştir.

Shi vd., (2021) Ocak 2013'ten Mayıs 2017'ye kadar olan karaciğer hasta verilerinden oluşan geriye dönük bir incelemedir. Çalışma, arkus aort ameliyatı geçirmiş 672 ardışık hastadan oluşan bu veri setine Logistic Regression Support Vector Machine (SVM), Random Forest (RF), Naive

Bayes algoritması ile sınıflandırma yapmışlardır. Bu çalışmanın sonucu Logistic Regression modeli, diğer sınıflandırma modellerinden daha yüksek performans göstermiştir.

Tablo 4. Karaciğer Hastalığı Alanında Yapılan Makine Öğrenmesi Çalışmaları.

Yazar ve Yayın Yılı	Algoritma	Ele Alınan Problem	Sonuçlar	Karşılaştırma Yapılan Algoritmalar	Doğruluk (%)
Gatos vd. (2017)	SVM	Kronik karaciğer hastalığını sınıflandırmak	Support Vector Machine (SVM)algoritmasının doğruluk oranını %93,5 bulmuşlardır.	-	93,5
Gogi vd. (2018)	LR	Karaciğer hastalarını sınıflandırmak	En iyi algoritmanın Logistic Regression olduğunu kanıtlamışlardır.	SVM, LR, LDA,DT	95,8
Wu vd. (2019)	RF	Yağlı karaciğer hastalığını tahmin etmek	Random Forest (RF), modeli eniyi sonuç veren algoritmadır.	RF, NB, ANN, LR	87,47
Thaiparnit vd. (2018)	RF	Karaciğer hastalarını sınıflandırmak	Random Forest (RF), modeli eniyi sonuç veren algoritmadır.	Veri Kuralları, OneR, Tree REPTree, Tree Decision Stump technique	75,76
Shi vd. (2021)	LR	Arkus aort ameliyatı geçirmiş karaciğer hastalarının veri seti incelenmek	Logistic Regression daha iyi performans göstermiştir.	LR, SVM, RF	77
Srivenkatesh (2019)	LR	Karaciğer enfeksiyon tahmin sınıflandırmak	Logistic Regression daha iyi performans göstermiştir.	NB, LR, SVM, RF,K-NN	76,27

2.2.4. Meme Kanseri Hastalığı İçin Ulaşılan Çalışmalar

Kadın hastalıklarının en başında gelen hastalıklarından biri olan meme kanseri ölümcül hastalıklardan biri olmasına rağmen erken teşhis de ölüm riskini azaltmaktadır. Bu nedenle bu alanda yapılan çalışmalar çok önem arz etmektedir. Derleme çalışması kapsamında taranan, meme kanseri için farklı algoritmaların bir araya geldiği makine öğrenmesi çalışmaları, doğruluk oranı en yüksek orana sahip olan algoritma, ele alınan problem, karşılaştırma yapılan algoritmalar ve sonuçlar bilgileriyle birlikte Tablo 5'te özetlenmiştir. Çalışmalarda, NB (Naive Bayes), RF (Random Forest), SVM (Support Vector Machine), DT (Decision Tree), LR Logistic Regression), K Nearest Neighbor (K-NN), LDA (Linear Discriminant Analysis), ANN (Artificial Neural Network) MLP (Multiple Linear Perceptron)), RBF SVM (Radial basis function kernel Support Vector Machine), DNNs (Deep Neural Network with Support Value) algoritmalarının uyguladığı görülmüştür.

Asri vd., (2016) Wisconsin Meme Kanseri veri setini kullanarak bu veri setine dört ana algoritma uygulamış ve en iyi makine öğrenmesi algoritmasını seçmeyi hedeflemişlerdir. Meme kanseri hastalığına sahip hastaların sınıflandırmasının analizi için Support Vector Machine (SVM), Naive Bayes (NB), K Nearest Neighbor (K-NN) ve C4.5 algoritmalarını kullanmışlardır. En iyi sınıflandırma doğruluğunu bulmak için bu algoritmaların verimliliğini ve etkililiğini doğruluk, kesinlik, duyarlılık ve özgüllük açısından karşılaştırma yapmışlar ve sonuç olarak SVM algoritmasının %97,13 doğruluğu elde ederek diğer makine öğrenmesi algoritmalarına göre en iyi performans gösteren algoritma olmuştur.

Karthikeyan vd., (2020). Wisconsin Meme Kanseri veri setini kullanarak üç ana algoritma kullanmış ve en iyi makine öğrenmesi algoritmasını seçmeyi hedeflemişlerdir. K Nearest

Neighbor (K-NN), Naive Bayes (NB), ve Random Forest (RF) algoritmalarını kullanmışlardır. En iyi sınıflandırma doğruluğunu bulmak için bu algoritmaların verimliliğini ve etkililiğini doğruluk, kesinlik, duyarlılık ve özgüllük açısından karşılaştırma yapmışlar ve sonuç olarak %95,90 doğruluk oranı ile K-NN algoritmasının diğer algoritmalara göre en iyi performans gösteren algoritma olduğunu belirlemişlerdir.

Vaka vd., (2020) makine öğrenmesi tekniklerini kullanarak meme kanserini tespit etmek için yeni bir yöntem sunmayı amaçlamışlardır. Önerilen yöntem, mevcut yöntemlerle karşılaştırıldığında oldukça doğru ve verimli sonuçlar üretmişlerdir. Deep Neural Network with Support Value (DNNS) yöntemi, önerilen yöntem derin bir sinir ağındaki destek değerine dayanmaktadır. Karşılaştırma yapmak için Support Vector Machine (SVM), Naive Bayes (NB) gibi makine öğrenmesi algoritmalarını kullanmışlardır. Deneysel sonuçlar, önerilen DNNS'nin mevcut yöntemlerden oldukça daha iyi performans gösterdiğini kanıtlamışlardır.

Sarkar vd., (2021) meme kanserini sınıflandırmak için MikroRNA analizini yapmışlardır. Meme kanserinde tümör baskılayıcı mikroRNA'lar, aşağı akım sinyal yolları ve bunlar tümörden sorumludur. Bu mikroRNA listesini elde etmek için makine öğrenmesi algoritmalarını kullanmışlardır. Çalışma 2 aşamadan oluşmaktadır. İlk aşamada, veri setini sınıflandırma doğruluğuna bakarak. Support Vector Machine (SVM), K Nearest Neighbor (K-NN), Artificial Neural Networks (ANN), Decision Tree (DT), Random Forest (RF), Naive Bayes (NB), DISCR algoritmalarını kullanarak en iyi doğruluğa sahip olanı seçmişlerdir. İkinci aşamada ise meme kanseri alt tiplerinin sınıflandırılmasını yapmışlardır. Random Forest (RF) algoritmasının diğer altı teknikten nispeten daha iyi doğruluk (%95 güven aralığında $76.5761 \pm 0,33$) sağladığını gözlemlemişlerdir.

Gopal vd., (2021) çalışmada nesnelerin interneti cihazlarıyla meme kanserini tahmin etme amacıyla makine öğrenmesi algoritmalarını kullanmışlardır. Wisconsin Meme kanseri Veri Kümesini (WBCD) kullanmışlardır. Veri setinde 32 öznitelik ve 569 örnek bulunmaktadır. Elde edilen sonuçları iyi huylu ve kötü huylu tümör olarak ayırmışlardır. Multiple Linear Perceptron (MLP), Logistic Regression (LR) ve Random Forest (RF), sınıflandırıcılarını kullanmışlar ve doğruluk oranı en yüksek gösteren algoritmanın MLP olduğunu bulmuşlardır.

Wu vd., (2021) Wisconsin Meme Kanseri veri kümelerinde, iki ana makine öğrenmesi algoritması kullanmışlardır. İçerisinde 32 öznitelik ve 569 örnek bulunmaktadır. Meme kanseri hastalığını sınıflandırmayı hedeflemişlerdir. Naive Bayes (NB) algoritması ve K Nearest Neighbor (K-NN) algoritması arasında bir karşılaştırma önermişler ve çapraz doğrulama kullanarak bunların doğruluğunu değerlendirmişlerdir. Algoritmalar arasında doğru bir karşılaştırma yaptıktan sonra, K-NN algoritmasının %97,51 gibi daha yüksek bir verimlilik elde ettiğini gözlemlemişlerdir.

Al-Azzam vd. (2021) meme kanseri tahmini için temel denetimli ve yarı denetimli makine öğrenme algoritmalarının performansını ve doğruluğunu karşılaştırmayı ve değerlendirmeyi amaçlamışlardır. Denetimli ve yarı denetimli öğrenme olmak üzere dokuz makine öğrenmesi sınıflandırma algoritması kullanmışlardır, Bunlar; Logistic Regression (LR), Gauss Naive Bayes (GNB), Linear Support Vector Machine (LSVM), Radial basis function kernel Support Vector Machine (RBF SVM), Decision Tree (DT), Random Forest (RF), Xgboost, Gradient Boosting son K Nearest Neighbor (K-NN) Wisconsin Tanı Kanseri veri kümesi kullanmışlardır. Modelin güvenilirliğini sağlamak için, K-kat çapraz doğrulama uygulamış ve hiperparametreleri optimize etmişlerdir.

Denetimli makine öğrenmesi algoritması olan için KNN modeli ve yarı denetimli makine öğrenmesi algoritması olarak Logistic Regression %98'lik en yüksek doğruluğu elde etmiştir.

Tablo 5. Meme Kanseri Alanında Yapılan Makine Öğrenmesi Çalışmaları.

Yazar ve Yayın Yılı	Algoritma	Ele Alınan Problem	Sonuçlar	Karşılaştırma Yapılan Algoritmalar	Doğruluk (%)
Sarkar vd. (2021)	RF	Meme kanseri alt tipini tahmin etmek	RF'nin diğer altı teknikten nispeten daha iyi olan doğruluk sağlamıştır.	SVM, ANN, K-NN, DT, RF, NB, DISCR	76,58
Gopal vd. (2021)	MLP	Meme kanserinin erken teşhisi sağlamak	En iyi performans gösteren algoritmanın MLP olduğunu bulmuşlardır.	MLP, LR, RF	98
Wu vd. (2021)	K-NN	Meme kanserini sınıflandırmak	En iyi performans gösteren algoritmanın K-NN olduğunu bulmuşlardır.	NB, K-NN, RF	97,51
Asri vd. (2016)	SVM	Meme kanserini sınıflandırmak	SVM algoritması iyi bir sonuç vermiştir.	NB, K-NN, C4.5	97,13
Karthikeyan vd. (2020)	K-NN	Meme kanserini sınıflandırmak	SVM algoritması iyi bir sonuç vermiştir.	NB, K-NN, RF	95,90
Al-Azzam vd. (2021)	K-NN ve LR	Meme kanseri tahmini için makine öğrenme algoritmalarının performansını karşılaştırmak	Denetimli makine öğrenmesi algoritması olan için KNN modeli ve yarı denetimli makine öğrenmesi algoritması olarak da Logistic Regression'un en yüksek doğruluğu verdiğini bulmuşlardır.	LR, Gauss Naive Bayes, SVM, RBF SVM, DT, RF, Xgboost, Gradient Boosting, K-NN	98
Vaka vd. (2020)	DNNS	Meme kanserinin erken teşhisi sağlamak	DNNS algoritması iyi bir sonuç vermiştir.	SVM, NB	97,21

2.2.5. Orak Hücre Anemi Hastalığı İçin Ulaşılan Çalışmalar

Orak hücre anemi hastalığına sahip kişiler kırmızı kan hücreleri hasarlı üretirler. Bu hücreler kan damarında sıkışır ve kan akışını engeller. Kan akışını engellemesinden dolayı vücut, organlar ve doku oksijen ihtiyacını karşılayamaz. Orak hücreli anemiye sahip hastaların çoğunda hastalığın bir tedavisi yoktur ancak mevcut tedaviler ile ağrılar hafifletilebilir ve hastalığın yol açabileceği sorunlar önlenir. Bu nedenle bu alanda yapılan çalışmalar çok önem arz etmektedir. Derleme çalışması kapsamında taranan, Orak hücreli anemi için farklı algoritmaların bir araya geldiği makine öğrenmesi çalışmaları, doğruluk oranı en yüksek olan algoritma, ele alınan problem, karşılaştırma yapılan algoritmalar ve sonuçlar bilgileriyle birlikte Tablo 6'da özetlenmiştir. Tablo 6'da NB (Naive Bayes), RF (Random Forest), SVM (Support Vector Machine), DT (Decision Tree), LR (Logistic Regression), K-NN (K Nearest Neighbor), LDA (Linear Discriminant Analysis), ANN (Artificial Neural Network), SMO (Sequential Minimal Optimization), ASC (Attribute selected classifier) algoritmaları bulunmaktadır.

Khalae vd., (2016) çalışma da orak hücreli anemi hastalığının tedavisi için ilaç dozaj seviyelerinin sınıflandırılmasını amaçlamışlardır. Kullanılan veri setinde 12 öznitelik bulunmaktadır. Veri setinde ağırlık, hemogloblin değeri, nötrofiller (beyaz kan hücre) değeri gibi veriler bulunmaktadır. H1, H2, LEVNN ve RFC gibi yaklaşımlarla doğruluk ve performansı en üst düzeye çıkarmak için

makine öğrenmesi algoritmaları kullanmıştır. En iyi performans gösteren algoritma, %98,2 doğruluk payı ile Random Forest (RF) olmuştur.

Abd vd., (2017a) çalışmada orak hücreli anemi hastalığının test edilmesi, takip edilmesi ve izlenmesi hizmetlerini sunan bir model oluşturmayı amaçlamışlardır. İngiltere’de bir hastaneye ait veri seti kullanmışlardır. Toplanan verileri orak hücre anemisine sahip olanlar ve olmayanlar olarak 2 gruba ayırmışlardır. Veri setinin içinde yaş, cinsiyet, hemoglobin değerleri gibi birçok veri yer almasıyla beraber 250 hasta ve toplam 12 öznitelik bulunmaktadır. Önerilen model için makine öğrenmesi algoritmaları uygulamışlardır. Sınıflandırma için Support Vector Machine (SVM), Sequential Minimal Optimization (SMO), Naive bayes (NB), Decision Tree (DT) gibi algoritmalar uygulamışlardır. En iyi performans gösteren makine öğrenmesi algoritması SMO olmuştur.

Chen vd., (2017) çalışmada orak hücreli anemi hastalığı olan ve aynı zamanda inme riski düşük olan çocuklar için yapısal manyetik rezonans (MR) görüntüleme analizi yapmayı hedeflemişlerdir. Veri için; inme geçirme öyküsü olmayan çocuklara 2-4 yıl aralık boyunca Kaufman kısa zekâ (K-BIT) testi uygulamışlardır. Test sonrasında MR görüntüler incelenmiş ve test sonucu düşüş grubu ve test sonucu düşüş olmayan grup olarak 2 grup oluşturmuşlar. Daha sonra, yapısal MR görüntülerinden hesaplanan bölgesel gri madde (GM) hacimlerine dayalı olarak K-BIT düşüş / düşüş olmadığını tahmin etmek için öngörücü modeller oluşturmuşlardır. GM hacmine ve temel K BIT'e dayanarak, K* algoritmasını kullanarak bir prognostik model oluşturmuşlardır. Doğruluk, duyarlılık ve özgüllük oranlarını sırasıyla %84, %78 ve %86 olarak tespit etmişlerdir.

Abd vd., (2017b) çalışmada akıllı telefona yüklenen uygulamaya dayalı olarak hastayı izleyebilen sistem önermeyi amaçlamışlardır. Sistemi, orak hücre anemi hastalarının kan testi sonucuna göre kritik olup olmadığı kararını verecek şekilde uygulama planlamışlardır. Orak hücre anemi hastalarının sınıflandırılmasını bu sistem sayesinde tamamlamışlardır. Jrip Kuralları, öznitelik seçili sınıflandırıcı, Voted Perceptron, Bayes Net, AdaboostM1, Attribute selected classifier LogitBoost algoritmaları kullanmışlardır. En iyi performans gösteren algoritma %98,2 doğruluk payı ile LogitBoost olmuştur.

Tablo 6. Orak hücre anemi Alanında Yapılan Makine Öğrenmesi Çalışmaları.

Yazar ve Yayın Yılı	Algoritma	Ele Alınan Problem	Sonuçlar	Karşılaştırma Yapılan Algoritmalar	Doğruluk (%)
Abd vd. (2017a)	SMO	Orak hücreli anemi hastalığı için ilaç tahmin etmek	Orak hücreli anemi hastalığını tahmin etmişlerdir.	SVM, SMO,NB, DT	99
Chen vd. (2017)	K*	Orak hücreli anemi hastalığının olan ve aynı zamanda inme riski düşük olan çocuklar için MR görüntüleme analizi sağlamak	K BIT'e dayanarak, K* algoritmasını kullanarak bir prognostik model oluşturulmuştur.	-	84
Abd vd. (2017b)	LogitBoost	Hastayı izleyebilen sistemi önermek	En iyi performans gösteren algoritma LogitBoost olarak belirlenmiştir.	Jrip Kuralları,ASC, Voted Perceptron, BayesNet, AdaboostM1, LogitBoost	98,2
Khalae vd. (2016)	RF	Orak hücreli anemi hastalığının tedavisi için ilaç dozaj seviyelerini sınıflandırmak	En iyi performans gösteren algoritmanın RF olduğu belirlenmiştir.	H1, H2, LEVNN ve RF, TREEC	98,2

2.2.6. Kalp Durması Alanında Yapılan Makine Öğrenmesi Çalışmaları

Kalbin pompa görevini yerine getirmemesine kalp durması denir. Kalp durması kalp krizinden farklıdır. Kalp krizi genellikle kalp damalarında biriken yağ sonucunda gösterdiği tepkidir. Kalp durması ani bir şekilde olur ve beklenmedik anda gerçekleşir. Bu nedenle bu alanda yapılan çalışmalar çok önem arz etmektedir. Derleme çalışması kapsamında taranan, Kalp durması için farklı algoritmaların bir araya geldiği makine öğrenmesi çalışmaları, doğruluk oranı en yüksek olan algoritma, ele alınan problem, karşılaştırma yapılan algoritmalar ve sonuçlar bilgileriyle birlikte Tablo 7’de özetlenmiştir. Tablo 7’de NB (Naive Bayes), RF (Random Forest), SVM (Support Vector Machine), DT (Decision Tree), LR (Logistic Regression), K- NN (K Nearest Neighbo), ANN (Artificial Neurol Network), GNB (Gaussian Naive Bayes) algoritmaları bulunmaktadır.

Chauhan vd., (2019) kalp durması ihtimalini öngörmek için ve farklı makine öğrenme algoritmalarını incelemeyi amaçlamışlardır. Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), Logistic Regression ve Artificial Neural Networks (ANN), algoritmalarını hastalarda kalp durması oluşumunu tahmin etmek için veri setine uyguladıktan sonra en iyi performans gösteren algoritmanın Artificial Neural Networks (ANN) olduğu ve doğruluğunun en yüksek %85 olduğunu bulmuşlardır.

Chang vd., (2019) çalışmada, kalp durmasını tespit etmek için 2 görev oluşturmuşlardır. Birinci görev de alıcı çalışma karakteristiğinin altındaki alan (AUROC) eğrisini kullanmışlardır, ikincinde ise kesinlik geri çağırma eğrisi altındaki alan için makine öğrenmesi algoritmaları kullanılmıştır. Statik veriler yaş, cinsiyet, boy, kilo, ateş gibi değerler yer alırken, dinamik verilerde ortalama arter basıncı, sistolik kan basıncı, nabız, solunum hızı, vücut ısısı vb. hayati belirtiler yer almaktadır. Naive Bayes (NB), Support Vector Machine (SVM), Random Forest (RF) algoritmalarını kullanmışlardır. En iyi performans gösteren algoritmanın Random Forest (RF) olduğunu bulmuşlardır.

Kwon vd., (2019) çalışmada, kalp durması yaşayan hastaların taburcu olduktan sonra sağkalım analizini yapmışlardır. Hastalıktan iyileşmeyi ve taburculuk için sağkalımı tahmin etmek için derin öğrenmeye dayalı hastane dışı kardiyak arest prognostik sistemi (DCAPS) geliştirmeyi ve onaylamayı amaçlamıştır. Doğrulama verilerinde nörolojik iyileşmeyi tahmin etmek için DCAPS'nin alıcı çalışma karakteristik eğrisinin (AUROC) altındaki alan %95,3, Logistic Regression %94,7, Random Forest (RF) %94,3, Support Vector Machine (SVM) %93,0 ve önceki bir çalışmanın geleneksel yöntemlerinden %81,7 doğruluk oranına ulaşılmıştır. DCAPS, OHCA hastalarının doğru bir şekilde taburcu edilmesi için nörolojik iyileşme ve sağkalımı öngörmüşlerdir ve geleneksel yöntem ve diğer makine öğrenme yöntemlerinden daha iyi performans gösterdiğini belirlemişlerdir.

Javan vd., (2019) çalışma da yetişkin hastalar için enfeksiyona bağlı olarak yaşanan kalp durmalarını tahmin etmeyi amaçlamışlardır. Ek olarak yaşamsal belirtilerin zaman serisi dinamiklerinin kalp durmanın öngörülmesine etkisini araştırmışlardır. Kullanılan makine öğrenmesi modelleri Support Vector Machine (SVM), Decision Tree (DT), Logistic Regression K Nearest Neighbor (K-NN), Gaussian Naive Bayes, Multilayer Perceptron Classifiers, XGBoost, Random Forest (RF) gibi sınıflandırıcıları içermektedir. En iyi performans gösteren algoritmanın Random Forest (RF) olduğunu bulmuşlardır.

Hirano vd., (2021) çalışmada hastanın kalp durmasını tahmin modeli geliştirmeyi hedeflemişlerdir. Japonya’da hastane dışı 43,350 kalbi duran hasta verisi toplanmış olup 18 yaşından küçük hastalar veya bir etkene bağlı kalbi duran hastalar dahil edilmemiştir. Logistic

Regression, Support Vector Machine (SVM), Random Forest (RF), ve Multilayer Perceptron Classifiers algoritmalarının sonuçları karşılaştırılmıştır. Sonuçlar, %95 güven aralığında olup %5 hata payı ile değerlendirilmiştir. Doğrulama analizlerinde, Support VectorMachine (SVM) için %86,6, Random Forest (RF) için %87,7 ve Multiple Linear Perceptron (MLP) sınıflandırıcılar için %88,8 doğruluk payı belirlenmiştir. En iyi tahmincinin Multilayer Perceptron Classifiers olduğu belirlenmiştir.

Safa vd., (2021) çalışma sıcaklık, kan basıncı, nabız ve strese bağlı kalp hastalarının verilerinden oluşmaktadır. Kalp hastalarının stres analizini hesaplamayı hedeflemişlerdir. Stres analizi için tahmin modeli, Decision Tree (DT), K Nearest Neighbor (K-NN) ve Support Vector Machine (SVM) algoritmalarını eğitmişlerdir. Hastalarının verilerine ön işleme için SQL kullanmışlardır. Kalp hastalığı veri seti üzerinde algoritmaları test etmişlerdir. Sonuç, K Nearest Neighbor (K-NN) sınıflandırmasının etkinliğinin SVM ve DT yöntemlerine kıyasla çok daha iyi olduğunu göstermektedir.

Tablo 7. Kalp Durması Alanında Yapılan Makine Öğrenmesi Çalışmaları.

Yazar ve Yayın Yılı	Algoritma	Ele Alınan Problem	Sonuçlar	Karşılaştırma Yapılan Algoritmalar	Doğruluk (%)
Chauhan vd. (2019)	ANN	Kalp durması olasılığını tahmin etmek	En iyi performans gösteren algoritmanın Artificial Neural Networks (ANN), olduğunu bulmuşlardır.	SVM, RF, DT,LR, ANN	85
Chang vd. (2019)	RF	Kalp durması olasılığını tahmin etmek	En iyi performans gösteren algoritmanın RF olduğunu bulmuşlardır.	NB, SVM, RF	88
Hirano vd. (2021)	MLP	Kalp durması olasılığını tahmin etmek	En iyi performans gösteren algoritma MLP algoritması bulmuşlardır.	LR, SVM, RF,MLP	88
Safa vd. (2021)	K-NN	Kalp hastalarının stresini analiz etmek	En iyi performans gösteren algoritmanın K-NN algoritması olduğunu bulmuşlardır.	K-NN, SVM, DT	85,98
Kwon vd. (2019)	DCAPS	Kalp durması yaşayan hastaların taburcu olduktan sonra sağkalım analizi yapmak	Hastalarının doğru bir şekilde taburcu edilmesi için nörolojik iyileşme ve sağkalımı öngörmüşlerdir.	LR, RF, SVM	97,6
Javan vd. (2019)	RF	Enfeksiyona bağlı olarak yaşanan kalp durmalarını tahmin etmek	En iyi performans gösteren algoritma RF olduğunu bulmuşlardır.	SVM, DT, LR, K-NN, GNB, Gradient boosting, XGBoost, RF	95

2.2.7. Tiroid Hastalığı Alanında Yapılan Makine Öğrenmesi Çalışmaları

Tiroid bezi boynun ön bölümünde yer alan kelebek şeklinde bir organdır. Tiroid bezinin sahip olduğu hormonlar, kan dolaşımına girerek metabolizmayı düzenler, hızını kontrol eder. Tiroid hastalığı insan hayatında büyük risk oluşturmaktadır. Bu nedenle bu alanda yapılan çalışmalar çok önem arz etmektedir. Derleme çalışması kapsamında taranan, Tiroid hastalığı için farklı algoritmaların bir araya geldiği makine öğrenmesi çalışmaları, doğruluk oranı en yüksek olan algoritma, ele alınan problem, karşılaştırma yapılan algoritmalar ve sonuçlar bilgileriyle birlikte Tablo 8’de özetlenmiştir. Tablo 8’de NB (Naive Bayes), RF (Random Forest), SVM (Support

Vector Machine), DT (Decision Tree), K-NN (K-En Yakın Komşu), ANN (Artificial Neural Network), ELM (Extreme Learning Machine), BPNN (Back Propagation Neural Network), MLP (Multiple Linear Perceptron) algoritmaları bulunmaktadır.

Pavya vd., (2017) tiroid hastalığının erken teşhisini tanımlamak için makine öğrenmesi algoritmalarından yararlanmışlardır. Filtre tabanlı (F-Skoru) ve sarmalayıcı tabanlı (Yinelemeli Özellik Eliminasyonu) özellik seçim algoritmalarının hastalık tanımlama ve sınıflandırma üzerindeki etkisini analiz etmeyi amaçlamışlardır. Performans değerlendirmesi doğruluk, duyarlılık ve özgüllük olmak üzere üç ölçü ile yapmışlardır. Deneysel sonuçlar, hem F-Skoru hem de Yinelemeli Özellik Eliminasyonunun tiroid hastalığı teşhisinin performansını iyileştirirken, sarmalayıcı tabanlı algoritmanın Extreme Learning Machine sınıflandırıcı ile maksimum verimlilik ürettiğini ve maksimum %98,14 doğruluk ürettiğini göstermiştir.

Duggal vd., (2018) tiroid bezinin iki yaygın hastalığı olan hipertiroidizm ve hipotiroidizm hastalıklarını ele almışlardır. Tiroid hastalıklarını dört sınıfa ayırmışlardır. Bunlar; hipotiroid bezi, hipertiroid bezi, hasta ötiroid, ötiroid. Bu dört hastalığını sınıflandırmayı hedeflemişlerdir. Veri Kümesi UCI makine öğrenmesi tabanlı olup 7200 hasta ve 27 öznitelik bulunmaktadır. Naive bayes (NB), Support Vector Machine (SVM), Random Forest (RF), algoritmalarını eğitmişlerdir. Bu teknikler arasından en iyi performans doğruluğuna sahip olan Support Vector Machine (SVM) sonuç vermiştir.

Tyagi vd., (2018) tiroid hastalığının erken teşhisini tanımlamak için makine öğrenmesi algoritmalarından yararlanmışlardır. Tiroid hasta bilgilerinin yer aldığı veri setini kullanmışlardır. Veri setinde yaş, cinsiyet, tiroid hormon değerleri gibi değerler bulunmaktadır. Support Vector Machine (SVM), K Nearest Neighbor (K-NN) Decision Tree (DT) gibi birçok makine öğrenmesi algoritmalarını, tiroid hastalığına yakalanma riski olanları tahmin etmek için kullanmışlardır. En iyi sonuç veren algoritma SVM olup %99,63 doğruluğuna sahiptir.

Olatunji vd., (2021) tiroid hastalığının çok erken aşamalarda (pre-semptomatik aşama) hastalığı tespit edecek bir sistem kurmayı planlamışlardır. Suudi Arabistan King Fahad hastanesinden alınan veri içerisinde kan hastalıkları olanlar dahil olmak üzere tiroid hastaları bulunmaktadır. Veri setinde normal hasta ve tiroid hastaları olmak üzere 2 sınıf bulunmaktadır. 14 nitelik içeren bu veride toplam 109 tiroid hastası 109 normal hasta bulunmaktadır. Bu çalışmada Random Forest (RF), Artificial Neural Networks (ANN), Support Vector Machine (SVM) ve Naive Bayes (NB) algoritmalar kullanılmıştır. Elde edilen en yüksek doğruluk oranı RF tekniği ile %90,91 olurken, SVM, YSA ve NB sırasıyla %84,09, %88,64 ve %81,82 doğruluk elde etmiştir.

Tablo 8. Tiroid Hastalığı Alanında Yapılan Makine Öğrenmesi Çalışmaları.

Yazar ve Yayın Yılı	Algoritma	Ele Alınan Problem	Sonuçlar	Karşılaştırma Yapılan Algoritmalar	Doğruluk (%)
Duggal vd. (2018)	SVM	Tiroid bezinde yaygın olan hastalığı sınıflandırmak	En iyi sonuç veren algoritma SVM olmuştur.	NB, SVM, RF	92,92
Tyagi vd. (2018)	SVM	Tiroid hastalığını tahmin etmek	En iyi sonuç veren algoritma SVM olmuştur.	SVM, K-NN, DT	99,63
Pavya vd. (2017)	ELM	Tiroid hastalığını tahmin etmek	En iyi sonuç veren algoritma ELM olmuştur.	MLP, SVM, BPNN, ELM	98,14
Olatunji vd. (2021)	RF	Tiroid hastalığını tahmin etmek	En iyi sonuç veren algoritma RF olmuştur.	RF, ANN, SVM, NB	90,91

2.2.8. Alzheimer Hastalığı Alanında Yapılan Makine Öğrenmesi Çalışmaları

Alzheimer hastalığı ileri yaşı hastalığıdır. Beynin bazı bölümleri zaman içinde giderek hasar görmeye başlar bunun sonucunda başta bellek olmak üzere tüm günlük faaliyetleri ve davranış bozukluğu gösteren bir hastalıktır. Derleme çalışması kapsamında taranan, Alzheimer hastalığı için farklı algoritmaların bir araya geldiği makine öğrenmesi çalışmaları, doğruluk oranı en yüksek olan algoritma, ele alınan problem, karşılaştırma yapılan algoritmalar ve sonuçlar bilgileriyle birlikte Tablo 9’da özetlenmiştir. Tablo 9’ da NB (Naive Bayes), RF (Random Forest), SVM (Support Vector Machine), DT (Decision Tree), K-NN (K Nearest Neighbor) ANN (Artificial Neural Network), MLP (Multilayer Perceptron Classifiers), ELM (Extreme Learning Machine), GPR (Extreme Learning Machine), PLS (partial least squares), SVM-DA (Support Vector Machine Discriminant Analysis) algoritmalar bulunmaktadır.

Lodha vd., (2018) Alzheimer hastalarının beyin görüntülerini analiz etmeyi amaçlamışlardır. Nöropsikolojik ve objektif değerlendirmelerinin aralarında ilişki olup olmadığını araştırmışlardır. Hastalığı erken teşhis etmek için makine öğrenmesi algoritmalarından yararlanmışlardır. Vishwakarma Teknoloji Enstitüsü ait veri tabanındaki Alzheimer hasta verileri kullanmışlardır. Artificial Neural Networks (ANN), Support Vector Machine (SVM), Random Forest (RF) algoritmalarını kullanmışlardır. En yüksek doğruluk oranına sahip algoritma Artificial Neural Networks (ANN) algoritması olmuştur ve doğruluğu %98,36 bulmuşlardır.

Zhang vd., (2019) Alzheimer olan hastaların erken teşhisini tanımlamayı amaçlamışlardır. Makine öğrenmesi algoritmalarından yararlanmışlardır. Extreme Learning Machine (ELM), Support Vector Machine (SVM), Extreme Learning Machine (GPR) ve partial least squares (PLS) modellerinin performansını değerlendirmek için 10 kat çapraz doğrulama kullanmışlardır. Nöro görüntüleme girişimi'ndeki Alzheimer hastalığı verileri içerisinde klinik bilgilerin yanı sıra Voxel-Bazlı Morfometri (VBM) parametreleri ve doku parametreleri kullanmışlardır. Önerilen yöntemleri Alzheimer 'li 58 hastadan ve 94 normal kontrollü verilere uygulamışlar ve ELM modelinin tüm sınıflandırma özellikleriyle %96'ya varan bir sınıflandırma doğruluğu elde ederken, diğer üç modelin sonuçları %82 (PLS), %79 (GPR) ve %75 (SVM). Alzheimer hastaları normal kontrollü hastalarla ayırmada optimal sonuç elde etmişler ve bu nedenle bu çalışma Alzheimer 'nin teşhisi için faydalı olabilir.

Neelavani vd., (2020) Alzheimer hastalarının yaş, ziyaret sayısı, MMSE (Mini Mental Test) ve eğitim gibi psikolojik parametreleri kullanarak Alzheimer hastalığını erken teşhis etmek için makine öğrenmesi algoritmalarını seçmişlerdir. MMSE skoru, yaş, ziyaret sayısı ve hastaların eğitimi gibi hastalığı tahmin etmede büyük ölçüde yardımcı olan önemli parametreler kullanmışlardır. SVM ve DT algoritmalarını %70 eğitim %30 test veri kümesiyle test etmişlerdir. En iyi performans gösteren algoritma SVM algoritması olmuştur ve doğruluğu %85 bulmuşlardır.

Khan vd., (2020) analiz için MRI (Manyetik Rezonans Görüntüleme) beyin görüntülerinin Açık Erişim Görüntüleme Çalışmaları Serisi (OASIS) veri tabanına ait verileri kullanmışlardır. Veri seti, 150 denek içeren 343 MRI seansından oluşmaktadır. Alzheimer teşhisi için boyamsal beyin MRG verilerini kategorize etmek için demanslı ve demansız olmak üzere iki sınıfa ayırmayı amaçlamışlardır. Analizde MMSE (Mini Mental Test), CDR (Clinical Dementia Rating) ve ASF (Atlas Scaling Factor) olmak üzere üç farklı skor kullanılmıştır. Önerilen makine öğrenmesi ardışık düzeni, deneysel ve veri analizi tasarımına gömülü olan veri dönüştürme ve özellik seçme teknikleriyle birlikte bir sınıflandırıcı sistemi oluşturmuşlardır. En iyi performans gösteren algoritmanın Random Forest (RF), ağının doğruluğunun en yüksek %86,84 olduğunu bulmuşlardır.

Rangaswamy vd., (2020) Alzheimer hastalığını teşhis etmek bunun için iyi makine öğrenmesi algoritmasını bulmayı amaçlamışlardır. Genom çapında ilişkilendirme çalışmaları (GWAS) ve Genotip-Doku İfadesi (GTEX) veri tabanlarından 20,401 zararlı ve 37,452 kontrol seti oluşturmuşlardır. Önemli özellikleri seçmek için çapraz doğrulama (RFECV) kullanılarak özellik eleme ve ardından ileri özellik seçme yöntemi kullanılmış ve zararlı ve nötr varyantları ayırt etmek için Random Forest (RF), sınıflandırıcısı kullanılmıştır. Random Forest (RF), algoritmasının doğruluğunu %81,21 bulmuşlardır.

Uysal vd., (2020) Alzheimer hastalığı (AD), hafif bilişsel bozukluk (MCI) ve bilişsel normal (CN) hastalarını birbirinden ayırt etmek için kullanılabileceği sistem önermişlerdir. Kurulan model için 482 hasta için veri 160 test set ve 322 eğitim set olarak ayırmışlardır. Tahmin performansları, tüm tanı grupları arasında sınıflandırma yapılarak değerlendirmişler. Her modelde sol hipokampus hacim değeri sağ hipokampustan daha başarılı tahmin sonuçları verdiği için, sağ hipokampus değerinin dahil edilmediği yaş ve cinsiyet parametreleri eklenerek yeni bir alan incelemişler ve bu durum da başarı oranının arttığını kanıtlamışlardır. %80 doğrulukla KNN iyi bir performans göstermiş ve sağ hipokampus hacim değeri dahil edilmediğinde ise %82 doğrulukla GNB iyi performans gösteren algoritma arasında yer almış. Yaş ve cinsiyet her algoritma için olumlu etkilere sahip bir sonuç çıkmıştır. Tüm bulgular değerlendirildiğinde hipokampus hacim bilgisi ile anlamlı tanı konulabileceği, yaş ve cinsiyet arasındaki hipokampal hacim farklılıklarının tanı kararında etkili olduğu sonucuna varmışlardır.

Ryzhikova vd., (2021) makine öğrenmesi algoritmaları beyin omurilik sıvısına dayalı Alzheimer Hastalığını teşhis etmek için yeni bir yöntem geliştirmişlerdir. Alzheimer Hastalığı teşhisi konmuş hastaların veri setini kullanmışlardır. Bu veri seti için 21 hastadan ve 16 sağlıklı denekten beyin omurilik sıvısı almışlardır. Artificial Neural Networks (ANN), ve Support Vector Machine Discriminant Analysis, (SVM-DA) istatistiksel yöntemleri farklılaştırma amaçları için kullanıldı ve en başarılı sonuçlar %84 duyarlılık ve özgüllük ile Alzheimer Hastalığı ve sağlıklı kontrol denekler konularının farklılaşmasını gözlemlemişlerdir.

Tablo 9. Alzheimer Hastalığı Alanında Yapılan Makine Öğrenmesi Çalışmaları.

Yazar ve Yayın Yılı	Algoritma	Ele Alınan Problem	Sonuçlar	Karşılaştırma Yapılan Algoritmalar	Doğruluk (%)
Lodha vd. (2018)	ANN	Alzheimer hastalarının beyin görüntülerini analiz etmek	En iyi performans gösteren algoritma ANN olmuştur.	SVM, RF	98,36
Neelavani vd. (2020)	SVM	Alzheimer hastalığını erken teşhis etmek	En iyi performans gösteren algoritma SVM olmuştur.	DT	85
Zhang vd. (2019)	ELM	Alzheimer hastalığını erken teşhis etmek	En iyi performans gösteren algoritma ELM olmuştur.	SVM, GPR, PLS	96
Khan vd. (2020)	RF	Alzheimer teşhisi için beyin MRG verilerini kategorize etmek ve demanslı ve demansız olmak üzere iki sınıfa ayırmak	En iyi performans gösteren algoritma RF olmuştur.	ExtraTree, DT, NuSVC, AdaBoost, GradientBoosting, GNB, Ridge, K-NN	86,84
Rangaswamy vd. (2020)	RF	Alzheimer hastalığını teşhis etmek	En iyi performans gösteren algoritma RF olmuştur.	XgBoost, AdaBoost, ANN,	81,21
Uysal ve Öztürk (2020)	GNB ve K-NN	Alzheimer Hastalığı, Hafif Bilişsel Bozukluk ve Bilişsel Normal hastalarını birbirinden ayırt etmek	Alzheimer hastalığının bilgisayar destekli teşhisinin performansını artırdığını ortaya koymuştur.	K-NN, LR, SVM, DT, RF, GNB	80 / 82
Ryzhikova vd. (2021)	SVM-DA ve ANN	Alzheimer Hastalığını teşhis etmek için yeni bir yöntem geliştirmek	Alzheimer Hastalığı ve sağlıklı kontrol denekler konularının farklılaşmasını gözlemlemişlerdir.	ANN, SVM-DA	84

2.2.9. Kronik Böbrek Hastalığı Alanında Yapılan Makine Öğrenmesi Çalışmaları

Böbrek zaman içerisinde işlevini ve fonksiyonlarını kaybetmesine kronik böbrek yetmezliği veya kronik böbrek hastalığı denmektedir. Böbreğin işlevlerini zaman içerisinde kaybetmesi insan yaşamını riske sokmaktadır. Bu alanda yapılan çalışmalar çok fazla önem arz etmektedir. Derleme çalışması kapsamında taranan, kronik böbrek hastalığı için farklı algoritmaların bir araya geldiği makine öğrenmesi çalışmaları doğruluk oranı en yüksek olan algoritma, ele alınan problem, karşılaştırma yapılan algoritmalar ve sonuçlar bilgileriyle birlikte Tablo 10'da özetlenmiştir. Tablo 10'da RF (Random Forest), SVM (Support Vector Machine), DT (Decision Tree), Logistic Regression ANN (Artificial Neural Network), MLP (Multilayer Perceptron Classifiers), Gaussian NB (Gaussian Naive Bayes), K Nearest Neighbor (K-NN), LD (Linear Discriminant Analysis) algoritmalar bulunmaktadır.

Rabby vd., (2019) kronik böbrek hastalığını teşhis edebilmek için makine öğrenimi algoritmalarını kullanmışlardır. Bir bölümde eğitim ve test olarak veri setini bölmüşlerdir. Birçok makine öğrenmesi algoritmalarını kullanmışlardır. Yüksek doğruluk oranına sahip algoritma Decision Tree (DT), Random Forest (RF), GNB algoritması olmuştur. %100 doğruluk gösteren bu algoritmalar çalışmadaki kronik böbrek hastalarını sınıflandırabilmektedir.

Almansour vd., (2019) çalışmanın amacı kronik böbrek hastalarını erken teşhis etmektir. 400 hastadan oluşan bir veri kümesine ve kronik böbrek hastalığının teşhisi ile ilgili 24 özelliğe makine öğrenmesi algoritmaları uygulayarak karşılaştırma yapmışlardır. Artificial Neural Networks (ANN) ve Support Vector Machine (SVM) algoritmalarıyla analizi tamamlamışlardır. Uygulamalardan elde edilen ampirik sonuçlara göre ANN algoritmasının doğruluğu %99,75 ve SVM algoritmasının doğruluğu %97,75 çıkmıştır. Her iki algoritmanın sonuçları iyi olmasına rağmen yüksek doğruluk oranına sahip algoritma Artificial Neural Networks (ANN) algoritması olmuştur.

Yashf vd., (2020) kronik böbrek hastalığı erken teşhisini araştırmışlardır. Kronik böbrek hastalığına yakalanma riski olan hastaları belirlenebileceği bir sistem oluşturmuşlardır. 445 kronik böbrek hastası verilerini kullanmışlardır. UCI makine öğrenmesi veri tabanından erişilen bu veri gerçek zamanlı bir veridir. Random Forest (RF), %97,12 doğruluğa sahip, Artificial Neural Networks (ANN) %94,5 doğruluğa sahip çıkmıştır. Bu çalışmanın sonucunda yüksek doğruluk olasılığına sahip algoritma Random Forest (RF) algoritması olmuştur.

He vd., (2021) böbrek hasarı, karaciğer transplantasyonundan sonra sık görülen komplikasyondur ve kötü prognoz göstergesidir. Beyin ölüm sonrası karaciğer transplantasyonu (DCDLT) yapılan toplam 493 hasta verisi kullanmışlardır. Akut böbrek hasarı olan hastalar ile olmayan hastaları karşılaştırmışlardır. Decision Tree (DT), Logistic Regression, Random Forest (RF), Support Vector Machine (SVM) algoritmaları ile analiz etmişlerdir. Yüksek doğruluk oranına sahip algoritma Random Forest (RF) olarak belirlenmiştir.

Tablo 10. Kronik Böbrek Hastalığı Alanında Yapılan Makine Öğrenmesi Çalışmaları.

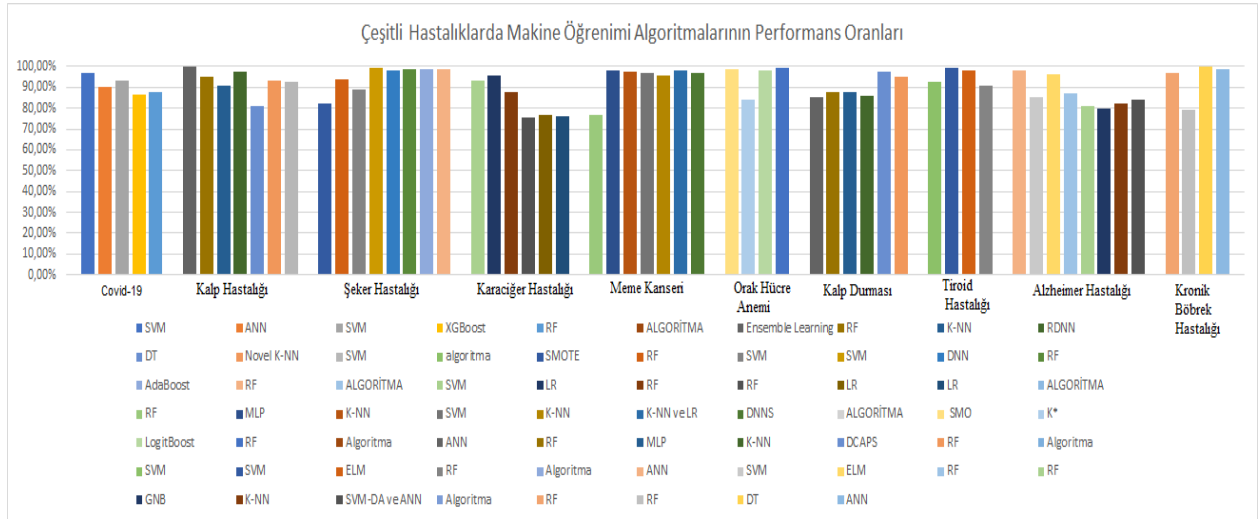
Yazar ve Yayın Yılı	Algoritma	Ele Alınan Problem	Sonuçlar	Karşılaştırma Yapılan Algoritmalar	Doğruluk (%)
Yashf vd. (2020)	RF	Kronik böbrek hastalığı erken teşhis etmek	En iyi performans gösteren algoritma RF algoritması olmuştur.	RF, ANN	97,12
He vd. (2021)	RF	Kronik böbrek hastalığının durumunu etkili bir şekilde tanıyabilen ve tahmin edebilen sistem oluşturmak	En iyi performans gösteren algoritma RF algoritması olmuştur.	LR, RF, SVM, DT	79
Rabby vd. (2019)	DT RF, GNB	Kronik böbrek hastalığı erken teşhis etmek	En iyi performans gösteren algoritma DT, RF, GNB algoritması olmuştur.	DT, GNB, RF, SVM, RF, K-NN, AdaBoost, LD, Gradient boosting, ANN, LR	100
Almansour vd. (2019)	ANN	Kronik böbrek hastalığı erken teşhis etmek	En iyi performans gösteren algoritma ANN algoritması olmuştur.	ANN, SVM	99

3. Sonuç

Bu çalışmanın amacı, sağlık araştırmalarında kullanılan makine öğrenimi algoritmalarını karşılaştırarak en iyi performans gösteren algoritmayı keşfetmektir. Bu kapsamda son 10 yıl içerisinde sağlık alanında gerçekleştirilen makine öğrenmesi çalışmaları özetlenmiş ve sınıflandırılmıştır. Yapılan ayrıntılı literatür çalışması ile halk sağlığını tehdit eden ve dünyada ölüm nedenleri listesinde ilk sıralarda yer alan hastalıklar için uygulanan makine öğrenmesi çalışmalarına odaklanılmıştır.

Ayrıca dünyanın en ölümcül hastalıklar listesinde yer alan ve son yıllarda halk sağlığı için acil durum ilan edilen COVID-19 hastalığına da yer verilmiştir. Hastalıklar için yapılan çalışmalarda erken teşhis, hastalık sınıflandırma, ilaç sınıflandırma, salgını önleme gibi çalışmaların sağlık sektörüne büyük katkılar sağladığı açıktır. Bu çalışmalar hastaların ölüm riskini azaltmakla birlikte yaşam kalitesini artırmaktadır.

Ancak, faaliyet alanı ve mimarisi nedeniyle, sistemin karmaşıklığına uygun bir tahmin ve talep karar modeli seçmek zordur (Tozan vd., 2018). Çalışmanın, bu alanda çalışan araştırmacılara ışık tutarak ilgili literatüre önemli katkısı olacağı düşünülmektedir. Çalışma kapsamında makine öğrenmesi algoritmalarının performans sonuçları farklı hastalıklar üzerinde özetlenmiştir (Şekil 2).



Şekil 2: Farklı Hastalıklarda Makine Öğrenmesi Algoritmalarının Performansı

Sağlık hizmetlerinde en iyi performans gösteren sınıflandırma algoritması ortalama %100 doğruluk başarıyla Decision Tree (DT), Random Forest (RF), Gaussian Naive Bayes (GNB) olmuştur.

Çalışma sonucunda, makine öğrenmesi algoritmalarının en çok hastalıkları teşhis etme ve hastalıkları sınıflandırma konularında yaygın kullanıldığı görülmektedir.

Kaynakça

1. Abd, D., Alwan, J. K., Ibrahim, M., & Naeem, M. B. (2017a). The utilisation of machine learning approaches for medical data classification and personal care system mangement for sickle cell disease. *2017 Annual Conference on New Trends in Information and Communications Technology Applications, NTICT 2017, March*, 213–218.
2. Abd, D. H., & Al-Mejibli, I. S. (2017b). Monitoring System for Sickle Cell Disease Patients by Using Supervised Machine Learning. *2017 2nd Al-Sadiq International Conference on Multidisciplinary in IT and Communication Science and Applications, AIC-MITCSA 2017*, 119–124.
3. Ahmed, H., Younis, E. M. G., Hendawi, A., & Ali, A. A. (2020). Heart disease identification from patients' social posts, machine learning solution on Spark. *Future Generation Computer Systems*, 111, 714–722.
4. Al-Azzam, N., & Shatnawi, I. (2021). Comparing supervised and semi-supervised Machine Learning Models on Diagnosing Breast Cancer. *Annals of Medicine and Surgery*, 62(December 2020), 53–64.
5. Almansour, N. A., Syed, H. F., Khayat, N. R., Altheeb, R. K., Juri, R. E., Alhiyafi, J., Alrashed, S., & Olatunji, S. O. (2019). Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study. *Computers in Biology and Medicine*, 109(April), 101–111.
6. Alves, M. A., Castro, G. Z., Oliveira, B. A. S., Ferreira, L. A., Ramírez, J. A., Silva, R., & Guimarães, F. G. (2021). Explaining machine learning based diagnosis of COVID-19 from routine blood tests with decision trees and criteria graphs. *Computers in Biology and Medicine*, 132(March).
7. Arul Jothi, K., Subburam, S., Umadevi, V., & Hemavathy, K. (2021). Heart disease prediction system using machine learning. *Materials Today: Proceedings*, 1–3.
8. Asri, H., Mousannif, H., Al Moatassime, H., & Noel, T. (2016). Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. *Procedia Computer Science*, 83(Fams).
9. Balaji, K., Lavanya, K., & Mary, A. G. (2020). Machine learning algorithm for clustering of heart disease and chemoinformatics datasets. *Computers and Chemical Engineering*, 143, 107068.
10. Bayat, V., Phelps, S., Ryono, R., Lee, C., Parekh, H., Mewton, J., Sedghi, F., Etmnani, P., & Holodniy, M. (2020). A Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Prediction Model From Standard Laboratory Tests. *Clinical Infectious Diseases*, 1–7.
11. Chang, H. K., Wu, C. T., Liu, J. H., Lim, W. S., Wang, H. C., Chiu, S. I., & Jang, J. S. R. (2019). Early detecting in-hospital cardiac arrest based on machine learning on imbalanced data. *2019 IEEE International Conference on Healthcare Informatics, ICHI 2019*, 1–10.
12. Chauhan, U., Kumar, V., Chauhan, V., Tiwary, S., & Kumar, A. (2019). Cardiac Arrest Prediction using Machine Learning Algorithms. *2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies, ICICICT 2019, Cvd*, 886–890.
12. Chen, R., Krejza, J., Arkuszewski, M., Zimmerman, R. A., Herskovits, E. H., & Melhem, E. R. (2017). Brain morphometric analysis predicts decline of intelligence quotient in children with sickle cell disease: A preliminary study. *Advances in Medical Sciences*, 62(1), 151–157.
14. Faiyaz Waris, S., & Koteeswaran, S. (2021). Heart disease early prediction using a novel machine learning method called improved K-means neighbor classifier in python. *Materials Today: Proceedings*, 1–7.
15. Ferdous, M., Debnath, J., & Chakraborty, N. R. (2020). Machine Learning Algorithms in Healthcare: A Literature Survey. *2020 11th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2020*.
16. Gambhir, E., Jain, R., Gupta, A., & Tomer, U. (2020). Regression Analysis of COVID-19 using Machine Learning Algorithms. *Proceedings- International Conference on Smart Electronics and Communication, ICOSEC 2020, Icosec*, 65–71.
17. Gatos, I., Tsantis, S., Spiliopoulos, S., Karnabatidis, D., Theotokas, I., Zoumpoulis, P., Loupas, T., Hazle, J. D., & Kagadis, G. C. (2017). A Machine-Learning Algorithm Toward Color Analysis for Chronic Liver Disease Classification, Employing Ultrasound Shear Wave Elastography. *Ultrasound in Medicine and Biology*, 43(9), 1797–1810.
18. Gogi, V. J., & Vijayalakshmi, M. M. (2018). Prognosis of Liver Disease: Using Machine Learning Algorithms. *2018 International Conference on Recent Innovations in Electrical, Electronics and Communication Engineering, ICRIEEECE 2018*, 875–879.
19. Gopal, V. N., Al-Turjman, F., Kumar, R., Anand, L., & Rajesh, M. (2021). Feature Selection and Classification in Breast Cancer Prediction using IoT and Machine Learning. *Measurement*, 178(February), 109442.
20. Grampurohit, S., & Sagarnal, C. (2020). Disease prediction using machine learning algorithms. *2020 International Conference for Emerging Technology, INCET 2020, November*.

21. He, Z. L., Zhou, J. Bin, Liu, Z. K., Dong, S. Y., Zhang, Y. T., Shen, T., Zheng, S. Sen, & Xu, X. (2021). Application of machine learning models for predicting acute kidney injury following donation after cardiac death liver transplantation. *Hepatobiliary and Pancreatic Diseases International*, 1–10.
22. Hirano, Y., Kondo, Y., Sueyoshi, K., Okamoto, K., & Tanaka, H. (2021). Early outcome prediction for out-of-hospital cardiac arrest with initial shockable rhythm using machine learning models. *Resuscitation*, 158(August), 49–56.
23. Jashwanth Reddy, D., Mounika, B., Sindhu, S., Pranayteja Reddy, T., Sagar Reddy, N., Jyothsna Sri, G., Swaraja, K., Meenakshi, K., & Kora, P. (2020). Predictive machine learning model for early detection and analysis of diabetes. *Materials Today: Proceedings*.
24. Kalender, Z. T., Tozan, H., & Vayvay, O. (2020). Prioritization of medical errors in patient safety management: Framework using interval-valued intuitionistic fuzzy sets. *Healthcare (Switzerland)*, 8(3).
25. Karadayı, M. A., Gökmen, Y. G., Kasap, L. G., & Tozan, H. (2019). Sağlıkta Güncel Simülasyon Yaklaşımları: Bir Derleme Çalışması. *International Journal of Advances in Engineering and Pure Sciences*, 1–21.
26. Karthikeyan, B., Gollamudi, S., Singamsetty, H. V., Gade, P. K., & Mekala, S. Y. (2020). Breast cancer detection using machine learning. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(2), 981–984.
27. Khalaf, M., Hussain, A. J., Keight, R., Al-Jumeily, D., Keenan, R., Fergus, P., & Idowu, I. O. (2016). The utilisation of composite machine learning models for the classification of medical datasets for sickle cell disease. *2016 6th International Conference on Digital Information Processing and Communications, ICDIPC 2016*, 37–41.
28. Khan, A., & Zubair, S. (2020). An Improved Multi-Modal based Machine Learning Approach for the Prognosis of Alzheimer's disease. *Journal of King Saud University- Computer and Information Sciences*.
29. Kwon, J. myoung, Jeon, K. H., Kim, H. M., Kim, M. J., Lim, S., Kim, K. H., Song, P. S., Park, J., Choi, R. K., & Oh, B. H. (2019). Deep-learning-based out-of-hospital cardiac arrest prognostic system to predict clinical outcomes. *Resuscitation*, 139(March 2019), 84–91.
30. Layeghian Javan, S., Sepehri, M. M., Layeghian Javan, M., & Khatibi, T. (2019). An intelligent warning model for early prediction of cardiac arrest in sepsis patients. *Computer Methods and Programs in Biomedicine*, 178, 47–58.
31. Li, J. P., Haq, A. U., Din, S. U., Khan, J., Khan, A., & Saboor, A. (2020). Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare. *IEEE Access*, 8(MI), 107562–107582.
32. Kafes, M. (2018). Tip 2 Diyabet Mellitus Hastalarının Cepten Sağlık Odeme Maliyetlerinin Belirlenmesi üzerine Bir Araştırma (Doctoral dissertation, Necmettin Erbakan University (Turkey)).
33. Lodha, P., Talele, A., & Degaonkar, K. (2018). Diagnosis of Alzheimer's Disease Using Machine Learning. *Proceedings- 2018 4th International Conference on Computing, Communication Control and Automation, ICCUBE 2018*, 1–4.
34. Lukmanto, R. B., Suharjito, Nugroho, A., & Akbar, H. (2019). Early detection of diabetes mellitus using feature selection and fuzzy support vector machine. *Procedia Computer Science*, 157, 46–54.
35. Karadayı, M. A. , Yılmaz, B. Ö. , Erol, B. E. & Tozan, H. (2020). Sağlık Teknolojisi Değerlendirmede Çok Kriterli Karar Verme Yaklaşımları Üzerine Bir Derleme Çalışması . *Düzce Üniversitesi Bilim ve Teknoloji Dergisi* , 8 (1) , 264-289 .
36. Mujumdar, A., & Vaidehi, V. (2019). Diabetes Prediction using Machine Learning Algorithms. *Procedia Computer Science*, 165, 292–299.
37. Mung, P. S., & Phyu, S. (2020). Effective Analytics on Healthcare Big Data Using Ensemble Learning. *2020 IEEE Conference on Computer Applications, ICCA 2020*.
38. Neelaveni, J. (2020). "Alzaeimer Disease Prediction using Machine Learning Algorithm", 6th International Conference on Advanced Computing & Communication Systems (I.C.A.C.C.S.).
39. Nguyen, B. P., Pham, H. N., Tran, H., Nghiem, N., Nguyen, Q. H., Do, T. T. T., Tran, C. T., & Simpson, C.
40. R. (2019). Predicting the onset of type 2 diabetes using wide and deep learning with electronic health records. *Computer Methods and Programs in Biomedicine*, 182(August).
41. Olatunji, S. O., Alotaibi, S., Almutairi, E., Alrabae, Z., Almajid, Y., Altabee, R., Altassan, M., Basheer Ahmed, M. I., Farooqui, M., & Alhiyafi, J. (2021). Early diagnosis of thyroid cancer diseases using computational intelligence techniques: A case study of a Saudi Arabian dataset. *Computers in Biology and Medicine*, 131(February), 104267.
42. Öztürk, N., Tozan, H., & Vayvay, Ö. (2016). *Comprehensive Needs Analysis For Health Technology Assessment Studies and Improvement Proposal*. 1(1), 69–76.
43. Öztürk, N., Tozan, H., & Vayvay, Ö. (2020). A new decision model approach for health technology

- assessment and a case study for dialysis alternatives in Turkey. *International Journal of Environmental Research and Public Health*, 17(10).
45. Nadesh, R. K., & Arivuselvan, K. (2020). Type 2: diabetes mellitus prediction using deep neural networks classifier. *International Journal of Cognitive Computing in Engineering*, 1, 55-61.
 46. Pavya, K., & Srinivasan, B. (2017). Feature selection algorithms to improve thyroid disease diagnosis. *IEEE International Conference on Innovations in Green Energy and Healthcare Technologies - 2017, IGEHT 2017*, 1-5.
 47. Pourhomayoun, M., & Shakibi, M. (2021). Predicting mortality risk in patients with COVID-19 using machine learning to help medical decision-making. *Smart Health*, 20(April 2020), 100178.
 48. Rabby, A. K. M. S. A., Mamata, R., Laboni, M. A., Ohidujjaman, & Abujar, S. (2019). Machine Learning Applied to Kidney Disease Prediction: Comparison Study. *2019 10th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2019*, 1-7.
 49. Rangaswamy, U., Dharshini, S. A. P., Yesudhas, D., & Gromiha, M. M. (2020). VEPAD - Predicting the effect of variants associated with Alzheimer's disease using machine learning. *Computers in Biology and Medicine*, 124(July), 103933.
 50. Ryzhikova, E., Ralbovsky, N. M., Sikirzhyski, V., Kazakov, O., Halamkova, L., Quinn, J., Zimmerman, E.A., & Lednev, I. K. (2021). Raman spectroscopy and machine learning for biomedical applications: Alzheimer's disease diagnosis based on the analysis of cerebrospinal fluid. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 248, 119188.
 51. Safa, M., & Pandian, A. (2021). Applying machine learning algorithm to sensor coupled IoT devices in prediction of cardiac stress – An integrated approach. *Materials Today: Proceedings*.
 52. Saha, P., Sadi, M. S., & Islam, M. M. (2021). EMCNet: Automated COVID-19 diagnosis from X-ray images using convolutional neural network and ensemble of machine learning classifiers. *Informatics in Medicine Unlocked*, 22, 100505.
 53. Sarkar, J. P., Saha, I., Sarkar, A., & Maulik, U. (2021). Machine learning integrated ensemble of feature selection methods followed by survival analysis for predicting breast cancer subtype specific miRNA biomarkers. *Computers in Biology and Medicine*, 131(January), 104244.
 54. Shi, S., Lei, G., Yang, L., Zhang, C., Fang, Z., Li, J., & Wang, G. (2021). Using Machine Learning to Predict Postoperative Liver Dysfunction After Aortic Arch Surgery. *Journal of Cardiothoracic and Vascular Anesthesia*, 000.
 55. Sridhar, G. M., & Prema Kirubakaran, A. (2021). Heart disease and optimal prediction of attacks using hybrid machine learning algorithm: A survey. *Materials Today: Proceedings*.
 56. Srivenkatesh, M. (2019). Performance Evolution of Different Machine Learning Algorithms for Prediction of Liver Disease. *International Journal of Innovative Technology and Exploring Engineering*, 9(2), 1115-1122.
 57. Sun, P., Lu, X., Xu, C., Sun, W., & Pan, B. (2020). Understanding of COVID-19 based on current evidence. *Journal of Medical Virology*, 92(6), 548-551.
 58. SunilaGodara, 2Sanjeev Kumar. (2018). *Prediction of Thyroid Disease Using Machine Learning Techniques*. 10(2), 787-793.
 60. Thaiparnit, S., Chumuang, N., & Ketcham, M. (2018). A Comparative Study of Classification Liver Dysfunction with Machine Learning. *2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing, ISAI-NLP 2018- Proceedings*, 283, 1-4.
 61. Tigga, N. P., & Garg, S. (2020). Prediction of Type 2 Diabetes using Machine Learning Classification Methods. *Procedia Computer Science*, 167(2019), 706-716.
 62. Tozan, H., Karatas, M., & Vayvay, O. (2018). Reducing demand signal variability via a quantitative fuzzy grey regression approach. *Tehnicki Vjesnik*, 25(September), 411-419.
 63. Tyagi, A., Mehra, R., & Saxena, A. (2018). Interactive thyroid disease prediction system using machine learning technique. *PDGC 2018- 2018 5th International Conference on Parallel, Distributed and Grid Computing*, 689-693.
 64. Uysal, G., & Ozturk, M. (2020). Hippocampal atrophy based Alzheimer's disease diagnosis via machine learning methods. *Journal of Neuroscience Methods*, 337(February), 108669.
 65. Vaka, A. R., Soni, B., & K., S. R. (2020). Breast cancer detection by leveraging Machine Learning. *ICT Express*, 6(4), 320-324.
 66. Vilorio, A., Herazo-Beltran, Y., Cabrera, D., & Pineda, O. B. (2020). Diabetes Diagnostic Prediction Using Vector Support Machines. *Procedia Computer Science*, 170, 376-381.
 67. WHO. (2018). The-Top-10-Causes-of-Death @ Wwww.Who.Int. In *The top 10 causes of death* (p. Consultado 23 de marzo de 2019). <https://www.who.int/es/news-room/fact-sheets/detail/the-top-10-causes->

of-death

68. Winter, G. (2019). Machine learning in healthcare. *British Journal of Health Care Management*, 25(2), 100–101.
69. Wu, C. C., Yeh, W. C., Hsu, W. D., Islam, M. M., Nguyen, P. A. (Alex), Poly, T. N., Wang, Y. C., Yang, H.C., & (Jack) Li, Y. C. (2019). Prediction of fatty liver disease using machine learning algorithms. *Computer Methods and Programs in Biomedicine*, 170, 23–29.
70. Wu, J., & Hicks, C. (2021). Breast cancer type classification using machine learning. *Journal of Personalized Medicine*, 11(2), 1–12.
71. Yashfi, S. Y., Islam, M. A., Pritilata, Sakib, N., Islam, T., Shahbaaz, M., & Pantho, S. S. (2020). Risk Prediction of Chronic Kidney Disease Using Machine Learning Algorithms. *2020 11th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2020*.
72. Zhang, F., Tian, S., Chen, S., Ma, Y., Li, X., & Guo, X. (2019). Voxel-Based Morphometry: Improving the Diagnosis of Alzheimer's Disease Based on an Extreme Learning Machine Method from the ADNI cohort. *Neuroscience*, 414, 273–279.