



Salgın Hastalıklarla Mücadelede Açık Kaynak Kodlu Çözümler

Melike Bektaş^{a,1}, Abdullah Yavuz^{a,2}, Faruk Bulut^{a,3}

^a Istanbul Rumeli Üniversitesi, Mühendislik ve Mimarlık Fakültesi, Bilgisayar Mühendisliği Bölümü, İstanbul, Türkiye

Istanbul Sabahattin Zaim Üniversitesi Fen Bilimleri Enstitüsü Dergisi (2021) 3 (1): 99-105

<https://doi.org/10.47769/izufbed.861541>

¹0000-0002-1944-1928; ²0000-0001-5950-7269; ³0000-0003-2960-8725

YAYIN BİLGİSİ

Yayın geçmişi:

Gönderilen tarih: 14 Ocak 2021

Kabul tarihi: 15 Nisan 2021

Anahtar kelimeler:

Covid-19

İstatistiksel Analizler

Sınıflandırma

Kümeleme

İlişkilendirme

Görüntü İşleme

ÖZET

İnsanlık tarihi boyunca salgın hastalıklar birçok can kaybına neden olmuştur. Bilgi teknolojileri ve Endüstri 4.0 çağında bu hastalıklarla mücadelenin farklı boyutları vardır. Tıbbi yaklaşımlar, kimyevi çözümler, laboratuvar çalışmaları elbette bu işin en önemli boyutu ve olmazsa olmazdır. Bunun yanında istatistik, matematik ve veri bilimi ile elde edilecek analizler, çıkarımlar ve öngörüler, salgın hastalıklar ile mücadelede önemli bir rol oynamaktadır. Bu alanda açık kaynak kodlu yazılımlar ve çözümlerle, salgın hastalıklarla daha iyi bir mücadele sergilenebilmektedir. Farklı algoritmik yaklaşımları içeren açık kaynak kodlu yazılımlar özgür geliştiricilerin desteği ile daha da ileri seviyelere götürülebilmektedir. Ayrıca bu tür yazılımlar ülkelere ve bölgelere göre özgülendirilebilir. Bu çalışmada, salgın hastalıklarla mücadelede kullanılan istatistiksel ve veri bilimi yöntemlerinin açık kaynak kodlu yazılımlarda nasıl kullanıldığı kategorilere ayrılarak incelenmiştir.

Open Source Based Solutions in Combating Epidemics

ARTICLE INFO

Article history:

Received: 14 January 2021

Accepted: 15 April 2021

Key words:

Covid-19

Statistical Analysis

Classification

Clustering

Association

Image processing

ABSTRACT

Epidemics have caused numerous casualties throughout human history. There are different dimensions of combating these types of diseases in the age of information technologies and Industry 4.0. Medical approaches, chemical solutions, laboratory studies are of course the most important aspect of this subject. In addition, analyzes, inferences and predictions obtained from statistics, mathematics and data science play crucial roles in combating epidemics. In this field, a better fight against epidemics can be conducted with open source software and their solutions. Open source software containing different algorithmic approaches can be taken to further levels with the support of free developers. In addition, such software can be customized according to countries and regions. In this study, statistical and data science methods in combating epidemic diseases in the open source software field are examined and categorised.

1. Giriş

Devrilen bir domino taşı, sırasıyla diğer domino taşlarını düşürür ve büyük bir kaos oluşturur. Benzer bir şekilde kanat çırpın bir kelebeğin oluşturduğu hava dalgası çok uzaklarda bir kasırganın çıkmasına ve büyük bir afet yaşanmasına neden olabilir (Vaidyanathan S., Azar A. T., 2016). Salgın hastalıklar, domino etkisi ve kelebek etkisi gibi kaos teoreminin birer örneğidir. Ütopik veya hamasi bir yaklaşım olarak değerlendirilen bu teoremlerin son birkaç ayda nasıl bir gerçekliğe dönüştüğünü bu gezegende yaşayan insanlar olarak gözlemledik. 2019 yılının Aralık ayından beri tüm dünya ülkelerini etkisi altına alan COVID-19 kod adlı Corona virüsünün yayılması ile birlikte Dünya Sağlık Örgütü tarafından

pandemi ilan edilmiştir. Özellikle tıbbi alanda bu virüsün vücutta yok edilmesi amacıyla birçok çalışma yapılmıştır ve yapılmaya da devam etmektedir.

Salgın hastalıklara neden olan virüsler ve etkileri hakkında elde edilen ham verilerden malumat ve bilgi çıkarımı yapılabilir. Hastalığa neden olabilen birçok faktör içerisinde en etkilisi tespit edilebilir. Hastalar farklı niteliklerine göre kümelenebilir ve değişik çıkarımlar elde edilebilir. Bunun yanında dağılım analizleri, hipotez testleri ve korelasyonlar da bu alanda değişik farkındalıklar sağlayabilir.

Bu çalışmada amaç, salgın hastalıkların oluşturacağı etkileri azaltmak için görüntü işleme, istatistik ve veri bilimi alanında şu ana

kadar yapılmış olan açık kaynak kodlu çalışmalarını bir araya getirerek bir farkındalık oluşturmaktır. Tek çatı altına alınıp gruplandırılan bu çalışmalar hem bir kaynak taraması hem de mücadelede amacıyla yapılan disiplinler arası çalışmalara ışık tutacak nitelikte olacaktır. Açık kaynak kodlu çözümlerin hem ücretsiz olması hem de gönüllü kişi ve kuruluşlarla geliştirilmeye açık olması gibi avantajları vardır. Diğer taraftan ticari yazılımların genellikle belirli bir şirket ve kuruluş çalışmaları tarafından geliştiriliyor olması, ilgili yazılımın veya projenin gelişim sürecinin daha yavaş olmasına neden olabilir. Esasen istatistik, görüntü işleme ve veri bilimi disiplinleri kendi içlerinde oldukça derin ve detaylı konulara sahiptir. Fakat bu çalışmada, meseleyi belirli bazı perspektiflerden ele alarak daha nicel bir değerlendirme ve raporlama imkânı sunmak istenmiştir.

Bu çalışma dört bölüme ayrılmıştır. Yukarıdaki ilk bölümde giriş, ikinci bölümde konu ile alakalı literatür çalışmalarına, üçüncü bölümde istatistik ve veri bilimindeki açık kaynak kodlu çözümlere ve son olarak da sonuç kısmına yer verilmiştir.

2. Literatür Taraması

Corona virüsünün tüm dünyada yayılması ile birlikte bilim camiasında son aylarda ciddi bir çalışma baş göstermiştir. Bilim insanlarının bu alanda yaptıkları değerli çalışmalar başka çalışmalara ışık göstermesi açısından literatürde yerini almıştır. İstatistik ve veri bilimi alanında yapılmış olan bu çalışmalardan en değerli bulduklarımızı bu bölümde inceleyeceğiz.

Abdulmunem ve arkadaşları (2020), hasta akciğer X-ray görüntülerine eşikleme tekniği uygulayarak hastaların COVID-19 olup olmadığını tespit eden bir metot önermişlerdir. Önerilen metot akciğerin sağlıklı bölge ile hastalıklı bölgenin birbirinden ayrılması ve hastalıklı bölgenin yerini belirlemek için kullanılmıştır.

Yıldırım ve Fındık (2020), çalışmalarında Türkiye'deki COVID-19 vaka sayılarındaki değişimleri öngörebilmek için Uzun-Kısa Süreli Bellek (LSTM) tabanlı bir sinir ağı yaklaşımı ile ileriye yönelik vaka sayısı tahmininde bulunabilen bir çalışma yapmışlardır. Yapılan çalışmanın sonucunda, LSTM algoritmasının önleyici adımlar atabilmek için vaka tahmininde önemli bir avantaj sağlayabileceğini göstermiştir.

Randhawa ve arkadaşları (2020), makine öğrenimi bileşenine bir karar ağacı yaklaşımı ve Spearman'in korelasyon katsayısı hesaplamasını adapte ederek taksonomik bir sınıflandırma yapmışlardır. Genom analizleri için dijital sinyal işlemeli denetimli makine öğrenimini (MLDSP: machine learning with digital signal processing) kullanılmıştır.

Elaziz ve arkadaşları (2020), yaptıkları bir çalışmada, göğüs röntgeni görüntülerini COVID-19 hastası veya COVID-19 hastası olmayan kişi olarak iki sınıfa ayırmak için yeni bir makine öğrenimi yöntemi önermişlerdir. Hesaplama sürecini hızlandırmak için paralel birçok çekirdekli hesaplama çerçevesi kullanmışlardır. Ardından, en önemli özellikleri seçmek için kullanılan evrime dayalı değiştirilmiş bir Manta-Ray Toplama Optimizasyonu yöntemi önermişler ve çalışmada iki farklı veri seti kullanmışlardır. Önerilen yöntem ile birinci ve ikinci veri seti için sırasıyla % 96.09 ve % 98.09 doğruluk oranlarına ulaşmışlardır.

Barstugan, Özkaya ve Öztürk (2020), COVID-19'un makine öğrenimi yöntemleri kullanarak erken faz tespiti üzerine çalışma gerçekleştirmişlerdir. Tespit işlemini abdominal Bilgisayarlı Tomografi (BT) görüntüleri üzerinde uygulamışlardır. Özellik çıkarma yöntemleri olarak Gri Seviye Eş Oluşum Matrisi (GLCM), Yerel Yön Modeli (LDP), Gri Seviye Çalışma Uzunluğu Matrisi (GLRLM), Gri Seviye Boyutlu Bölge Matrisi (GLSZM) ve Ayrık

Dalgacık Dönüşümü (DWT) algoritmalarını kullanmışlardır. Sınıflandırma sürecinde 2'li, 5'li ve 10'lu çapraz doğrulama uygulamışlardır. Sınıflandırma performansını değerlendirmek için duyarlılık, doğruluk, kesinlik ve F-skor ölçütleri kullanılmıştır. En iyi sınıflandırma doğruluğunu, 10 kat çapraz doğrulama ve GLSZM öznelik çıkarma yöntemi ile % 99.68 olarak elde ettiklerini belirtmişlerdir.

Akgül ve arkadaşları (2021), çalışmalarında Konvolüsyonel Sinir Ağlarını kullanarak insan resimlerinde maske olup olmadığını tespit etmeye çalışmışlardır. Simulated Masked Face Dataset (SMFD) veri setini kullanarak çalışmalarını test etmişlerdir. Çalışma sonucunda %97.07 doğruluk başarısına ulaştıklarını belirtmişlerdir.

Özen ve arkadaşları (2021), çalışmalarında makine öğrenimi algoritmalarını kullanarak ABD'nin gelecek günlerdeki COVID-19 vaka sayılarını tahmin etmeye çalışmışlardır. Çalışma sonucunda Polinom Regresyon algoritması ile %1.86 ortalama mutlak yüzde hata oranı elde etmişlerdir.

Wang ve arkadaşları (2021), tipik viral pnömoni teşhisi konmuş olanlar ile birlikte patojen onaylı COVID-19 vakalarının olduğu BT görüntüleri üzerinde çalışma yapmışlardır. Dahili doğrulamada 0,87 hassasiyet ve 0,88 özgüllük kullanarak %89,5 doğruluk başarısına ve harici veri setinde 0,67 hassasiyet ve 0,83 özgüllük kullanılarak %79,3 doğruluk başarısına ulaştıklarını bildirmişlerdir.

Sedik ve arkadaşları (2021), BT ve X-ray resimlerinden oluşan veri setlerine, Konvolüsyonel Sinir Ağları ve Konvolüsyonel Uzun-Kısa Süreli Bellek metodlarını birleştirerek uygulamıştır. Bazı durumlarda %100 doğruluk ve %100 F1 skoruna ulaştıklarını açıklamışlardır.

3. Açık Kaynak Kodlu Çözümler

Bu çalışmada salgın hastalıklarla ilgili mücadele üç ana kategoride incelenmiştir. İlki temel istatistik yöntemler, ikincisi ise yapay zekânın alt dalları olan makine öğrenmesi, veri madenciliği ve derin öğrenme alanlarında sıklıkla kullanılan sınıflandırma, öbekleme (kümeleme) ve ilişkilendirme yöntemleridir. Üçüncüsü de görüntü işleme ile akciğer tomografisinden COVID-19'un tespit edilmesi ile ilgilidir.

İstatistik ve veri bilimi alanında sıklıkla kullanılan bazı açık kaynak kodlu genel yazılımları içeren platformlar mevcuttur. Öncelikle bunları şu şekilde listelemek istiyoruz.

Weka bu alanda sıklıkla tercih edilen görsel içerikleri zengin Java tabanlı ve açık kaynak kodlu bir ortamdır. Temel istatistiksel yöntemlerin tamamını barındırdığı gibi Veri Bilimi alanında kullanılan neredeyse tüm algoritmaları da içermektedir. Grafik ve görsel çıktılar ile de sonuçların daha net ve doğru bir şekilde değerlendirilmesine olanak tanımaktadır. Verilerin kullanılabilir hale getirilebilmesi, sınıflandırma, kümeleme ve ilişkilendirme yöntemleri sekmeler halinde mevcuttur. COVID-19 ile ilgili hazırlanmış ARFF, CSV veya TXT gibi birçok formattaki dosya türleri Weka'da rahatlıkla açılıp üzerinde çalışmalar yapılabilir.

MATLAB programlama ortamında hazır olarak gelen birçok istatistik, veri bilimi ve görüntü işleme yöntemleri de salgın hastalıklar alanında hazırlanmış veri setlerini işlemek için kullanılabilir. Hatta ihtiyaç duyulduğunda www.mathworks.com gibi açık kaynak kodu paylaşan sitelerden de basit bir üyelik aşaması ile bu alandaki farklı ve geliştirilmiş yazılımlar temin edilebilir.

MATLAB, Python, Java, RapidMiner ve R gibi programlama dilleri ve ortamları, açık kaynak kodlu yazılımları geliştirme imkânı sunmaktadır. www.github.com gibi sitelerden de elde edilebilecek

yazılımlar bu gibi programlama dilleri ortamlarında rahatlıkla geliştirilebilmektedir.

KNIME (Konstanz Information Miner), Konstanz Üniversitesi'nde Java programlama dili kullanılarak geliştirilmiş açık kaynak kodlu veri ön işleme, istatistiki veri analizi ve interaktif grafik görüntüleme gibi işlemleri yapmak için kullanılan bir programdır (Tekerek A., 2011).



Şekil 1. Açık kaynak kodlu veri bilimi kütüphaneleri

Burada belirtmek istediğimiz önemli bir konu da istatistik, makine öğrenmesi, veri madenciliği ve görüntü işleme alanında hazırlanmış açık kaynak kodlu yazılımların bolluğu ve çeşitliliği hakkında farkındalık oluşturulmasıdır. Şekil 1'de makine öğrenmesi, veri madenciliği ve görüntü işleme alanlarında kullanılan popüler açık kaynak kodlu kütüphaneler gösterilmektedir.

Bu kütüphanelerden önemli olan bazılarını sırayla açıklamak gerekirse TensorFlow, Google tarafından geliştirilmiş 2015 yılında kullanıma açılmış olan ücretsiz, açık kaynak kodlu, derin öğrenme ve yapay zekâ uygulamaları için kullanılan bir kütüphanedir. Bu kütüphane birçok programlama dilinde rahatlıkla kullanılabilir.

Bu kütüphanelerden önemli olan bazılarını sırayla açıklamak gerekirse TensorFlow, Google tarafından geliştirilmiş 2015 yılında kullanıma açılmış olan ücretsiz, açık kaynak kodlu, derin öğrenme ve yapay zeka uygulamaları için kullanılan bir kütüphanedir. Bu kütüphane birçok programlama dilinde rahatlıkla kullanılabilir.

Pandas ise csv, txt, xls gibi farklı formatlarda bulunan verilerin okunması ve işlenmesi için kullanılan açık kaynak kodlu bir kütüphanedir.

Keras Python programlama dili kullanılarak François Chollet tarafından geliştirilmiş kolay bir şekilde yapay sinir ağları ve derin öğrenme modelleri oluşturmak ve modelleri eğitmek için kullanılan açık kaynak kodlu bir kütüphanedir.

OpenCV Intel tarafından geliştirilmiş açık kaynak kodlu gerçek zamanlı görüntü işleme kütüphanesidir. OpenCV Kütüphanesi C, C++, Matlab, Python programları dilleri ile birlikte kullanılabilir. NumPy, dizi ve matris işlemleri için kullanılan açık kaynak kodlu Python programlama dilinin desteklediği bir kütüphanedir.

Scikit-Learn, sınıflandırma, kümeleme, regression gibi veri analizleri için kullanılan açık kaynak kodlu makine öğrenmesi kütüphanesidir.

Pytorch Facebook tarafından geliştirilen açık kaynak kodlu bir kütüphanedir. Bilgisayarlı görme, derin öğrenme ve doğal dil işleme gibi alanlarda kullanılmaktadır. Python, Java ve C++ programlama dillerinde Pytorch Kütüphanesi kullanılarak geliştirme yapılabilir.

NLTK yapay zekânın alt kategorilerinden biri olan doğal dil işleme uygulamalarında kullanılan açık kaynak kodlu Python kütüphanesidir.

3.1 Temel İstatistiki Yöntemler

Temel istatistik alanındaki çalışmalarını iki ana grupta inceleyebiliriz. Salgın hastalığa yakalanan kişiler bir örneklem uzayına alınır, bu uzay içerisindeki veriler en küçük, en büyük, ortalama, medyan, standart sapma, ortanca, histogram gibi temel ölçütlerle analiz edilebilir. Bireylere ait yaş, kilo, nabız, kan değerleri, idrar tahlilleri gibi sayısal verilerden her biri biraz önce verilmiş olan ölçütlerle analiz edilebilir. Bu sayede belirli bir bölgedeki kişiler genel olarak değerlendirilmiş olur. Bu ölçütler sayesinde yapılan temel analizler, hastalıkla mücadelede önemli faydalar sağlamaktadır. Ayrıca bu verilerden elde edilen çubuk, pasta ve histogram grafikleri de meseleyi görsel anlamda değerlendirmeyi sağlar ve büyük resmin açığa çıkmasına yardımcı olur (Nesteruk I., 2020).



Şekil 2. Türkiye'deki Hasta Yoğunluğunun Kümelenecek Gösterilmesi (Wikipedia, 2020)

Şekil 2'de 2 Aralık 2020 tarihine ait Türkiye'de bulunan Covid-19 hastalarının bölgelere göre yoğunlukları gösterilmektedir. Daha koyu renkler daha fazla hastanın bulunduğu bölgeleri simgelemektedir.

Hastalığın neden olduğu sonuçların coğrafi ve demografik açıdan incelenmesi birçok fayda sağlayabilir. Çubuk ve pasta grafiklerinden farklı bilgiler elde edilebilir.

Ayrıca Ki-Kare, T-Test, McNemar gibi hipotez testleri ile elde edilen verilerin anlamlılıkları, güvenilirlikleri gözlemlenebilir. Sağlık Bakanlığı tarafından bu alandaki güncel veriler düzenli olarak paylaşılmaktadır (Sağlık Bakanlığı, 2020). Paylaşılan veriler ile bu tür istatistiksel çalışmalar rahatlıkla yapılabilir (Roser M. ve ark., 2020).

Pankowski ve arkadaşları (2020), tarafından yapılan bir çalışmada 2020 yılı içindeki salgın hastalık ile ilgili açık kaynak kodlu ve kıyaslamalı çalışmalar listelenmiştir.

Sabic ve arkadaşları (2020), tarafından yapılan başka bir çalışmada R programlama platformu üzerinde yapılan istatistiksel çalışmalar listelenmiştir. Multidisipliner bir araştırma ile PubMed'den toplanan 60500'den fazla tıbbi bibliyografik referans ile COVID-19 hakkında bir veri seti çalışması gerçekleştirilmiştir. Bu pakette sunulan referansların bibliyometrik analizine dayalı kısa bir kullanım örneği de sunulmuştur. Veri setini filtrelemek için ağ analizi kullanılmıştır. Bu çalışma, COVID-19 ile ilgili epidemiyologların araştırmalarını hızlandırmayı amaçlamışlardır.

3.2 Veri Bilimi Yöntemleri

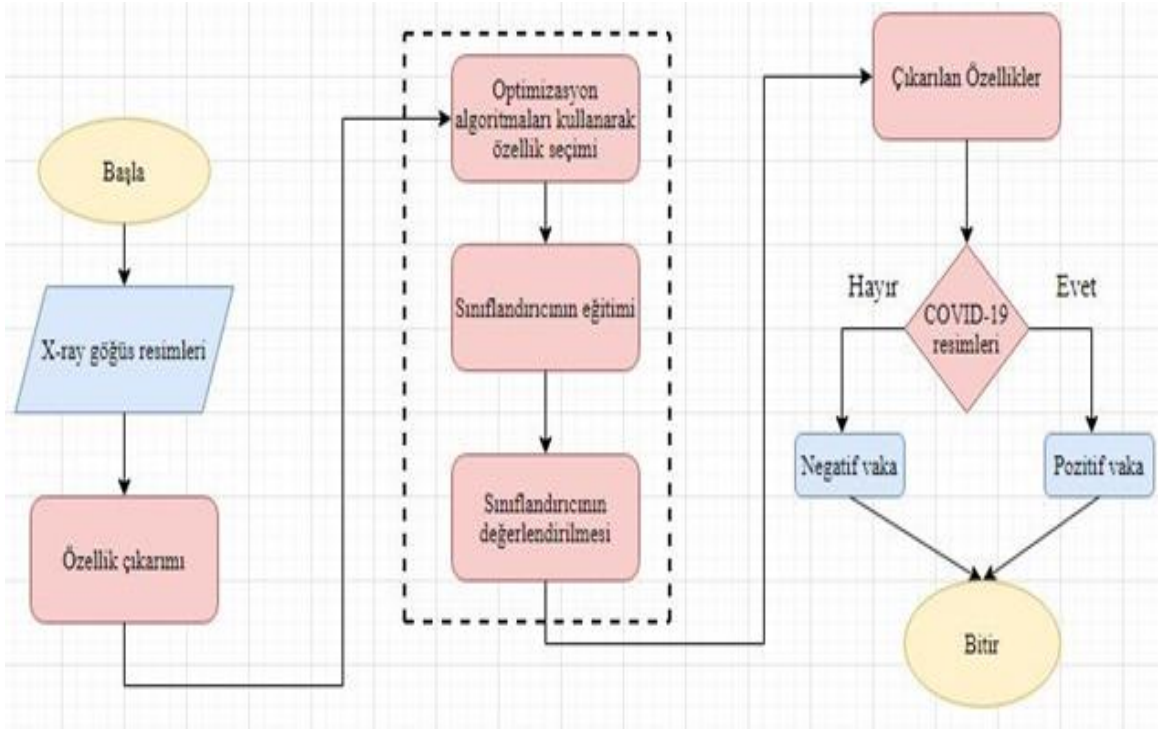
Veri Bilimi (Data Science), yığınla bir araya getirilmiş verilerden malumatın ve bilginin çıkarılması işlemine denir. Simyacılık mesleği ile benzerlik göstermektedir. TeraByte, hatta PetaByte boyutundaki işlenmemiş verilerden çok değerli birkaç cümlelik çıkarımlar ve analizler yapılmaktadır. Veri madenciliği (data mining), şekil tanıma (pattern recognition) ve makine öğrenmesi (machine learning) disiplinlerinin ortak konuları olan sınıflandırma (classification), öbekleme (clustering) ve ilişkilendirme (association) konuları bu çalışmada ele alınacaktır (Yadav M. ve ark., 2020).

Sınıflandırma benzer içeriğe sahip verilerin kategorilere ayrıştırılması işlemidir. Sarıman ve Mutaf (2020), çalışmalarında COVID-19 sürecinde alınan önlem ve hizmete sunulan uygulamalardan oluşan Türkçe tweet mesajlarını maske, eba, sokağa çıkma yasağı, devlet desteği ve kısa çalışma ödeneği olmak üzere 5 başlık altında toplamış ve bu uygulamaların insanlar üzerindeki

etkilerini olumlu ve olumsuz olarak sınıflandırmışlardır. Çalışmalarında Logistic Regression algoritmasını kullanmışlardır. Çalışma sonucunda maske uygulamasının genelde olumlu değerlendirildiği ve diğer uygulamaların ise genelde olumsuz değerlendirildiği sonucuna ulaşmışlardır.

Zhang ve arkadaşları (2020), çalışmalarında, göğüs X-ray görüntülerinden oluşan veri setini kullanarak Viral Pneumonia tespiti gerçekleştirmişlerdir. Çalışmalarında, bir özellik çıkarıcı, bir anormallik algılama modülü ve bir güven tahmin modülünden oluşan, güvene duyarlı anormallik algılama (CAAD) derin öğrenme modeli önermişlerdir. Model başarısı test edildiğinde %83.61 AUC ve %71.70 hassasiyet başarısı elde ettiklerini bildirmişlerdir. Şekil 3'deki akış diyagramında salgın hastalık tespiti için bir sınıflandırma işleminin aşamaları görülmektedir. X-ray göğüs resimlerinden oluşan bir veri seti kullanılarak bir kişinin COVID-19 hastalığına sahip olup olmadığı geliştirilen görüntü işleme temelli hipotez ile tespit edilebilir.

Bütüner ve Calp (2020), çalışmalarında derin öğrenme ve makine öğrenmesi yöntemlerini kullanarak akciğer tomografi görüntülerinden COVID-19 tespiti gerçekleştirmişlerdir. Çalışmalarında Evrişimli Sinir Ağları ve Derin Sinir Ağları ile birlikte makine öğrenmesi algoritması olan K-En Yakın Komşuluk yöntemini kullanmışlardır (Bulut F., 2017). Çalışma sonucunda Evrişimli sinir ağlarının kullanılan diğer algoritmalara kıyasla daha başarılı olduğu gözlemlenmiştir.



Şekil 3. Görüntü işleme temelli bir sınıflandırma işleminin öğrenme ve karar verme fazları



Şekil 4. Çalışmada Kullanılan İlgili Etikete Sahip Göğüs Röntgeni (Brunese ve ark., 2020)

Brunese ve arkadaşları (2020), çalışmalarında, göğüs X-ray görüntülerini COVID-19 tespiti için kullanmışlardır. Çalışmalarında derin öğrenmenin kullanıldığı üç aşamalı bir yöntem geliştirmişlerdir. Birincisi, akciğer grafisinde pnömoni olup olmadığını tespit etmek, ikincisi COVID-19 ile zatürre arasında ayırım yapmak ve son adım ise COVID-19 varlığının X-ışını semptomatikindeki alanların lokalize edilmesi işlemidir. Şekil 4'de çalışmada kullanılan ilgili etikete sahip göğüs röntgeni bulunmaktadır.

Öbekleme bir veri setinde belirli nitelikler açısından benzer özellikler gösteren verilerin gruplandırılması işlemine denir. Bu sayede birbiri ile bazı açılardan benzeşen öğelerin bir arada tutulması ile önemli kestirimler yapılabilmektedir.

Demircioğlu ve Eşiyok (2020), çalışmalarında K-means algoritmasını kullanarak COVID-19 salgını sonuçlarını ülkeler bazında incelemişlerdir. Salgın hastalığın seyrinde benzerlik gösteren ülkeler ve Türkiye'nin bu ülkeler içerisindeki yerini tespit etmeye çalışmışlardır. Çalışmada ülkeler ikili, üçlü ve dördü kümelere bölünmüştür. Singhal ve arkadaşları (2020), vakanın eğilimini belirlemek ve ayrıca önümüzdeki günlerde vakaları tahmin etmek için Gaussian Mixture Model (GMM) kullanarak iki farklı model geliştirmişlerdir. Birincisi, virüsün yayılması ile ilgili çeşitli parametreleri açıklayan matematiksel bir model iken, ikincisi, mevcut verilere uyan Fourier ayrıştırma yöntemine (FDM) dayalı parametrik olmayan bir modeldir.

İlişkilendirme, bir A olayının B olayına etkisini analitik açıdan incelemektedir. Bu alanda bazı hastalıklar salgın hastalıklara zemin hazırlamakta ya da salgın hastalıklar bazı hastalıkların vuku bulmasını tetiklemektedir. Bu tespitlerin yapılabilmesi için ilişkilendirme yöntemlerinden yararlanılmalıdır. Apriori algoritması ilişkilendirme çalışmalarında sıkça kullanılan bir yöntemdir.

Örneğin, ilişkilendirme algoritmaları kullanılarak kronik rahatsızlıkları bulunan kişilerin ve sigara içen kişilerin COVID-19 hastalığına yakalanma oranı incelenmiştir (Garufi ve ark., 2020). Ayrıca, Esper ve arkadaşları (2005), bir çalışmada Kawasaki

hastalığı ile Coronavirüsün verdiği hasarlar arasında ciddi bir ilişki olduğunu gözlemlemiştir. Ayrıca ters yönlü olarak COVID-19'un , kalp rahatsızlıklarını tetiklediği ile alakalı da çalışmalar bulunmaktadır (Bonow ve ark., 2020). Ayrıca bir çalışmada havaalanlarının pandemi etkilerinden korunması ve salgın hastalığın yayılımını durdurmak için yolcuların vücut sıcaklıklarını periyodik olarak termal görüntüleme tekniği kullanarak izlenecek bir sistem önerilmiştir (Kumar ve ark., 2021)

3.3 Görüntü İşleme

Görüntü sürekli görünse bile parçalardan oluşan yapıdadır. Bu parçalara piksel denmektedir. Her görüntü 2 boyutlu piksel dizisi olarak tanımlanabilir. Görüntü işleme, elimizde bulunan görüntülerden anlamlı bir şeyler çıkarmak için kullanılmaktadır. Sağlık sektörü, savunma sanayi, cisim tanıma ve görüntü iyileştirme gibi alanlarda kullanılmaktadır.

COVID-19 ile daha hızlı ve etkin bir mücadele için görüntü işleme teknikleri bu alana uyarlanmış ve 2020 yılı içerisinde bazı değerli çalışmalar gerçekleştirilmiştir. Çok sayıda tıbbi görüntü kaynağı (X-ışını, BT ve MR görüntüleri) ve derin öğrenme tabanlı görüntü işleme yöntemleri COVID-19 salgınıyla mücadelede daha etkin bir rol oynamaktadır. Bu gerçeğin motivasyonu, 2020'nin ilk ayları için çok sayıda araştırma çalışması önerilmiş ve geliştirilmiştir. Sağlık hizmetleri alanında derin öğrenme, diyabetik retinopati tespiti, akciğer nodülü sınıflandırması, fetal lokalizasyon ve tiroid teşhisi gibi birçok uygulamada vardır.

Gozes ve arkadaşları bilgisayarlı akciğer tomografisinden (BT) COVID-19 virüsünün tespit edilebilmesi ile ilgili bir çalışma gerçekleştirmiştir (Gozes ve ark., 2020). Radyologları bu zorlu mücadelede desteklemek için, göğüs BT taramalarından COVID-19 tezahürünün ciddiyetini tespit edebilen, lokalize edebilen ve ölçebilen derin öğrenmeye dayalı bir algoritma geliştirmişlerdir. Algoritma, akciğer segmentasyonu, 2B dilim sınıflandırması ve lokalizasyonu içeren görüntü işleme algoritmalarından oluşan bir açık kaynak kodlu yazılımdır.



Şekil 5. Covid-19'lu bir hastanın akciğerlerine ait X-ışını ve BT görüntüleri (Bhattacharya ve ark., 2020)

Ayrıca bazı çalışmalarda derin öğrenme tabanlı görüntü işleme yöntemleri kullanılarak BT, MR ve X-ışını görüntüleri ile virüs tespiti çalışmaları yapılmıştır (Bhattacharya ve ark., 2020), (Afshar ve ark., 2020), (Elaziz ve ark., 2020). Şekil 5'de Corona virüse yakalanmış bir hastanın akciğerlerine ait X-ışını ve BT görüntüsü verilmektedir. Bu çalışmada BT görüntüsünde bulunan virüs öbekleri görüntü işleme teknikleri ile tespit edilebilmektedir.

Elbetteki bir hekimin tıbbi görüntüleri tespit etmesi daha doğru olan

bir seçimidir fakat görüntü işleme teknikleri ile anomali durumun tespit ve teyit edilmesi iyi bir seçenek olabilir. Bu durumda epidemiyologların da hata yapabilme ihtimalleri göz önüne alındığında, veri bilimi ve görüntü işleme teknikleri ile doğrulama ve geçirme çalışmaları daha fazla bir önem kazanmaktadır. Söz konusu insan hayatı ve sağlığı olunca, tüm bilim dallarının ve disiplinlerin bu alana katkı sağlaması gerekmektedir.

4. Sonuçlar

Aniden beliren ve bilim dünyasının hazırlıksız yakalandığı salgın hastalıklar insanlık için büyük bir tehlikedir. Bu tür hastalıklar hem can kayıplarına, hem vakit kaybına, hem de ciddi oranda maddi kayıplara neden olmaktadır. Yaşadığımız bilgi çağında, salgın hastalıklarla mücadelenin sadece tıbbi ve laboratuvar çalışmalarındaki biyolojik ve kimyevi yaklaşımlarla çözülemeyeceği de bir gerçektir. Bu çalışmada matematik, istatistik, veri madenciliği ve makine öğrenmesi gibi bilim dallarında hazırlanmış olan açık kaynak kodlu çözümlerden faydalanılarak salgın hastalıklarla nasıl daha hızlı ve iyi mücadele edileceği ile ilgili bir farkındalık oluşturulmaya çalışılmıştır.

Kaynaklar

- Abdulmunem, A., A., Metip, M., H., & Abdulhussein, M., F. (2020). Corona Virus Detection Using Image Segmentation Techniques. II. International Conference On Covid-19 Studies.
- Afshar, P., Heidarian, S., Naderkhani, F., Oikonomou, A., Plataniotis, K. N., & Mohammadi, A. (2020). Covid-caps: A capsule network-based framework for identification of covid-19 cases from x-ray images. arXiv preprint arXiv:2004.02696.
- Akgül, İ., Kaya, V., & Baran, A. (2021). Koronavirüse Karşı Yüz Maskesi Tespitinin Derin Öğrenme Yöntemleri Kullanılarak İncelenmesi.
- Barstugan, M., Ozkaya, U., & Ozturk, S. (2020). Coronavirus (covid-19) classification using ct images by machine learning methods. arXiv preprint arXiv:2003.09424.
- Bhattacharya, S., Maddikunta, P. K. R., Pham, Q. V., Gadekallu, T. R., Chowdhary, C. L., Alazab, M., & Piran, M. J. (2020). Deep Learning and Medical Image Processing for Coronavirus (COVID-19) Pandemic: A Survey. Sustainable cities and society, 102589.
- Bonow, R. O., Fonarow, G. C., O'Gara, P. T., & Yancy, C. W. (2020). Association of coronavirus disease 2019 (COVID-19) with myocardial injury and mortality. JAMA cardiology.
- Brunese, L., Mercaldo, F., Reginelli, A., & Santone, A. (2020). explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays. Computer Methods and Programs in Biomedicine, 196, 105608.
- Bulut, F. (2017). A new clinical decision support system with instance based ensemble classifiers. Journal of the Faculty of Engineering and Architecture of Gazi University, 32(1), 65-76.
- Bütüner, R., & Calp, M. H. (2020). COVID-19 Detection from Lung Tomography Images Using Deep Learning and Machine Learning Methods (No. 4097).
- Demircioğlu, M., & Eşiyok, S. Covid-19 Salgını İle Mücadelede Kümeleme Analizi İle Ülkelerin Sınıflandırılması. İstanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi, 19(37), 369-389.
- Elaziz, M. A., Hosny, K. M., Salah, A., Darwish, M. M., Lu, S., & Sahlol, A. T. (2020). New machine learning method for image-based diagnosis of COVID-19. Plos one, 15(6), e0235187.
- Elaziz, M. A., Hosny, K. M., Salah, A., Darwish, M. M., Lu, S., & Sahlol, A. T. (2020). New machine learning method for image-based diagnosis of COVID-19. Plos one, 15(6), e0235187.
- Esper, F., Shapiro, E. D., Weibel, C., Ferguson, D., Landry, M. L., & Kahn, J. S. (2005). Association between a novel human coronavirus and Kawasaki disease. The Journal of infectious diseases, 191(4), 499-502.
- Garufi, G., Carbognin, L., Orlandi, A., Tortora, G., & Bria, E. (2020). Smoking habit and hospitalization for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)-related pneumonia: The unsolved paradox behind the evidence. European Journal of Internal Medicine.
- Gozes, O., Frid-Adar, M., Sagie, N., Zhang, H., Ji, W., & Greenspan, H. (2020). Coronavirus detection and analysis on chest ct with deep learning. arXiv preprint arXiv:2004.02640.
- JR, D. K., Kumar, I. J., & Kiran, J. S. (2021). Smart Surveillance System for Controlling Pandemic Diseases in Airports.
- Nesteruk, I. (2020). Statistics-based predictions of coronavirus epidemic spreading in mainland China.
- Özen, N. S., Saraç, S., & Koyuncu, M. (2021). COVID-19 Vakalarının Makine Öğrenmesi Algoritmaları ile Tahmini: Amerika Birleşik Devletleri Örneği. Avrupa Bilim ve Teknoloji Dergisi, (22), 134-139.
- Pankowski, R., & Dziegielewska, J. (2020). COVID-19 crisis and hate speech: Transnational report.
- Randhawa, G. S., Soltysiak, M. P., El Roz, H., de Souza, C. P., Hill, K. A., & Kari, L. (2020). Machine learning using intrinsic genomic signatures for rapid classification of novel pathogens: COVID-19 case study. Plos one, 15(4), e0232391.
- Roser, M., Ritchie, H., Ortiz-Ospina, E., & Hasell, J. (2020). Coronavirus disease (COVID-19)—Statistics and research. Our World in data.
- Sabic, I., Dodig, L., & Luić, L. (2020, November). COVID-19 IMPACT ON THE USE OF ICT TOOLS IN THE CROATIAN EDUCATION SYSTEM. In Proceedings of ICERI2020 Conference (Vol. 9, p. 10th).
- Sağlık Bakanlığı, (2020), <https://covid19.saglik.gov.tr/>, Alınma tarihi: 01/10/2020
- Sariman, G., & Mutaf, E. (2020). Covid-19 Sürecinde Twitter Mesajlarının Duygu Analizi.
- Sedik, A., Hammad, M., Abd El-Samie, F. E., Gupta, B. B., & Abd El-Latif, A. A. (2021). Efficient deep learning approach for augmented detection of Coronavirus disease. Neural Computing and Applications, 1-18.
- Singhal, A., Singh, P., Lall, B., & Joshi, S. D. (2020). Modeling and prediction of COVID-19 pandemic using Gaussian mixture model. Chaos, Solitons & Fractals, 138, 110023.
- Tekerek, A. (2011). Veri madenciliği süreçleri ve açık kaynak kodlu veri madenciliği araçları. Akademik Bilişim, 11, 2-4.
- Vaidyanathan, S., & Azar, A. T. (2016). Adaptive control and synchronization of Halvorsen circulant chaotic systems. In Advances in chaos theory and intelligent control (pp. 225-247). Springer, Cham.
- Wang, S., Kang, B., Ma, J., Zeng, X., Xiao, M., Guo, J., ... & Xu, B. (2021). A deep learning algorithm using CT images to

screen for Corona Virus Disease (COVID-19). European radiology, 1-9.

Wikipedia. (2020). Alınma Tarihi: 20/12/2020
https://en.wikipedia.org/wiki/File:Turkey_total_COVID-19_cases_by_NUTS-1_regions.png

Yadav, M., Perumal, M., & Srinivas, M. (2020). Analysis on novel coronavirus (COVID-19) using machine learning methods. Chaos, Solitons & Fractals, 139, 110050.

Yıldırım, Ö., & Fındık O. (2020). Derin Öğrenme Yöntemleri Kullanılarak Türkiye’de Covid-19 Yayılımı Tahmini. II. International Conference On Covid-19 Studies.

Zhang, J., Xie, Y., Li, Y., Shen, C., & Xia, Y. (2020). Covid-19 screening on chest x-ray images using deep learning based anomaly detection. arXiv preprint arXiv:2003.12338.