

ENFEKSİYON HASTALIKLARI ALANINDA YAPAY ZEKA UYGULAMALARI

Dr.Ahmet Rıza Şahin

Adana Şehir Eğitim ve Araştırma Hastanesi

SUNUM PLANI:

Sağlıkta Yapay Zeka

Pandemide Yapay Zeka

Pandemide kullanım alanları

Pandemide kullanılan modellemeler

HIV alanında kullanımı

Antiretroviral dirence yönelik kullanımları

Koruyucu önlemlere dair kullanım modellemeleri

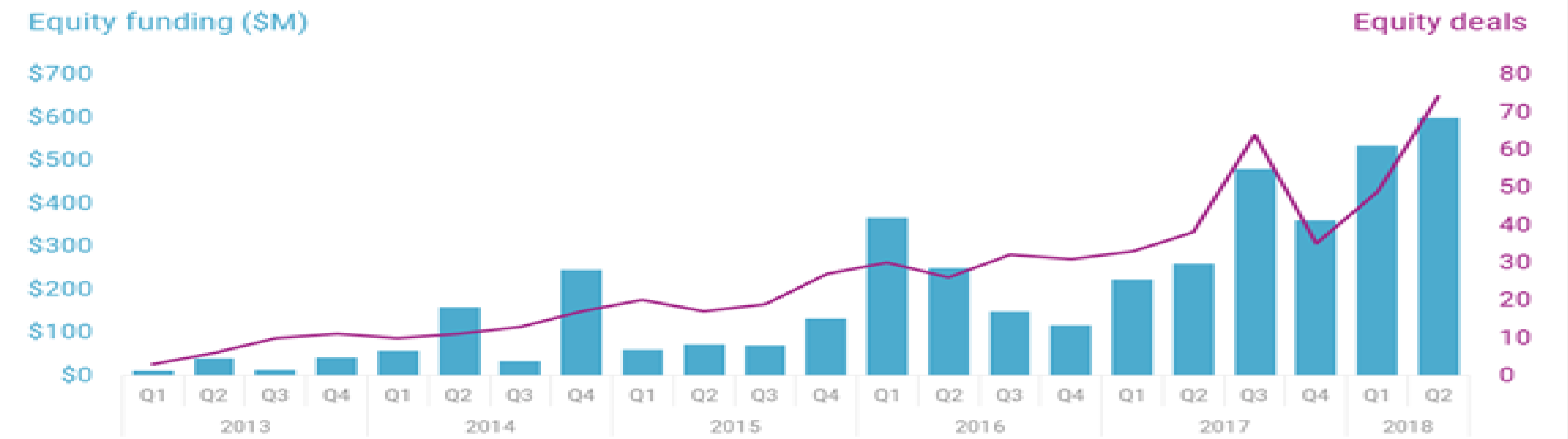
Aşı geliştirmede Yapay Zeka

Gelecek beklentileri



AI in healthcare funding hit a historic high in Q2'18

Disclosed equity funding, Q1'13 - Q2'18



Source: cbinsights.com

CBINSIGHTS





Turk J Gastroenterol 2015; 26: 447-9

Show Me Echo – Hepatitis C: A telemedicine mentoring program for patients with hepatitis C in underserved and rural areas in Missouri as a model in developing countries

Veysel Tahan¹, Ashraf Almashhrawi¹, Rachel Mutrux², Jamal A. Ibdah¹

¹Division of Gastroenterology and Hepatology, University of Missouri, Columbia, Missouri, United States

²Missouri Telehealth Network and Missouri Health IT Assistance Center, Columbia, Missouri, United States

HEPATITIS C: COUNSELING, TESTING, AND CLINICAL CARE

Project ECHO: Linking University Specialists with Rural and Prison-Based Clinicians to Improve Care for People with Chronic Hepatitis C in New Mexico

SANJEEV ARORA, MD
KARLA THORNTON, MD
STEVEN M. JENKINS, MD
BROOKE PARISH, MD
JOSEPH V. SCALETTI, PhD

SYNOPSIS

Project Extension for Community Healthcare Outcomes (Project ECHO) is a telemedicine and distance-learning program designed to improve access to quality health care for New Mexicans with hepatitis C. Project ECHO links health-care providers from rural clinics, the Indian Health Service, and prisons with specialists at the University of New Mexico. At weekly clinics, partners present and discuss patients with hepatitis C with specialists. Partners can receive continuing education credits for participating. Since June 2003, 173 hepatitis C clinics have been conducted with 1,843 case presentations. Partners have received 390 hours of training and 2,997 hours of continuing education credits. And in 2006, the State Legislature approved \$1.5 million in annual funding for the project. Project ECHO has increased access to state-of-the-art hepatitis C virus care for patients living in rural areas or prisons. Because of its success with hepatitis C, this project is being expanded to other chronic medical conditions.

Acad Med. Author manuscript; available in PMC 2013 Dec 7.

PMCID: PMC3855463

Published in final edited form as:

NIHMSID: NIHMS461626

Acad Med. 2007 Feb; 82(2): 10.1097/ACM.0b013e31802d8f68.

PMID: 17264693

doi: 10.1097/ACM.0b013e31802d8f68

Academic Health Center Management of Chronic Diseases through Knowledge Networks: Project ECHO

Dr. Sanjeev Arora, MD, professor and executive vice-chair, Dr. Cynthia M. A. Gappert, MD, PhD, chief consultation-liaison, Dr. Summers Kalishman, PhD, director, Dr. Denise Dion, PhD, research scientist, Dr. Frank Pullara, MD, health services director, Ms. Barbara Bjeletich, CFNP, MS, RN, nurse practitioner, Dr. Gary Simpson, MD, PhD, MPH, medical director, Dr. Dale C. Alverson, MD, professor of pediatrics and regents' professor, and medical director, Dr. Lori B. Moore, PharmD, pharmacist, Dr. Dave Kuhl, PharmD, pharmacist, and Dr. Joseph V. Scaletti, PhD, emeritus professor

DERLEME / Review Article

Klinik Mikrobiyoloji Laboratuvarlarında Yapay Zekanın Temel İşleyiş Modelleri

Basic Processing Models of Artificial Intelligence In Clinical Microbiology Laboratories

Ahmet Rıza Şahin¹, Selma Ateş¹, Mücahid Günay²

¹ Kahramanmaraş Sütçü İmam Üniversitesi Tıp Fakültesi Enfeksiyon Hastalıkları ve Klinik Mikrobiyoloji Anabilim Dalı

² Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Anabilim Dalı

Yazışma Adresi / Correspondence:

Ahmet Rıza Şahin

Kahramanmaraş Sütçü İmam Üniversitesi Tıp Fakültesi Enfeksiyon Hastalıkları ve Klinik Mikrobiyoloji Anabilim Dalı, Kahramanmaraş, Türkiye

T: +90 505 541 37 65

E-mail: : drahmetrizasahin@gmail.com

Orcid

Ahmet Rıza Şahin: <https://orcid.org/0000-0002-4415-076X>

Selma Ateş: <https://orcid.org/0000-0002-2515-8578>

Mücahid Günay: <https://orcid.org/0000-0003-1190-4016>

Geliş Tarihi / Received : 07-08-2019

Kabul Tarihi / Accepted : 22-08-2019

Yayın Tarihi / Online Published: 29-08-2019

Şahin A.R., Ateş S., Günay M., Klinik Mikrobiyoloji Laboratuvarlarında Yapay Zekanın Temel İşleyiş Modelleri
J Biotechnol and Strategic Health Res. 2019;3(2):66-71 DOI: bshr.602790

'Tsunami'

İtalya'daki yoğun bakım ünitelerinde 5200 kadar yatak bulunuyor, ancak kış aylarında bu yatakları zaten solunum güçlüğü yaşayan hastalar dolduruyor.

Ükenin kuzeyindeki Lombardiya ve Veneto bölgelerindeki özel ve kamu hastanelerinde sadece 1800 yatak var.



GETTY IMAGES

İtalya'nın kuzeyindeki hastaneler, yatak sayısını artırmak için geçici koşullar kurdu.

Lombardiya'daki bir hastanede çalışan Dr. Stefano Magnone, BBC'ye yaptığı açıklamada kapasitelerinin dođduđunu söyledi.



Magnone "Durum her geçen gün kötüleşiyor. Koronavirüs pozitif hastaları tedavi etmek için yoğun bakımda ve normal servislerdeki yatak kapasitemizin sınırlarına yaklaşıyoruz" dedi.

"Bölgemizde kaynaklarımız tamamen tükendi , hem insan hem de teknoloji anlamında. Yeni solunum cihazları bekliyoruz".

Bu hafta başı, Bergamo'daki yoğun bakım doktoru Daniele Macchini'nin anlatımları Twitter'da büyük ilgi görmüştü.

Ekibinin "bir tsunamiyle" karşı karşıya olduğunu ve solunum cihazı gibi tıbbi ekipmanların "altın gibi" değerli olduğunu söyledi.

"Vakalar çoğalıyor. Günde 15-20 hasta geliyor, hepsi aynı nedenle. Testlerin sonuçları peş peşe geliyor; pozitif, pozitif, pozitif. Acil servis birden çöktü."

"Enfekte olan iş arkadaşlarımız akrabalarına da bulaştırdı ve bazı yakınları ölüm kalım savaşı veriyor."

Pandemide yapay zeka kullanımı neden ön plana çıktı?

SARS-CoV2 :

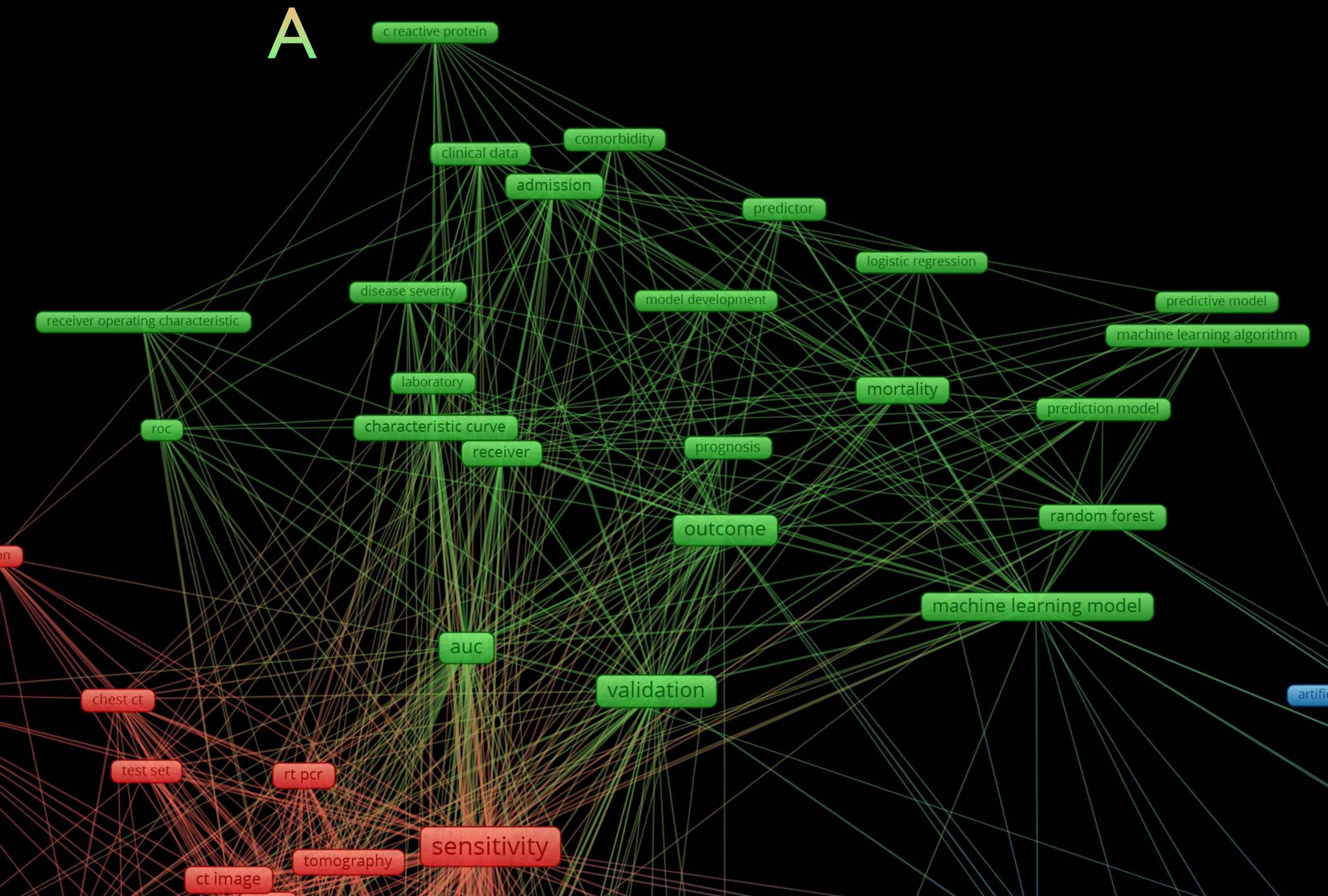
- Hastaneler bulaşın ana kaynaklarındandır.
- Hastaların erken izolasyonu önemlidir.
- Hastanelerde yatan hasta hizmetlerinde planlama ve verimlilik
- Yoğun bakım ünitelerinde doğru planlama
- Anti-sitokin ajanları kullanırken uygun maliyetli tedavi yönetimi
- COVID taşıyıcı asemptomatik personelin erken tespiti

PANDEMİDE KULLANIM ALANLARI:

- Tanı ve prognoz belirleyici yöntemler,
- Tedavi yöntemi geliştirme sürecinde uygulanan algoritmalar
- Sürveyans sistemleri,
- Toplumsal bulaşı kontrol etmeye yardımcı algoritmalar
- Bu sürecin farklı boyutlarını ele alan araştırmalar

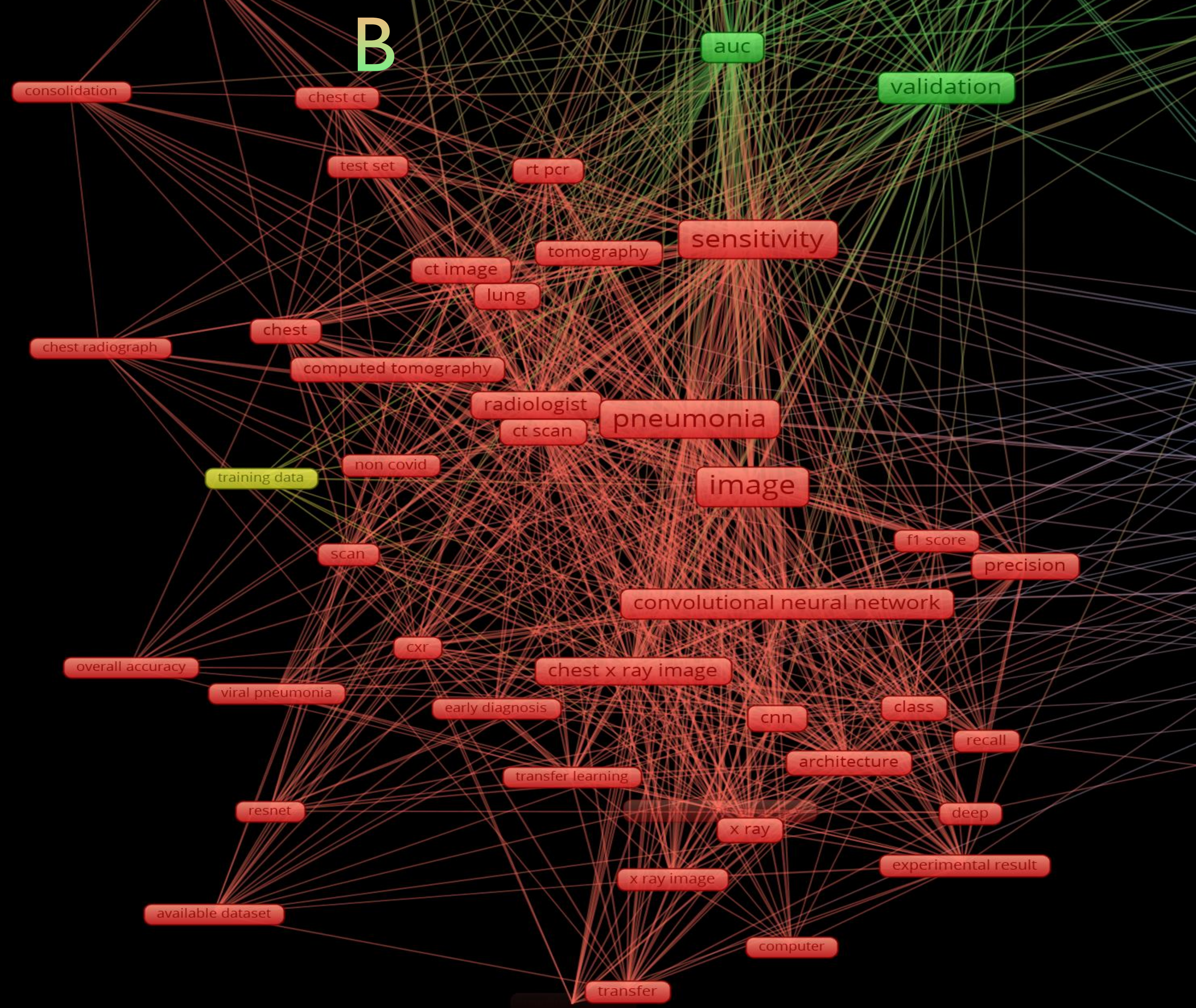
Web Of Science veri tabanında 2020-2022 yılları arasında [“covid” AND (“artificial intelligence” OR “machine learning” OR “deep learning”)] anahtar kelimeleri kullanılarak tespit edilen 694 makalenin kelime içerik ilişkisinin görselleştirilmesi hasta başvurusu, prognozu, yönetimi ve mortalitesinin laboratuvar değerleriyle birlikte değerlendirilmesini sağlayacak yapay zeka çalışmaları (A); Görsellerde radyolojik görüntüleme yöntemleri ile yapılan yapay zeka analizleri (B); ve sosyal medya, trend analizi, big data veri inceleme yöntemleri ile kurgulanan yapay zeka çalışmaları (C) üç ayrı ortak sözcük kümesinde incelenebilmektedir.

A

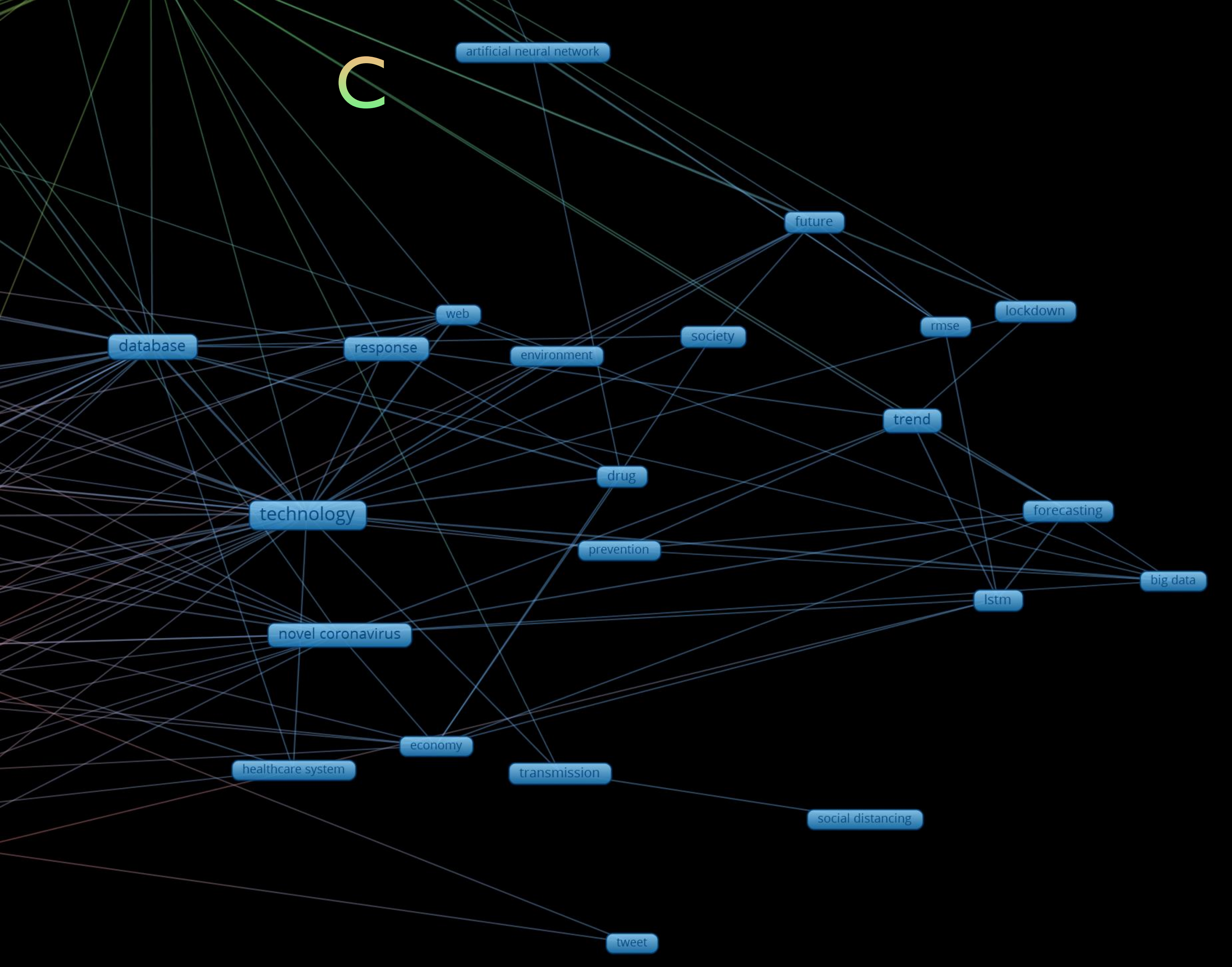


- Validation
- AUC
- Sensitivity
- RT- PCR
- Clinical data
- Random forest
- Prediction
- Machine learning
- Disease severity
- prognosis

B



- Pneumonia
- Convolutional neural network
- Image
- Chest x ray
- CT scan
- Test set
- AUC
- Sensitivity
- RT-PCR
- Clinical data



Validation

AUC

Sensitivity

RT-PCR

Clinical data

Random forest

Prediction

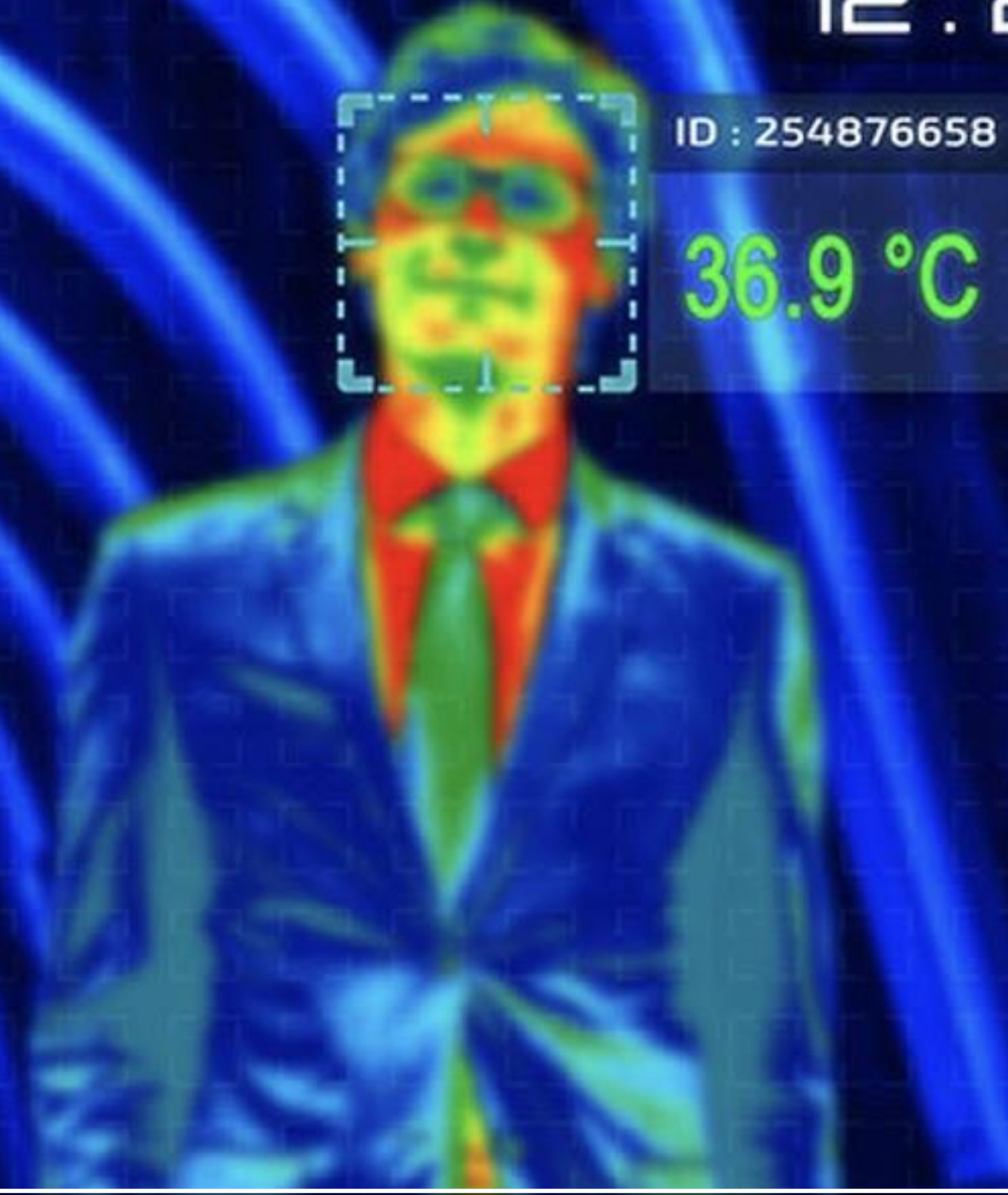
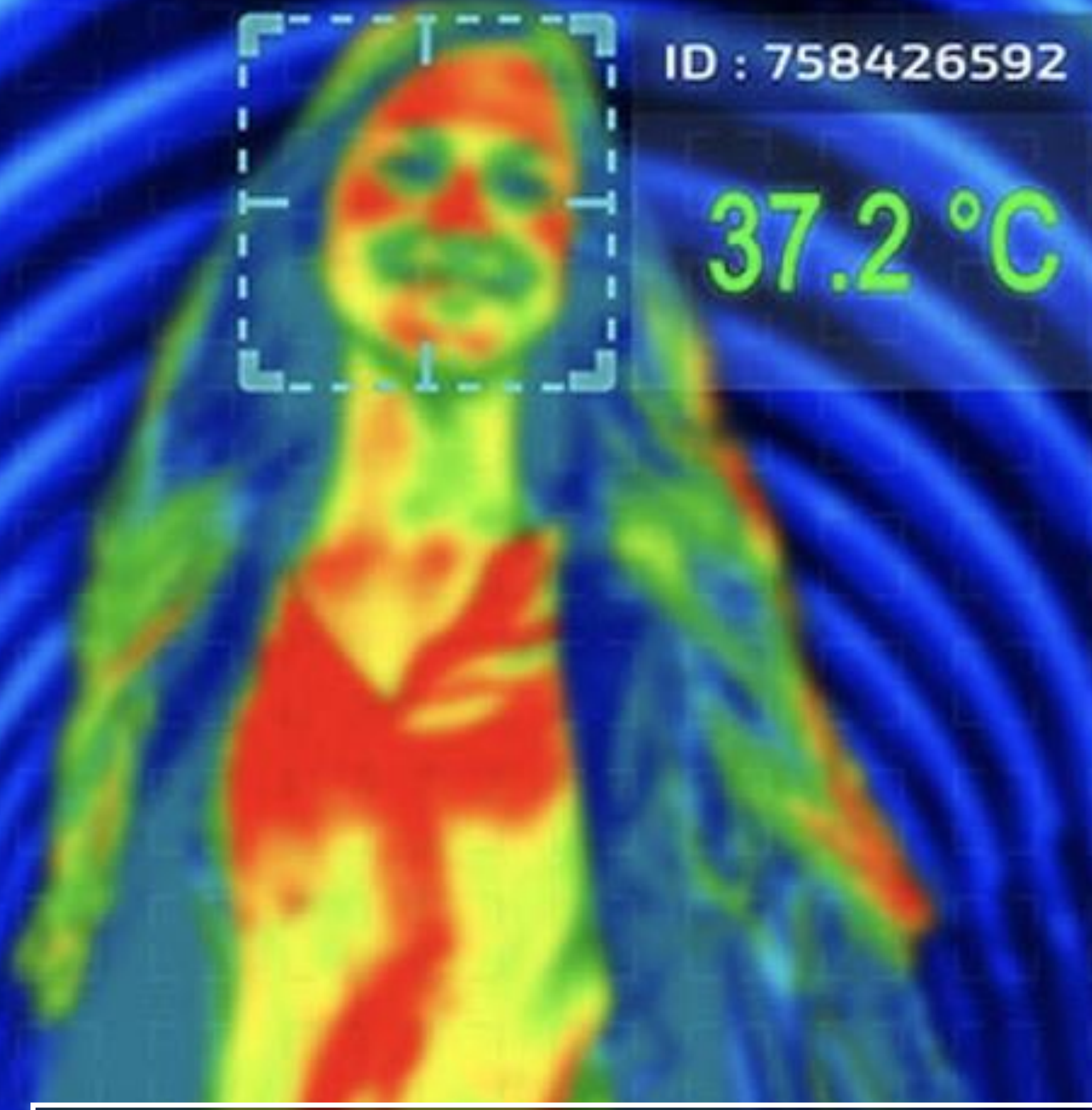
Machine learning

Disease severity

prognosis

CAM 5

12 : 28 : 32



**ISI SENSÖRLÜ KAMERALAR İLE POTANSİYEL COVID 19'LU
KİŞİLERİN TESPİTİ**

HASTALIK SÜRVEYANS SİSTEMLERİ

Erken Tespit ve Alarm Sistemleri

- Bir lokal çevre erken dönem bulguları
- 31 Aralık 2019 pandemi alarmı v
- Benzer şekilde pandemi sinyalleri
- BlueDot şirketin
- Sistem kabaca bitkilerle ilişkili hastalıkların kaçınması için uyarılar

It was an AI that first saw it coming, or so the story goes. On December 30, an artificial-intelligence company called BlueDot, which uses machine learning to monitor outbreaks of infectious diseases around the world, alerted clients—including various governments, hospitals, and businesses—to an unusual bump in pneumonia cases in Wuhan, China. It would be another nine days before the World Health Organization officially flagged what we've all come to know as Covid-19.

BlueDot wasn't alone. An automated service called HealthMap at Boston Children's Hospital also caught those first signs. As did a model run by Metabiota, based in San Francisco. That AI could spot an outbreak on the other side of the world is pretty amazing, and early warnings save lives.

n tehditlerin, tehde özgü

i yapay zeka modeli

yapay zeka uygulaması da

eki veriler, hayvanlar ve riskli bölgelerden

HASTALIK SÜRVEYANS SİSTEMLERİ

Yayılim Tahmin ve Takip Algoritmaları

- COVID19 öncesinde Akhtar ve ark. 2019 yılında zika virüsün zaman içerisinde yayılımını tahmin etme üzerine bir çalışma yapmıştır.
- Bu model, Amerika Kitasındaki tüm etkilenen ülkeler ve bölgeler için;
 - ✓ epidemiyolojik veriler,
 - ✓ yolcu hava yolculuğu hacim verileri ve vektör habitat uygunluğu,
 - ✓ sosyoekonomik ve nüfus verileri kullanılarak
 - ✓ haftalık zamansal çözünürlükte ve ülke uzamsal çözünürlüğünde, Amerika'daki Zika salgınına uygulanmıştır.
- Modelin, 12 haftaya kadar tahmin pencereleri için bile, genel ortalama doğruluk oranı %85'in üzerinde kalmış, Amerika'daki Zika'nın coğrafi genişlemesini doğru bir şekilde tahmin edebildiği gösterilmiştir.

HASTALIK SÜRVEYANS SİSTEMLERİ

Risk Tahmin Algoritmaları

- DeCaprio ve ark. bulaş indeksi hesaplayan Makine öğrenmesi tabanlı bir algoritma geliştirdiğini bildirdiler
- Yaş, seyahat hikayesi, hijyen alışkanlıkları, güncel ve geçmiş sağlık durumu ve aile hikayesi
- Bu faktörlerin direkt olarak bir matematik modellemesi mümkün değildir ancak yapay zeka tabanlı modellemelerde bu faktörler kullanılarak (“**vulnerability index**”) belirlenebilir.
- Enfekte olma riski
- 2 milyon hastaya ait geçmiş hastalık verileri kullanılmış ve hastanın gelecek 3 ay içerisinde hastane yatış ihtimali hesaplanması.



PANDEMi SÜRECİNDE ACABA MAHALLEMDE COVID-19 RISKİ YÜKSEK Mİ ?





- **Sağlık Bakanlığı** iOS ve Android işletim sistemi olan telefon ve tabletlerde kullanılmak üzere **Nisan 2020 tarihinde HES'i** yayınlamıştır
- Sadece yoğunluk haritası
- **Günlük vaka yoğunluğu** ile ilgili bir fikir veriyor ancak **objektif bir risk hesabı yok**

Editöre Mektup

Ankara Med J, 2021;(1):217-219 // doi 10.5505/amj.2021.48379

HAYAT EVE SIĞAR UYGULAMASI YOĞUNLUK HARİTASINDAN İSTANBUL'UN İLÇELERİ VE BÖLGELERİNİN COVID-19 RİSKİNİN SKORLANMASI

SCORING THE COVID-19 RISK OF DISTRICTS AND REGIONS OF ISTANBUL FROM THE DENSITY MAP OF HAYAT EVE SIGAR APPLICATION

 Bekir Aktura¹,  Kursad Koc²,  Mustafa Andaç Derinpınar³
 Nilüfer Aktura⁴

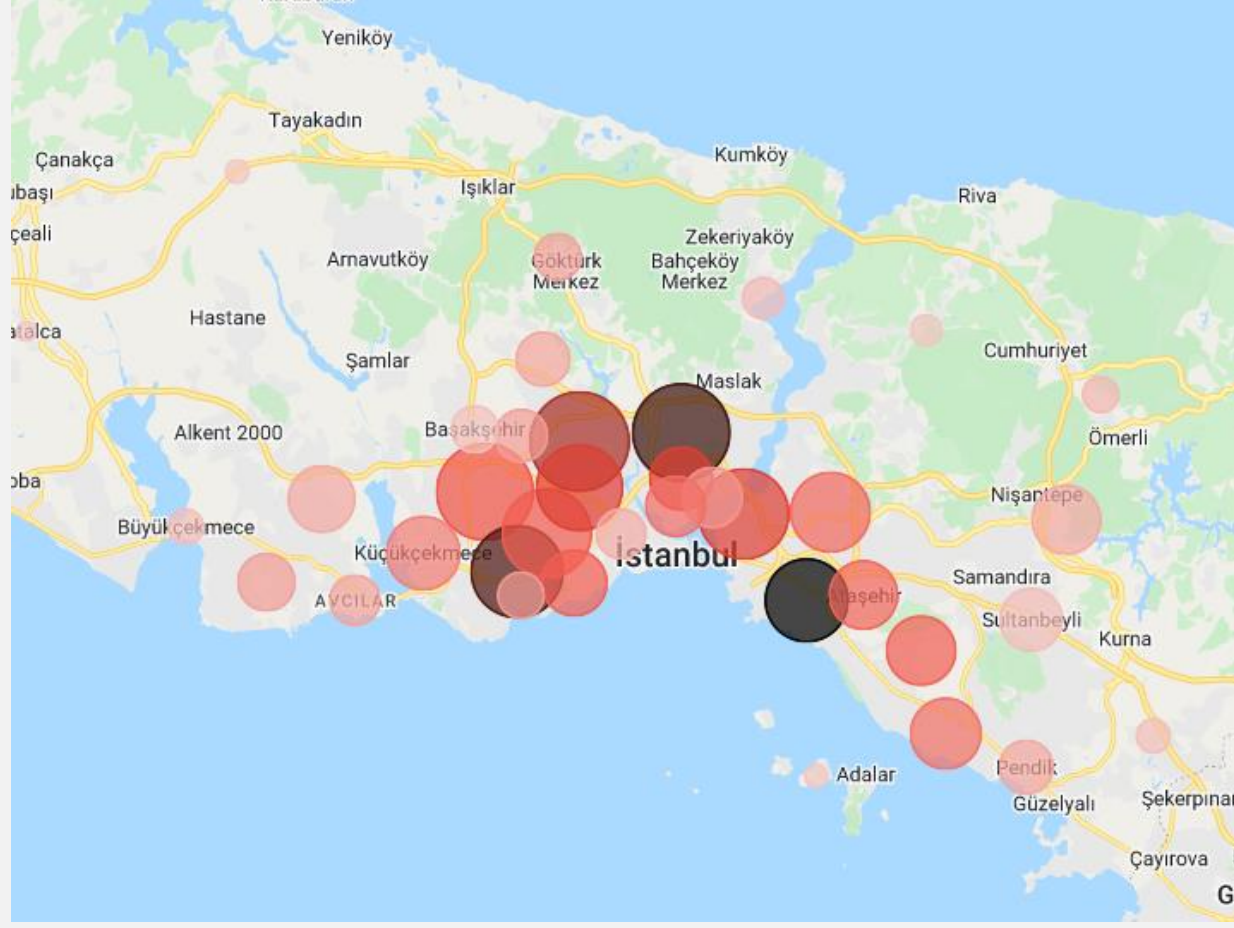
¹İstanbul Üniversitesi, Sağlık Bilimleri Enstitüsü Halk Sağlığı AD, İstanbul

²Bağımsız Araştırmacı, Ankara

³Reidin - Data Analytics, İstanbul

⁴İstanbul Medipol Üniversitesi, Tıp Fakültesi Aile Hekimliği AD, İstanbul

- Python tabanlı bir program
- Günlük olarak HES yoğunluk haritasını(İstanbul sınırları) tarıyor.
- Görüntü işleme yöntemi ile her mahalle için objektif bir risk skoru hesaplıyor
- Google datastudio altyapısı ile halka açık şekilde hesapladığı risk skorunu ve haftalık aylık değişimini raporluyor(İstanbul-ilçeler-mahalleleri olarak tüm katmanlarda)



Kırmızı Risk
16.49
↓ -54.7%

İstanbul'un son 7 günlük
Kırmızı Risk puanı ve bir
önceki yedi güne göre
değişimi

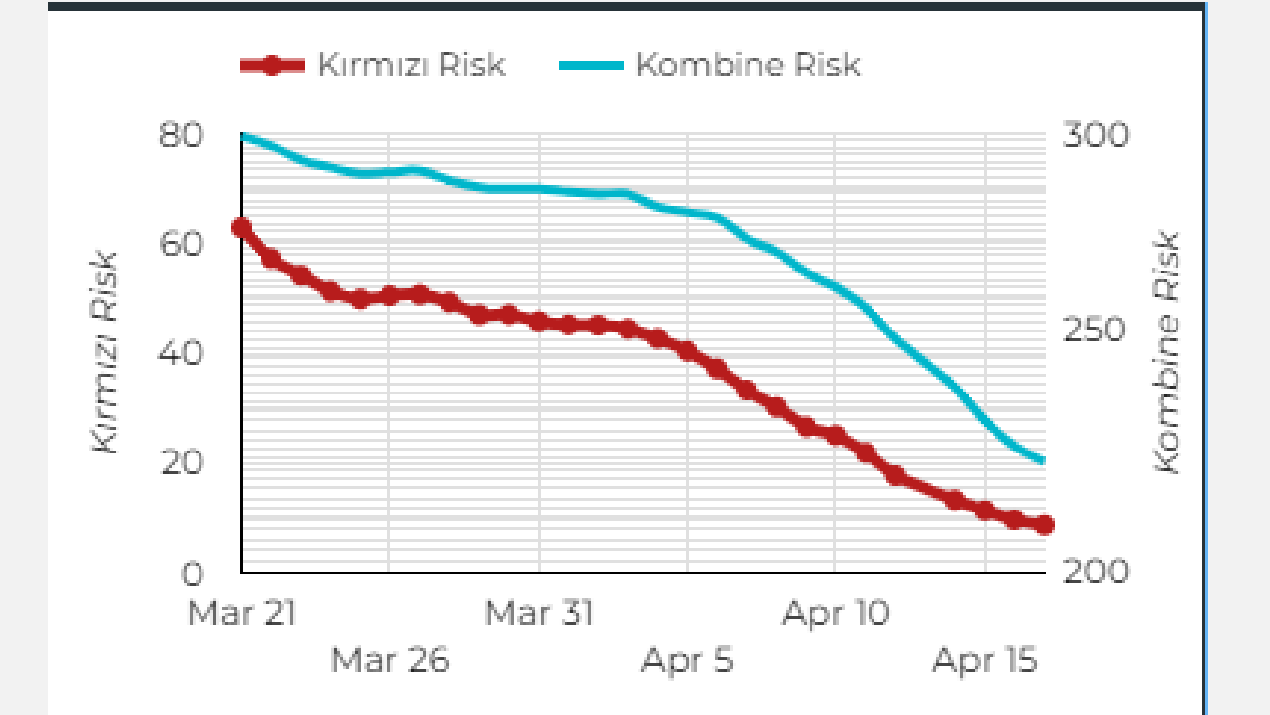
Kombine Risk
126
↓ -3.8%

İstanbul'un son 7 günlük Kombine Risk
puanı
ve
bir önceki yedi güne göre değişimi

istahed.org.tr/diger/covid-risk-takibi/

Tablo: İlçelerin son 14 günlük risk skor ortalamalarını ve bir önceki 14 güne göre değişimlerini gösterir.

İlçe	Kırmızı Risk	% Δ	Kombine Risk	% Δ
1. Kadıköy	299.41	-27.1%	746.42	-0.5%
2. Kâğıthane	235.98	-38.9%	895.61	-3.6%
3. Bahçelievler	221.86	-40.9%	848.11	-1.0%
4. Gaziosmanpaşa	165.36	-44.9%	920.68	-3.2%
5. Şişli	165.19	-24.4%	516.58	-2.7%
6. Üsküdar	149.06	-36.6%	815.11	-3.9%
7. Bayrampaşa	135.96	-45.0%	778.3	-1.7%
8. Cüngören	108.35	-40.1%	796.86	-5.4%
9. Bağcılar	101.19	-53.3%	891.64	-3.2%
10. Ataşehir	97.1	-41.8%	576.72	-3.0%
11. Maltepe	97.02	-47.7%	576.87	-6.0%
12. Zeytinburnu	95.9	-48.6%	541.92	-4.6%
13. Beyoğlu	85.83	-43.3%	490.93	0.2%
14. Ümraniye	82.9	-51.8%	686.26	-7.0%



HASTALIK SÜRVEYANS SİSTEMLERİ

Yayılim Tahmin ve Takip Algoritmaları

- Yayılim tahmin ve takip algoritmalarında
- “gürültü” çoğalması,
- Yeni gelişen bir pandemide (COVID19 pandemisi) geçmişe dönük verilerin ve tarafsız verilerin olmayışı sayılabilir.
- “Google Flu Trends” algoritması 2013 yılında grip salgının tespitinde “gürültülü veri” yüzünden başarısız olmuştur.
- Algoritmalarının kullanımını kritik olarak etkileyenlerden biri false flag gürültüler
- Bu kısıtlılığı aşmak adına “Google”, “Facebook” ve “YouTube” gibi büyük şirketler “big data” üzerinde sahte haberleri kontrol eden içerik moderasyon algoritmaları gibi veri düzenleyici yeni yöntemlere başvurmaktadır

MEDİKAL TANİVE PROGNOZ ALGORİTMALARI

MEDİKAL TANIVVE PROGNOZ ALGORİTMALARI

Tanı ve Prognoz Belirleyici Yöntemler

- Hastalara erken ve doğru müdahale için gereklidir.
- COVID19 tanısı konulduğunda, hastanın hangi düzeyde etkileneceğini tahmin etmek önemli bir sorun haline gelmiştir.
- Bu nedenle hastalara ait klinik, laboratuvar ve/veya radyolojik verilerin kullanılarak hasta prognozunu tahmin etmede çeşitli algoritmalar geliştirilmiştir.
- Bu alan yürütülen çalışmaların en ilgi çekici ve ticari kapasitesi en yüksek olan kısmını temsil eder.

Sex

Male

Body Temperature in Celsius

38.9

Cough?

No

Sputum?

No

Muscle pain?

Yes

Leucocyte Count

7800

Lymphocyte Count

1650

Lymphocyte percentage

21.2

Neutrophile Count

4540

CRP

4.54

Hemoglobin

15 - 55

Hematocrit

0 - 100

Diarrhea?

No

Headache?

No

Loss of sense of smell?

No

SUBMIT

:29:39)

**Covid -19 Test Result:**

High likelihood of COVID-19. Please contact a health care provider within 12 hours!

OK

3,39 - 8,86
1,5 - 5
40,1 - 71,4
1,05 - 3,17
21,6 - 49
0,22 - 0,68
4,3 - 10,2
0,04 - 0,6
2 - 4
0 - 0,2
0 - 1
0,01 - 0,04
0,16 - 0,62
0 - 0,015
0 - 0,03
3,91 - 6
11,1 - 16,6
36,9 - 54,5
80 - 98
25,6 - 30,8
31 - 36
36,3 - 50
36,3 - 50
11,2 - 14
150 - 400
0,19 - 0,41
9,5 - 15,5
17,9 - 43,8
9,2 - 12,2

Table 1 | Characteristics of patient's clinical information

	COVID-19 positive (n = 419)	COVID-19 negative (n = 486)	P value
Sex			
Men	208 (49.6)	280 (57.6)	0.36
^{a,b} Age (years)	43.0 ± 16.4 (32, 54)	38.6 ± 16.3 (27, 48)	0.00086
^b Temperature (°C)	37.6 ± 0.9	37.5 ± 1.0	0.60
Exposure history	319 (76.1)	254 (52.3)	<1 × 10 ⁻⁴
Clinical symptoms			
Fever	314 (75.0)	318 (65.4)	0.0089
Cough	210 (50.1)	128 (26.3)	<1 × 10 ⁻⁴
Cough with sputum	83 (20.0)	177 (36.4)	<1 × 10 ⁻⁴
Laboratory findings			
^{a,b} White blood cells (10 ⁹ l ⁻¹)	5.4 ± 2.2 (4.0, 6.4)	8.3 ± 3.4 (6.1, 10.0)	0.025
^{a,b} Neutrophils (10 ⁹ l ⁻¹)	3.5 ± 1.9 (2.3, 4.2)	5.9 ± 3.1 (3.7, 7.4)	0.11
^{a, b} Percentage neutrophils	63.0 ± 14.7 (53.8, 74.2)	68.9 ± 13.1 (60.7, 78.8)	0.30
^{a,b} Lymphocytes (10 ⁹ l ⁻¹)	1.4 ± 0.8 (0.9, 1.7)	1.6 ± 0.9 (1.0, 2.0)	0.12
^{a,b} Percentage lymphocytes	26.4 ± 12.5 (17.3, 34.5)	21.3 ± 11.1 (12.7, 28.9)	0.53

LETTERS<https://doi.org/10.1038/s41591-020-0931-3>nature
medicine

Check for updates

Artificial intelligence-enabled rapid diagnosis of patients with COVID-19

Xueyan Mei^{1,16}, Hao-Chih Lee^{2,16}, Kai-yue Diao^{3,16}, Mingqian Huang⁴, Bin Lin⁵, Chenyu Liu¹, Zongyu Xie⁶, Yixuan Ma¹, Philip M. Robson^{1,4}, Michael Chung⁷, Adam Bernheim⁴, Venkatesh Mani^{1,4}, Claudia Calcagno^{1,4}, Kunwei Li⁷, Shaolin Li⁷, Hong Shan⁷, Jian Lv⁸, Tongtong Zhao⁹, Junli Xia¹⁰, Qihua Long¹¹, Sharon Steinberger⁴, Adam Jacobi⁴, Timothy Deyer^{12,13}, Marta Luksza¹⁴, Fang Liu¹⁵, Brent P. Little^{15,17}, Zahi A. Fayad^{14,17} and Yang Yang^{14,17}✉

For diagnosis of coronavirus disease 2019 (COVID-19), a SARS-CoV-2 virus-specific reverse transcriptase polymerase chain reaction (RT-PCR) test is routinely used. However, this test can take up to 2 d to complete, serial testing may be required to rule out the possibility of false negative results and there is currently a shortage of RT-PCR test kits, underscoring the urgent need for alternative methods for rapid and accurate diagnosis of patients with COVID-19. Chest computed tomography (CT) is a valuable component in the evaluation of patients with suspected SARS-CoV-2 infection. Nevertheless, CT alone may have limited negative predictive value for ruling out SARS-CoV-2 infection, as some patients may have normal radiological findings at early stages of the disease. In this study, we used artificial intelligence (AI) algorithms to integrate chest CT findings with clinical symptoms, exposure history and laboratory testing to rapidly diagnose patients who are positive for COVID-19. Among a total of 905 patients tested by real-time RT-PCR assay and next-generation sequencing RT-PCR, 419 (46.3%) tested positive for SARS-CoV-2. In a test set of 279 patients, the AI system achieved an area under the curve of 0.92 and had equal sensitivity as compared to a senior thoracic radiologist. The AI system also improved the detection of patients who were positive for COVID-19 via RT-PCR who presented with normal CT scans, correctly identifying 17 of 25 (68%) patients, whereas radiologists classified all of these patients as COVID-19 negative. When CT scans and associated clinical history are available, the proposed AI system can help to rapidly diagnose COVID-19 patients.

The COVID-19 pandemic has rapidly propagated due to widespread person-to-person transmission¹⁻⁴. Laboratory confirmation of SARS-CoV-2 is performed with a virus-specific RT-PCR, but the test can take up to 2 d to complete. Chest CT is a valuable component of evaluation and diagnosis in symptomatic patients with

suspected SARS-CoV-2 infection⁵. Nevertheless, chest CT findings are normal in some patients early in the disease course and therefore chest CT alone has limited negative predictive value to fully exclude infection⁶, highlighting the need to incorporate clinical information in the diagnosis. We propose that AI algorithms may meet this need by integrating chest CT findings with clinical symptoms, exposure history and laboratory testing in the algorithm. Our proposed joint AI algorithm combining CT images and clinical history achieved an area under the curve (AUC) of 0.92 and performed equally well in sensitivity (84.3%) as compared to a senior thoracic radiologist (74.6%) when applied to a test set of 279 cases. While the majority of suspected patients currently have little option but to wait for RT-PCR test results, we propose that an AI algorithm has an important role for the rapid identification of patients with COVID-19, which could be helpful in triaging the health system and combating the current pandemic.

The COVID-19 pandemic has resulted in over 3 million cases worldwide. Early recognition of the disease is crucial not only for individual patient care related to rapid implementation of treatment, but also from a larger public health perspective to ensure adequate patient isolation and disease containment. Chest CT is more sensitive and specific than chest radiography in evaluation of SARS-CoV-2 pneumonia and there have been cases where CT findings were present before onset of clinical symptomatology⁷. In the current climate of stress on healthcare resources due to the COVID-19 outbreak, including a shortage of RT-PCR test kits, there is an unmet need for rapid, accurate and unsupervised diagnostic tests for SARS-CoV-2.

Chest CT is a valuable tool for the early diagnosis and triage of patients suspected of SARS-CoV-2 infection. In an effort to control the spread of infection, physicians, epidemiologists, virologists, phylogeneticists and others are working with public health officials and policymakers to better understand the disease pathogenesis. Early investigations have observed common imaging patterns on

Computers in Biology and Medicine
Early Diagnosis with Artificial Intelligence in COVID-19 Screening
--Manuscript Draft--

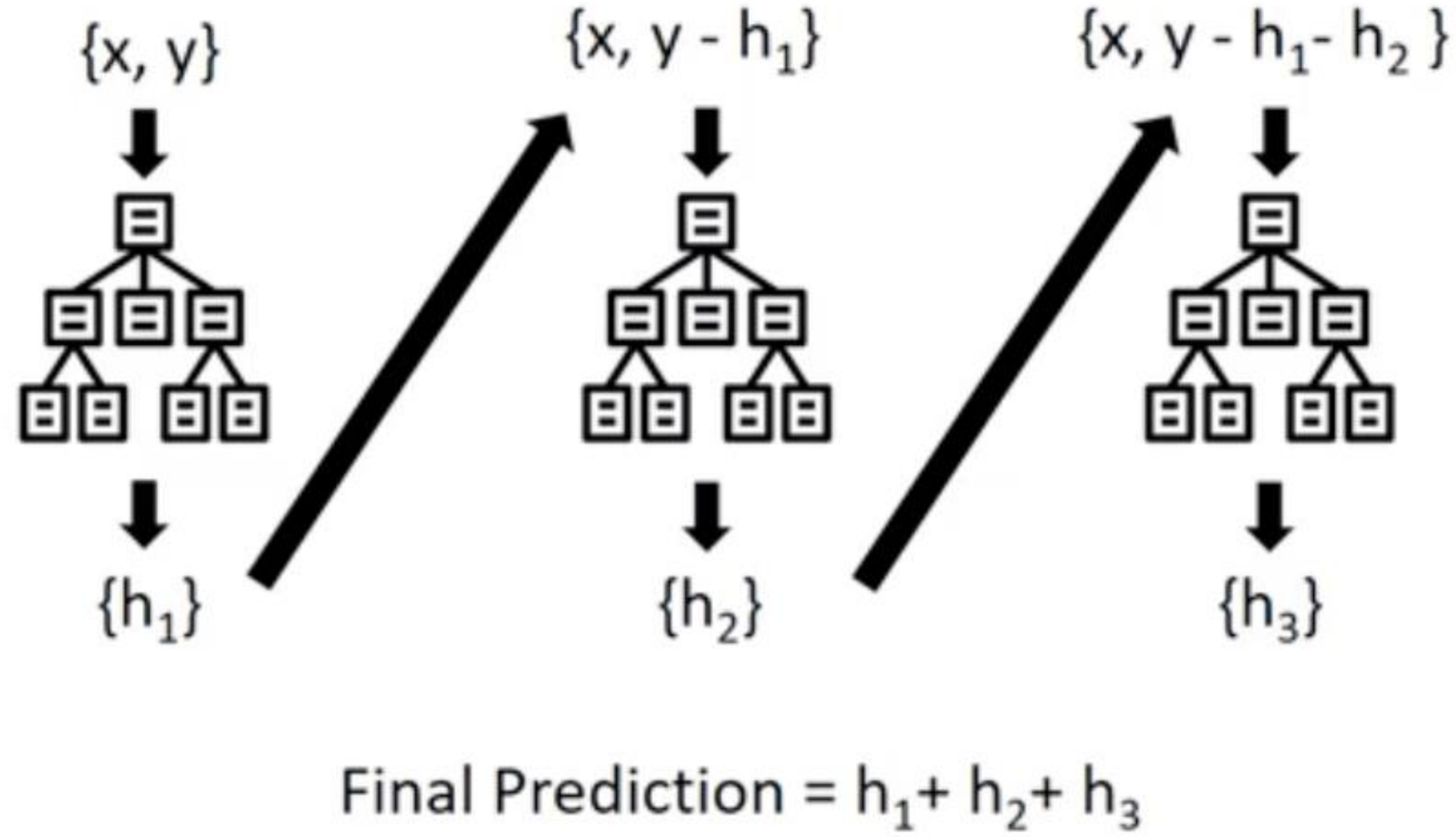
Manuscript Number:	CIBM-D-21-01442
Article Type:	Full Length Article
Keywords:	COVID-19; Deep Learning; pneumonia; Machine Learning
Corresponding Author:	Ahmet Rıza Şahin, associated professor Kahramanmaraş Sütçü İmam Üniversitesi Tıp Fakültesi Kahramanmaraş, Türkiye TURKEY
First Author:	Ahmet Rıza Şahin, Associate professor
Order of Authors:	Ahmet Rıza Şahin, Associate professor Taha Yusuf Kuzan, M.D. Mahmut Egemen Şenel, Associate professor Murat Yüce, M.D. Doğan Çakan, M.D. Ömer Dokumacı, Assistant professor
Abstract:	Objective: Artificial intelligence (AI) methods can analyze complex and large data under the conditions of the coronavirus disease 2019 (COVID-19) pandemic and reveal meaningful relationships within data sets. In this study, we evaluated the potential value of an AI method based on routine blood tests and clinical information to distinguish between COVID-19 and non-COVID-19 patients. Materials and methods: 451 patients, evaluated with the suspicion of COVID-19 from a single-center, included in this retrospective study. Patients' clinic and routine blood tests were noted and then patients divided into 2 groups as COVID-19 (242) and non-COVID-19 (209) based on the reverse transcription-polymerase chain reaction (RT-PCR) test. Gradient Boosting framework with Python implementation of LightGBM were used to distinguish COVID-19 patients from non-COVID-19. Twenty percent of the data was set aside as a test set. Results: The model achieved a 96.4% sensitivity (95% confidence interval (CI) 94.3% - 98.5%) and a 90.2% specificity (95% CI 87.2% - 93.3%) on the test set. Furthermore, the model had a 93.6% (95% CI 92.1% - 95.1%) accuracy and an AUC of 0.976 (95% CI 0.954-0.998). Conclusion: This study shows that AI-based decision tree systems can help detection of COVID-19 patients by following a new methodology depending on the

MEDİKAL TANİ VE PROGNOZ ALGORİTMALARI

Tanı ve Prognoz Belirleyici Yöntemler

- Nükleik asit testi (PCR) sonuçlarına göre teşhis edilen bireylerin ortak değişkenleri veya laboratuvar parametreleri üzerinde çalışılarak, gradient boosting yolağı ile karar ağacı tabanlı çalışan machine learning modeli kullanılarak bir yapay zeka tasarlanmıştır.
- Tasarlanan yapay zeka, 4 şikayet, vücut ısısının sayısal değeri, hemogram analizinden 5 parametre ve CRP değeri okuyarak yüksek hassasiyetle (%92) teşhis koyar.
- Algoritma, delta varyantı gibi tanısal sorun içeren vakaları bulmak için uygulanabilir ve pratiktir.
- Riskli bireyler, mevcut düşük maliyetli testlere entegre edilerek tanımlanabilir.
- Hastaneye yatan her bireyin riskini ortaya ek maliyete ihtiyaç duymadan tanımlar.

Dataset: (x_i, y_i) for $i=1$ to m



Şekil 9: Gradyen Artırma modeli.

- X bağımsız değişkenlerine karşılık gelen Y bağımlı değişkenlerinden oluşan bir veri setinde karar destek ağacı sonucu “ h ” tahminleri elde edilir ve bu tahminler her aşamada bir önceki tahmin sürecinden çıkarılarak hata “error” sonuçları ile tekrar besleme yapılır.
- Sonuçta ortaya çıkan tahmin her bir karar destek ağacından elde edilen tahminlerin toplamıdır.

Veri Yönetimi

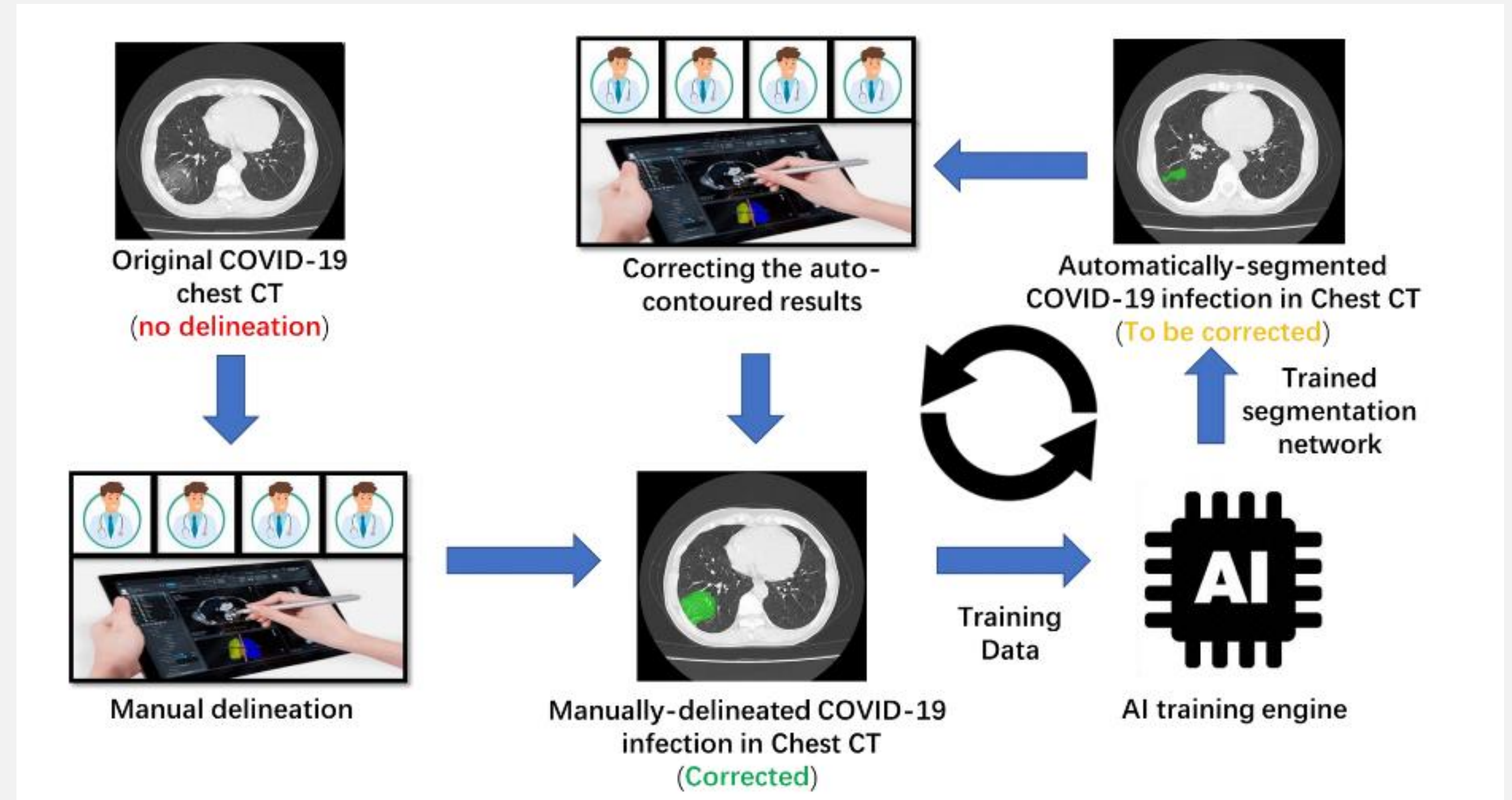
Uygulamalarda veri!!!

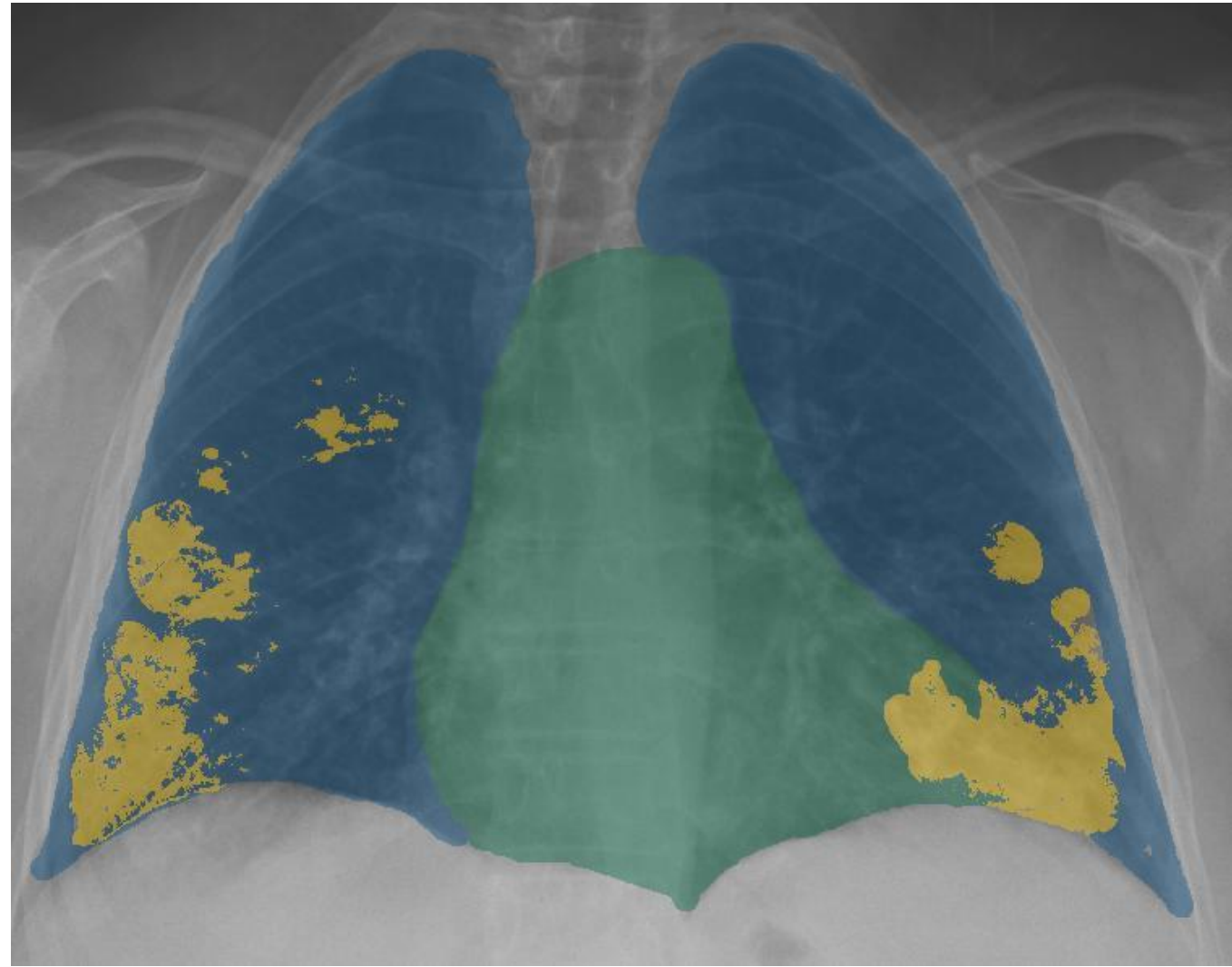
- Yapay zeka algoritmalarında kullanılacak verilere ulaşabiliyorsa açık kaynaklar tercih edilebilir.
- Veri satan şirketler mevcut.
- Hasta tanımlayıcıları (dosya numarası-hasta adı) kullanmadan doğrudan etik kurul başvurusu yaparak ve anonim hale getirilerek çalışır.
- Ayrıca hibrit yöntemlerin kullanıldığı projeler var.
- Bütün veriler ISO-27001 standartlarında korunmalıdır.

BAŐLİCA SORUNLAR:

- Uygulamaların halen geliştirme sürecinde olması
- Örneklem büyüklüklerinin ve verilerin henüz yeterli boyutlara ulaşmaması
- Olası yanlış tanı, tedavi veya bilgi kirliliđi gibi yönlerinin değeriendirmesi
- Medikolegal sorunlarla ilgili dünyada henüz standart bir yaklaşımın bulunmaması

Tomografi yapay zekasının veri sonrası geçen zorlu eğitim süreci

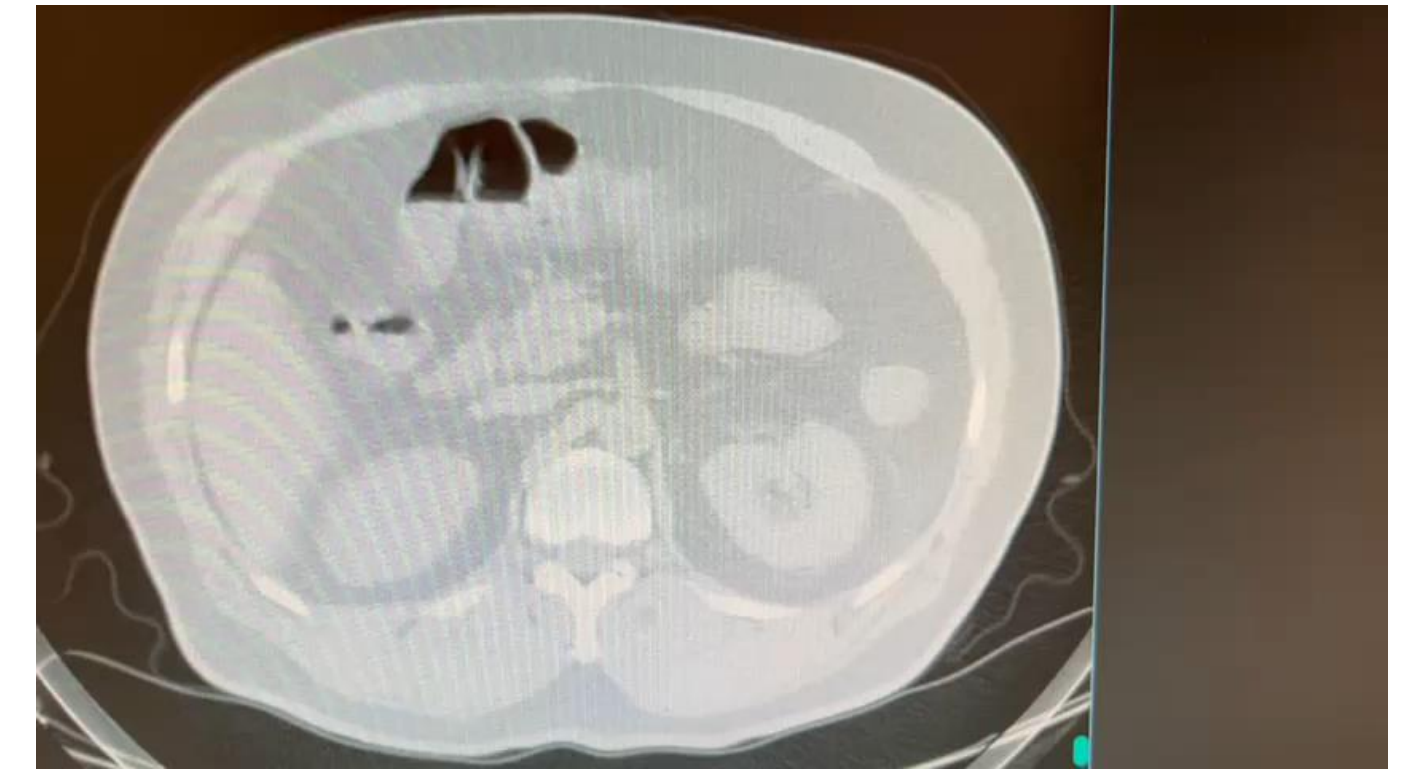


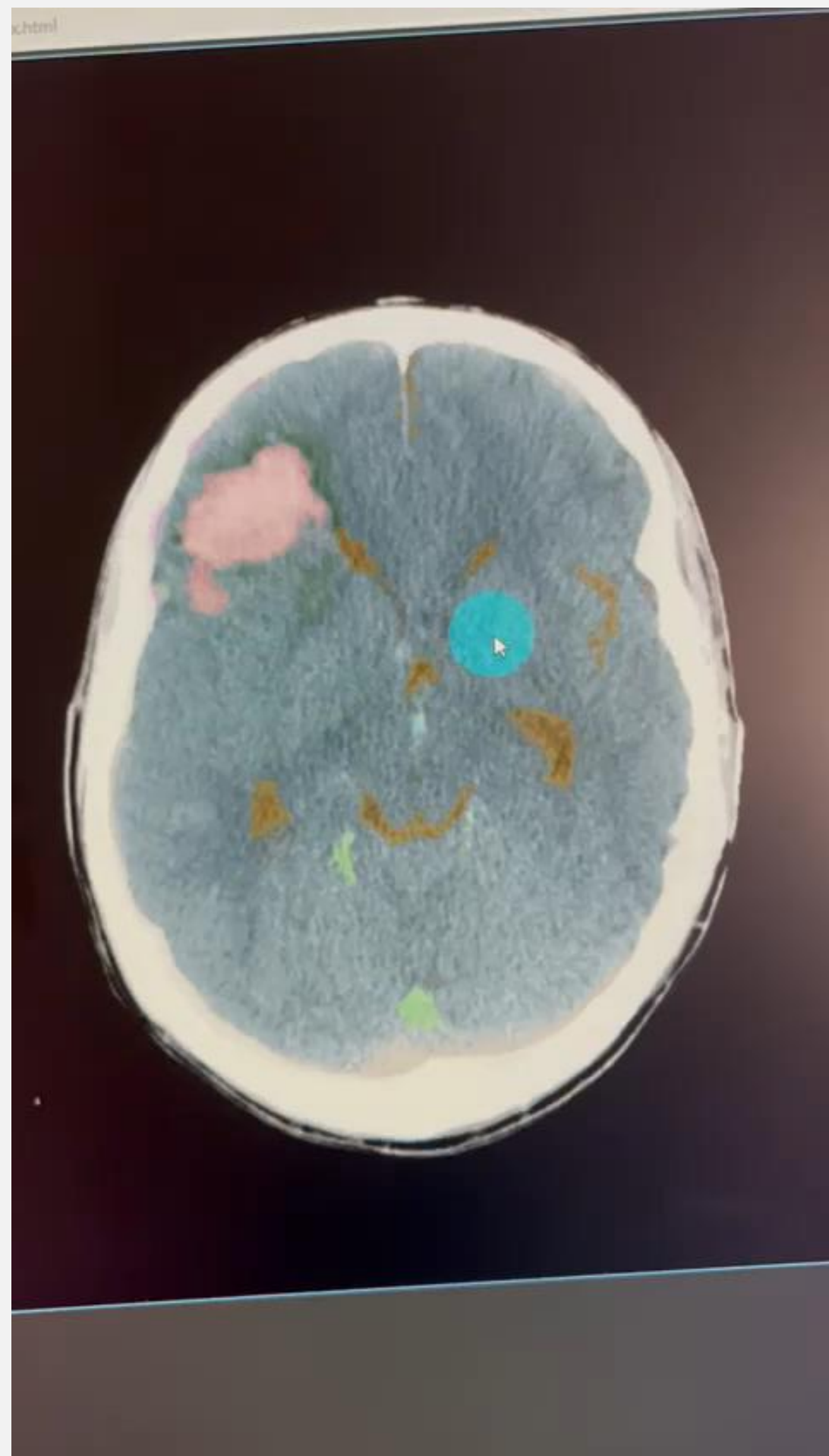
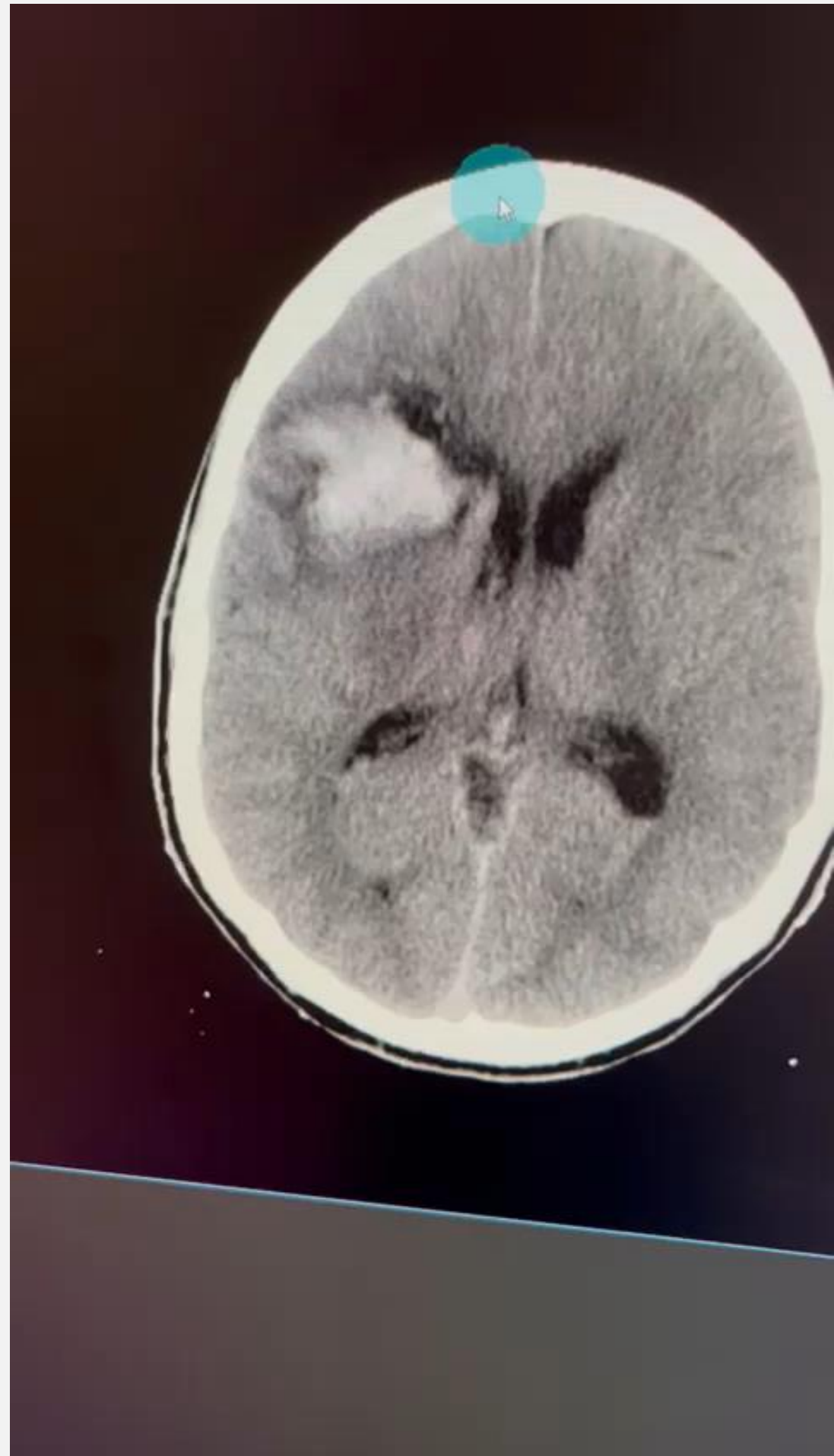


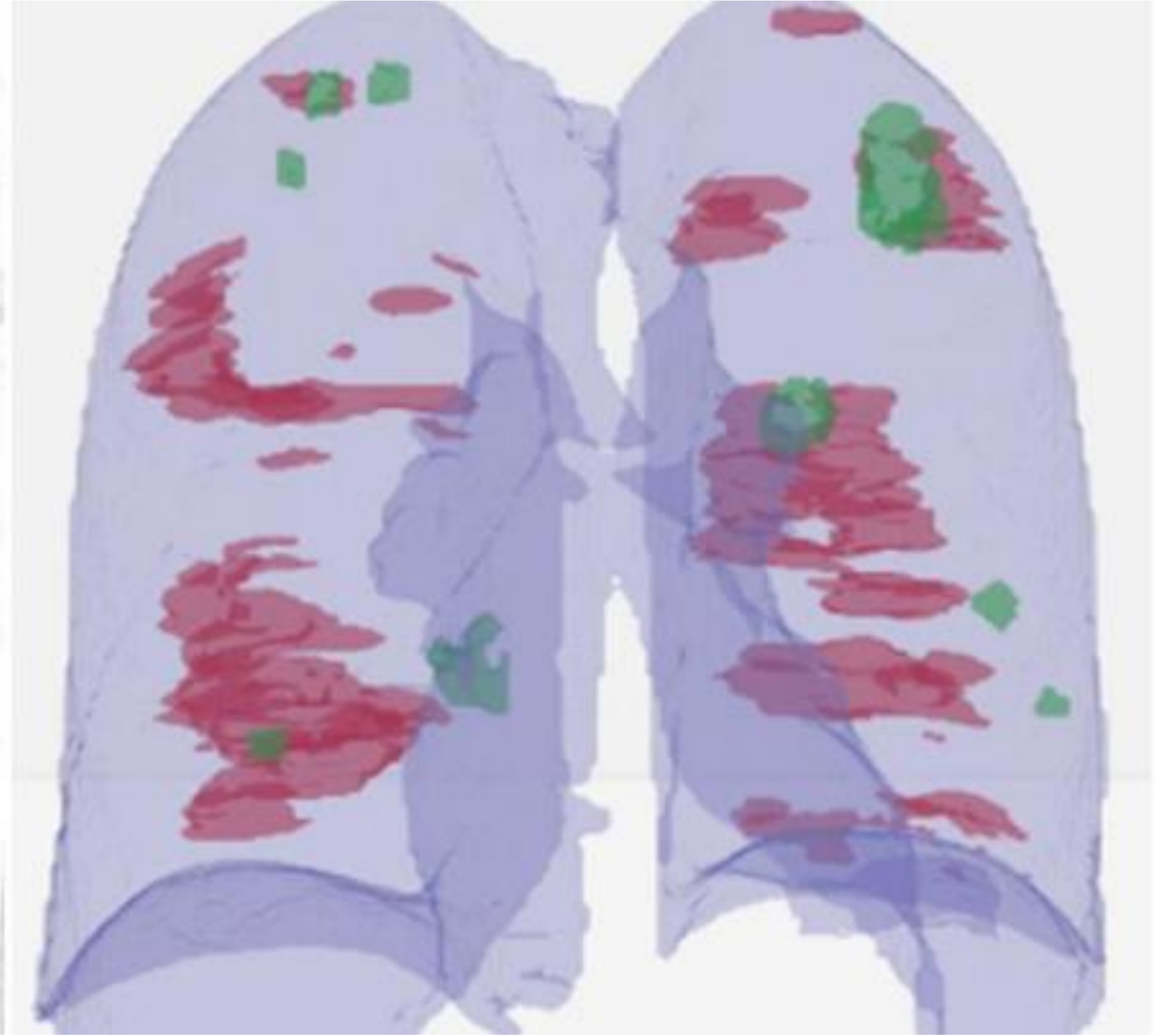
Total segmented volume: 4349.878 ml

1- Lungs volume: 4241.730 ml

2- COVID volume: 108.148 ml







Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks

Ali Narin¹, Ceren Kaya^{2,*}, Ziyne Pamuk²

¹Department of Electrical and Electronics Engineering, Zonguldak Bulent Ecevit University, 67100, Zonguldak, Turkey. alinarin@beun.edu.tr

² Department of Biomedical Engineering, Zonguldak Bulent Ecevit University, 67100, Zonguldak, Turkey. ceren.kaya@beun.edu.tr, ziyne.pamuk@beun.edu.tr

*Corresponding author: ceren.kaya@beun.edu.tr, crnkaya@hotmail.com

Abstract

The 2019 novel coronavirus (COVID-19), with a starting point in China, has spread rapidly among people living in other countries, and is approaching approximately 305,275 cases worldwide according to the statistics of European Centre for Disease Prevention and Control. There are a limited number of COVID-19 test kits available in hospitals due to the increasing cases daily. Therefore, it is necessary to implement an automatic detection system as a quick alternative diagnosis option to prevent COVID-19 spreading among people. In this study, three

Automated Detection of COVID-19 Cases Using Deep Neural Networks with X-ray Images

Tulin Ozturk^a, Muhammed Talo^b, Eylul Azra Yildirim^c, Ulas Baran Baloglu^d, Ozal Yildirim^{e*}, U Rajendra Acharya^{f,g,h}

^aDepartment of Radiology, Medikal Park Hospital, Elazığ, Turkey

^bDepartment of Software Engineering, Firat University, Elazığ, Turkey

^cComputer Engineer, Ministry of Health, Ankara, Turkey

^dDepartment of Computer Engineering, University of Bristol, Bristol, UK

^eDepartment of Computer Engineering, Munzur University, Tunceli, Turkey

^fDepartment of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore

^gDepartment of Bioinformatics and Medical Engineering, Asia University, Taichung, Taiwan

^hInternational Research Organization for Advanced Science and Technology (IROAST), Kumamoto University, Kumamoto, Japan.

*Corresponding author: oyildirim@munzur.edu.tr

Abstract

The novel coronavirus 2019 (COVID-2019), which first appeared in Wuhan city of China in December 2019, spread rapidly around the world and became a pandemic. It has caused a devastating effect on both daily lives, public health, and the global economy. It is critical to detect the positive cases as early as possible so as to prevent the further spread of this epidemic and to quickly treat affected patients. The need for auxiliary diagnostic tools has increased as there are no accurate automated toolkits available. Recent findings obtained using radiology imaging techniques suggest that such images contain salient information about the COVID-19 virus. Application of advanced artificial intelligence (AI) techniques

MEDİKAL TANI VE PROGNOZ ALGORİTMALARI

Prognoz Belirleyici Yöntemler

- Modellerde kullanılan hasta verileri yaş, cinsiyet, sigara öyküsü, vücut ısı, komorbiditeler, gündelik yaşam aktiviteleri ve semptomlardır.
- Model çıktısı olarak ise yoğun bakım ihtiyacı, ekstrakorporal yaşam desteği, mekanik ventilasyon veya vazopressör ihtiyacı veya 30 günlük yatış sonrası ölüm durumuna ilişkin sonuçlar elde edilmiştir.

Development of a prognostic model for mortality in COVID-19 infection using machine learning

Adam L. Booth¹ · Elizabeth Abels¹ · Peter McCaffrey¹

Received: 6 June 2020 / Revised: 28 September 2020 / Accepted: 28 September 2020 / Published online: 16 October 2020
© The Author(s), under exclusive licence to United States & Canadian Academy of Pathology 2020

Abstract

Coronavirus disease 2019 (COVID-19) is a novel disease resulting from infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which has quickly risen since the beginning of 2020 to become a global pandemic. As a result of the rapid growth of COVID-19, hospitals are tasked with managing an increasing volume of these cases with neither a known effective therapy, an existing vaccine, nor well-established guidelines for clinical management. The need for actionable knowledge amidst the COVID-19 pandemic is dire and yet, given the urgency of this illness and the speed with which the healthcare workforce must devise useful policies for its management, there is insufficient time to await the conclusions of detailed, controlled, prospective clinical research. Thus, we present a retrospective study evaluating laboratory data and mortality from patients with positive RT-PCR assay results for SARS-CoV-2. The objective of this study is to identify prognostic serum biomarkers in patients at greatest risk of mortality. To this end, we develop a machine learning model using five serum chemistry laboratory parameters (c-reactive protein, blood urea nitrogen, serum calcium, serum albumin, and lactic acid) from 398 patients (43 expired and 355 non-expired) for the prediction of death up to 48 h prior to patient expiration. The resulting support vector machine model achieved 91% sensitivity and 91% specificity (AUC 0.93) for predicting patient expiration status on held-out testing data. Finally, we examine the impact of each feature and feature combination in light of different model predictions, highlighting important patterns of laboratory values that impact outcomes in SARS-CoV-2 infection.

Initial chest radiographs and artificial intelligence (AI) predict clinical outcomes in COVID-19 patients: analysis of 697 Italian patients

Junaid Mushtaq^{1,2} · Renato Pennella^{1,2} · Salvatore Lavalle^{1,2} · Anna Colarieti^{1,2} · Stephanie Steidler¹ · Carlo M. A. Martinenghi¹ · Diego Palumbo^{1,2} · Antonio Esposito^{1,2} · Patrizia Rovere-Querini^{2,3} · Moreno Tresoldi⁴ · Giovanni Landoni^{2,5} · Fabio Ciceri^{2,6} · Alberto Zangrillo^{2,5} · Francesco De Cobelli^{1,2}

Received: 12 June 2020 / Revised: 30 July 2020 / Accepted: 8 September 2020 / Published online: 18 September 2020
© The Author(s) 2020

Abstract

Objective To evaluate whether the initial chest X-ray (CXR) severity assessed by an AI system may have prognostic utility in patients with COVID-19.

Methods This retrospective single-center study included adult patients presenting to the emergency department (ED) between February 25 and April 9, 2020, with SARS-CoV-2 infection confirmed on real-time reverse transcriptase polymerase chain reaction (RT-PCR). Initial CXRs obtained on ED presentation were evaluated by a deep learning artificial intelligence (AI) system and compared with the Radiographic Assessment of Lung Edema (RALE) score, calculated by two experienced radiologists. Death and critical COVID-19 (admission to intensive care unit (ICU) or deaths occurring before ICU admission) were identified as clinical outcomes. Independent predictors of adverse outcomes were evaluated by multivariate analyses.

Results Six hundred ninety-seven 697 patients were included in the study: 465 males (66.7%), median age of 62 years (IQR 52–75). Multivariate analyses adjusting for demographics and comorbidities showed that an AI system-based score ≥ 30 on the initial CXR was an independent predictor both for mortality (HR 2.60 (95% CI 1.69 – 3.99; $p < 0.001$)) and critical COVID-19 (HR 3.40 (95% CI 2.35–4.94; $p < 0.001$)). Other independent predictors were RALE score, older age, male sex, coronary artery disease, COPD, and neurodegenerative disease.

Conclusion AI- and radiologist-assessed disease severity scores on CXRs obtained on ED presentation were independent and comparable predictors of adverse outcomes in patients with COVID-19.

Trial registration [ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/NCT04318366) NCT04318366 (<https://clinicaltrials.gov/ct2/show/NCT04318366>).

REGISTERED REPORT PROTOCOL

Machine learning model for predicting severity prognosis in patients infected with COVID-19: Study protocol from COVID-AI Brasil

Flávia Paiva Proença Lobo Lopes¹, Felipe Campos Kitamura^{1*}, Gustavo Faibischew Prado², Paulo Eduardo de Aguiar Kuriki¹, Marcio Ricardo Taveira Garcia¹, COVID-AI-Brasil[†]

¹ Departments of Radiology and Innovation, Diagnósticos da América (Dasa), São Paulo, São Paulo, Brasil,
² Department of Innovation, Hospital Alemão Oswaldo Cruz, São Paulo, São Paulo, Brasil

[†] Membership of the COVID-AI Brasil Consortium is provided in the Acknowledgments.
* kitamura.felipe@gmail.com



This is a Registered Report and may have an associated publication; please check the article page on the journal site for any related articles.

OPEN ACCESS

Citation: Paiva Proença Lobo Lopes F, Kitamura FC, Prado GF, Kuriki PEdA, Garcia MRT, COVID-AI-Brasil (2021) Machine learning model for predicting severity prognosis in patients infected with COVID-19: Study protocol from COVID-AI

Abstract

The new coronavirus, which began to be called SARS-CoV-2, is a single-stranded RNA beta coronavirus, initially identified in Wuhan (Hubei province, China) and currently spreading across six continents causing a considerable harm to patients, with no specific tools until now to provide prognostic outcomes. Thus, the aim of this study is to evaluate possible findings on chest CT of patients with signs and symptoms of respiratory syndromes and positive epidemiological factors for COVID-19 infection and to correlate them with the course of the disease. In this sense, it is also expected to develop specific machine learning algorithm for this purpose. through pulmonary segmentation, which can predict possible prognostic fac-

MEDİKAL TANI VE PROGNOZ ALGORİTMALARI

Prognoz Belirleyici Yöntemler

- Ülkemizde geliştirilmiş COVID19 prognoz öngördürücü algorİtmalardan birinde ise COVID19 tanısı almış kişilerde prognostik süreç;
 - hastalığın sorunsuz geçmesi,
 - yoğun bakım ihtiyacı olması,
 - entübasyon ihtiyacı veya ölüm ihtimali olarak 3 sınıfa ayrıştırılmış ve yapay zeka modeliyle hastanın bu sınıflardan hangisine dahil olduğu belirlenmek istenmiştir.
 - Bu çalışmada rutinde bakılan kan parametreleri gibi veriler kullanılmış olup öncelikle hali hazırdaki analiz yöntemlerine ek maliyet oluşturmamak istenmiştir.



Tianqi Chen

XGBoost: A scalable tree boosting system

Yazarlar Tianqi Chen, Carlos Guestrin

Yayın tarihi 2016/8/13

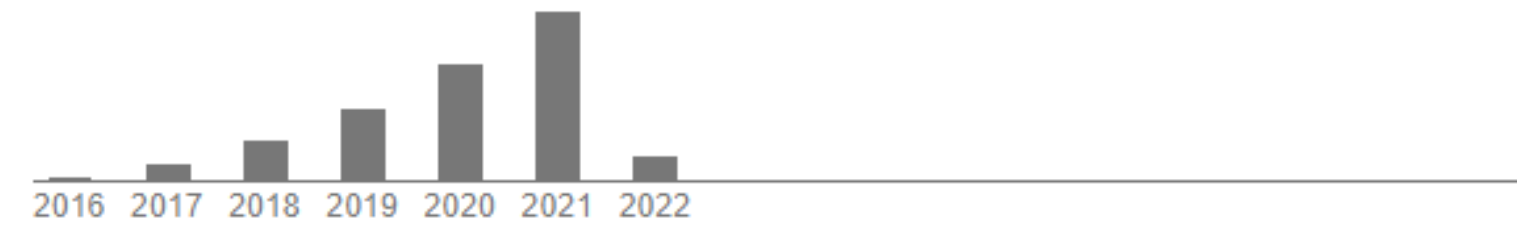
Konferans KDD'16

Sayfalar 785-794

Yayıncı ACM

Açıklama Tree boosting is a highly effective and widely used machine learning method. In this paper, we describe a scalable end-to-end tree boosting system called XGBoost, which is used widely by data scientists to achieve state-of-the-art results on many machine learning challenges. We propose a novel sparsity-aware algorithm for sparse data and weighted quantile sketch for approximate tree learning. More importantly, we provide insights on cache access patterns, data compression and sharding to build a scalable tree boosting system. By combining these insights, XGBoost scales beyond billions of examples using far fewer resources than existing systems.

Toplam alıntı sayısı **Alıntılanma sayısı: 15969**



**COVID19
PROGNOZ
MODELİNE AIT
DOĞRULAMA
(VALIDASYON)
VAKALARI (N=88)
ÜZERİNDE ÖN
SONUÇLAR.**

Birinci gün değerleri ile koşulan sistem	Yapay zekanın serviste yatıp taburcu öngördükleri	Yapay zekanın yoğun bakıma girip taburcu olacak dedikleri	Yapay zekanın <u>entübe</u> olacak ve <u>ölecek</u> dedikleri
Serviste yatıp taburcu	50	3	2
Yoğun bakımda taburcu	4	14	5
<u>Entübasyon</u> ve ölüm	1	1	8

TEDAVI YÖNTEMİ GELİŞTİRME SÜRECİNDE UYGULANAN ALGORİTMALAR

- Tedavi protokollerinin bulunmayışı
- AI yöntemleri ile yeni ilaçları
- Var olan ilaçların COVID19 da kullanımı da araştırılabilmektedir.
- COVID19 pandemisinde ilaçların yeniden yapılandırma ile potansiyel kullanımı yapay zeka yöntemleri ile araştırılmaktadır.
 - Beck ve ark. ML yöntemi kullanarak atazanavirin COVID19 tedavisinde potansiyel taşıdığını
 - Stebbing ve ark. ise İngiltere merkezli bir yapay zeka start up şirketi ile kullandıkları bir algoritma vasıtasıyla romatoid artrit ve myelofibrozis gibi hastalıklarda kullanılan Baricitinin COVID19 tedavisinde etkin olabileceğini öne sürmüştür.

TOPLUMSAL
BULAŐI KONTROL
ETMEYE YARDIMCI
ALGORİTMALAR

TOPLUMSAL BULAŐI KONTROL ETMEYE YARDIMCI ALGORİTMALAR

- Denetlenen bölgelerde alınan tedbirlere uyumun kontrolü için yapay zeka yöntemlerinden faydalanılabilmektedir.
- Çin'de, kızıl ötesi kameralar kullanılarak vücut sıcaklığı ve cerrahi maske kullanımı otomatik olarak denetlenebilmiştir.
- İngiltere'de insanların sosyal mesafe kurallarına uyumunu monitörize etmek için halka açık alanları tarayan yapay zeka
- ABD'de de bir start up şirketi, sosyal mesafe kuralları bozulduğunda tespit için kamera görüntülerini kullanan ve uyarı bildiri yapan bir "sosyal mesafe tespit" yazılımı geliştirmiştir
- Pandemi süreci gibi toplum sağlığını ve güvenliğini tehdit eden olaylar dışında bireylerin otoriteler tarafından sürekli izlenmesi fikri korkutucu gelebilmektedir.
- Bu verilerin farklı amaçlarla kullanılması ihtimali nedeniyle bu tarz uygulamalara sıcak yaklaşılabilir.
- Bu nedenle geliştirilen uygulamalarda veri kullanımının şeffaf ve uygulama amacıyla sınırlı kalması, gerekmeyen hallerde bu verilerin güvenli bir şekilde depolanması veya ortadan kaldırılması elzemdir.

YAPAY ZEKANIN BİLEŞENLERİ

YAPAY ZEKANIN BİLEŞENLERİ

- Çok sayıda alt ünitenin bir iç dolaşım ve dış dolaşım bağlantısı ile kusursuz olarak çalışması gerekir.
 - Elektronik sağlık kayıtları
 - Sağlık bilgi değişimi (SBD) ağlarıdır.
 - Tıp terminolojileri, kapsamaları ve ölçüm yöntemleri farklılık gösterebilir.
 - Tıp terminolojileri en önemli gereksinim ölçütlerin, analizlerin, numunelerin ideal olarak ortak bir kodu paylaşmasıdır.
- **NPU:**
 - **LOINC:**
 - **SNOMED CT:**
 - **UCUM:**
- Örnek vermek gerekirse Escherichia coli'i temsil eden I I2283007, genişletilmiş spektrumlu beta laktamaz enzimi üretiyor ise 40980000, Kapbapenemaz enzim geni taşıyorsa 737528008 direnç özellikleri için SNOMED CT kodu ile tanımlanır.
 - Sağlıkta, UCUM kodları elektronik iletişimde (Dijital hastane 7. Seviye normları tarafından tanımlanan formatlardaki mesajlar veya belgeler gibi) kullanılmak üzere tasarlanmıştır ve genellikle insan yorumuna aşina olan diğer birim dizeleri de bulunur.

YAPAY ZEKANIN BİLEŞENLERİ

Code System Concept

Code System Concept Code	86406008
Code System Concept Name	Human immunodeficiency virus infection (disorder)
Code System Preferred Concept Name	Human immunodeficiency virus infection (disorder)
Concept Status	Published
Concept Status Date	03/01/2019
Code System Name	SNOMED-CT

LOINC
VERSION 2.66

LOINC CODE 69668-2	LONG COMMON NAME HIV 1 and 2 Ab [Identifier] in Serum or Plasma by Rapid immunoassay	LOINC STATUS Active
FULLY-SPECIFIED NAME		
Component	HIV 1 & 2 Ab	
Property	Prid	
Time	Pt	
System	Ser/Plas	
Scale	Nom	
Method	IA.rapid	
Additional Names		
Short Name	HIV 1 & 2 Ab SerPI IA.rapid	
Display Name <small>BETA</small>	HIV 1 and 2 Ab IA.rapid Nom	
Consumer Name <small>ALPHA</small>	HIV 1 and 2 Antibody, Blood	

Most Common Healthcare Units

Valid UCUM Code	Descriptive Name	Common Synonym (non-UCUM)	Dimension (IUPAC)
%	Percent	%	1
/uL	PerMicroLiter	/uL	L-3
[IU]/L	InternationalUnitsPerLiter	IU/L	L-3[arb]
10 ³ /uL	ThousandsPerMicroLiter	K/uL, x10 ³ /mm ³	L-3
10 ⁶ /uL	MillionsPerMicroLiter	M/uL, x10 ⁶ /mm ³	L-3
fL	FemtoLiter	fL	L-3
g/dL	GramsPerDeciLiter	g/dL	L-3M
g/L	GramsPerLiter	g/L	L-3M
g/mL	GramsPerMilliLiter	g/mL	L-3M
kPa	KiloPascal	kPa	L-1MT-2
m[IU]/mL	MilliInternationalUnitsPerMilliLiter	mIU/mL	L-3[arb]
meq/L	MilliEquivalentsPerLiter	mEq/L	L-3N
mg/dL	MilliGramsPerDeciLiter	mg/dL	L-3M
mm[Hg]	MilliMetersOfMercury	mm Hg	L-1MT-2
mmol/kg	MilliMolesPerKiloGram	mmol/kg	M-1N
mmol/L	MilliMolesPerLiter	mmol/L	L-3N
mosm/kg	MilliOsmolesPerKiloGram	mOsm/kg	M-1N
ng/mL	NanoGramsPerMilliLiter	ng/mL	L-3M
nmol/L	NanoMolesPerLiter	nmol/L	L-3N
pg	PicoGrams	pg	M

[Asymptomatic human immunodeficiency virus infection \(disorder\) {91947003, SNOMED-CT}](#)

[Candidiasis of mouth co-occurrent with human immunodeficiency virus infection \(disorder\) {713497004, SNOMED-CT}](#)

[Cognitive impairment co-occurrent and due to human immunodeficiency virus infection \(disorder\) {15928141000119107, SNOMED-CT}](#)

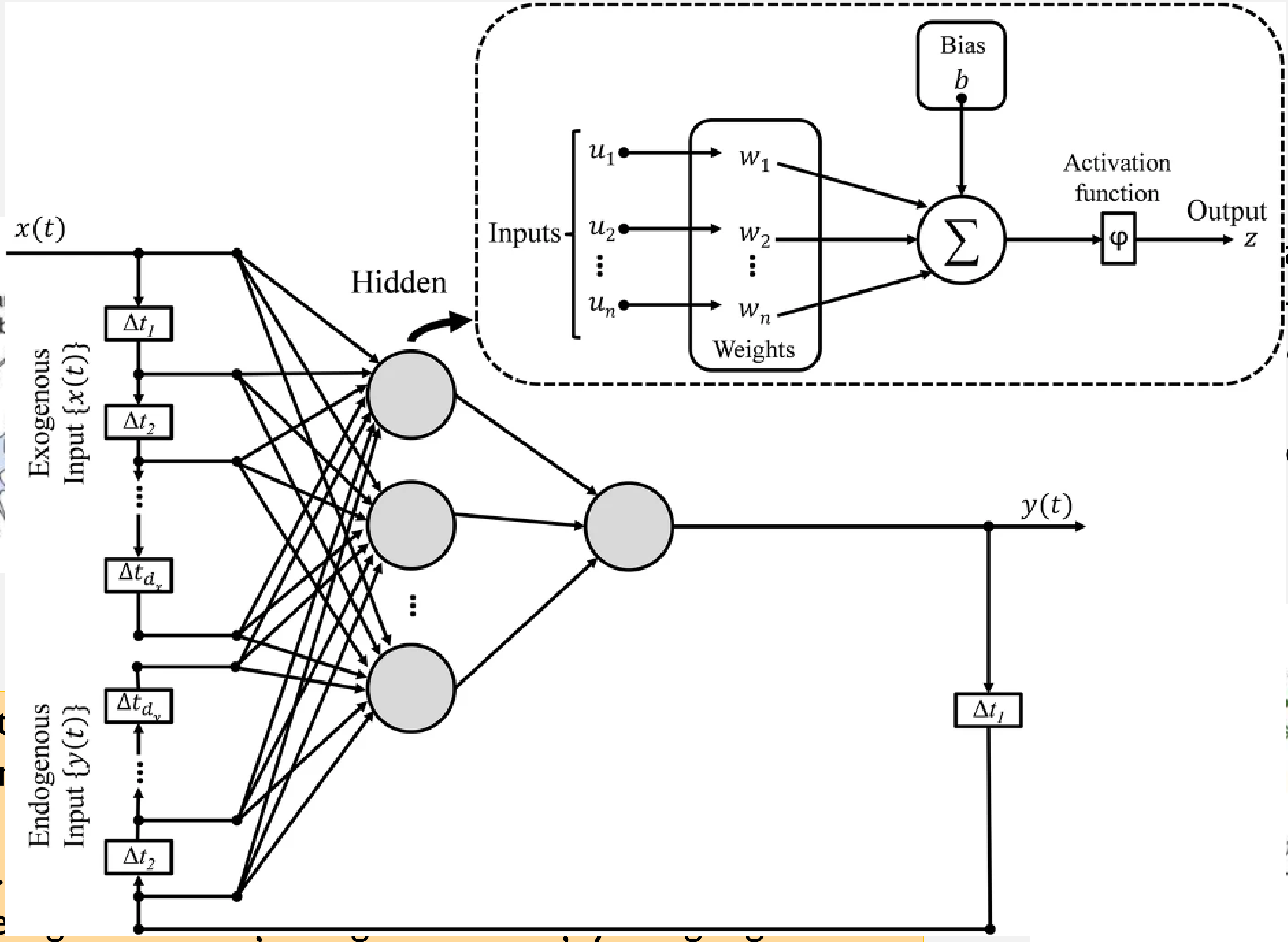
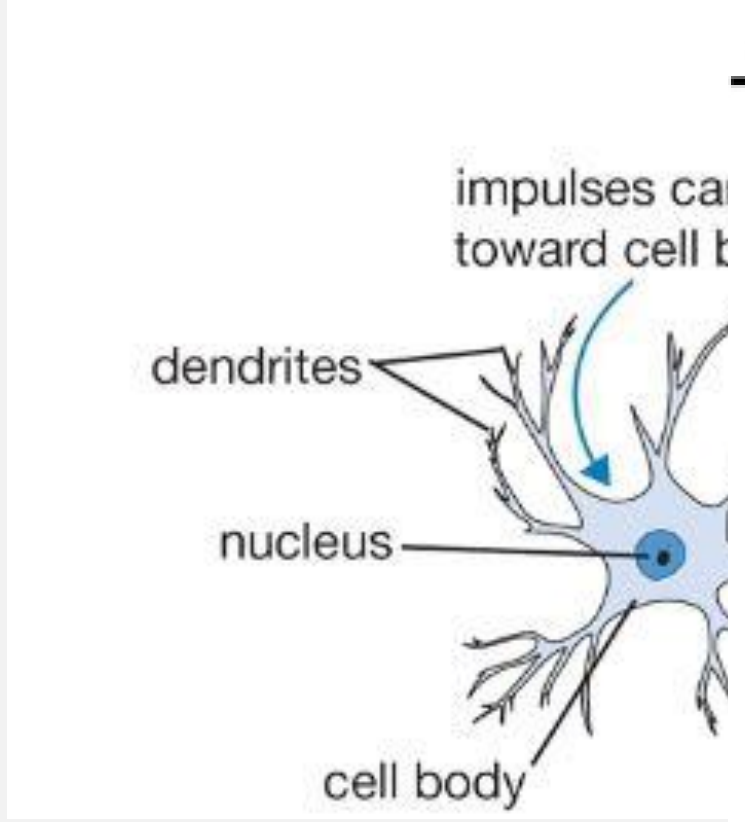
[Congenital human immunodeficiency virus infection \(disorder\) {52079000, SNOMED-CT}](#)

[Disorder of central nervous system co-occurrent with human immunodeficiency virus infection \(disorder\) {713571008, SNOMED-CT}](#)

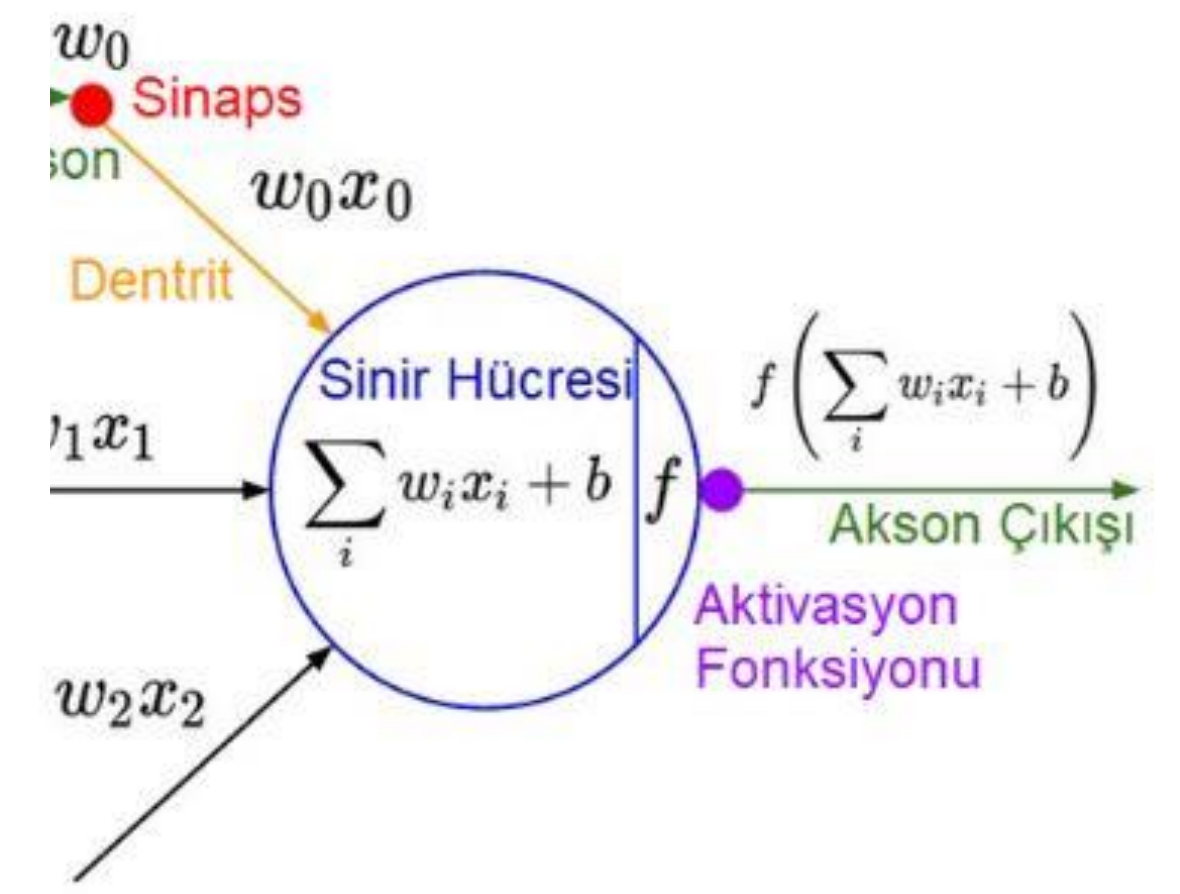
[Disorder of eye proper co-occurrent with human immunodeficiency virus infection \(disorder\) {713299003, SNOMED-CT}](#)

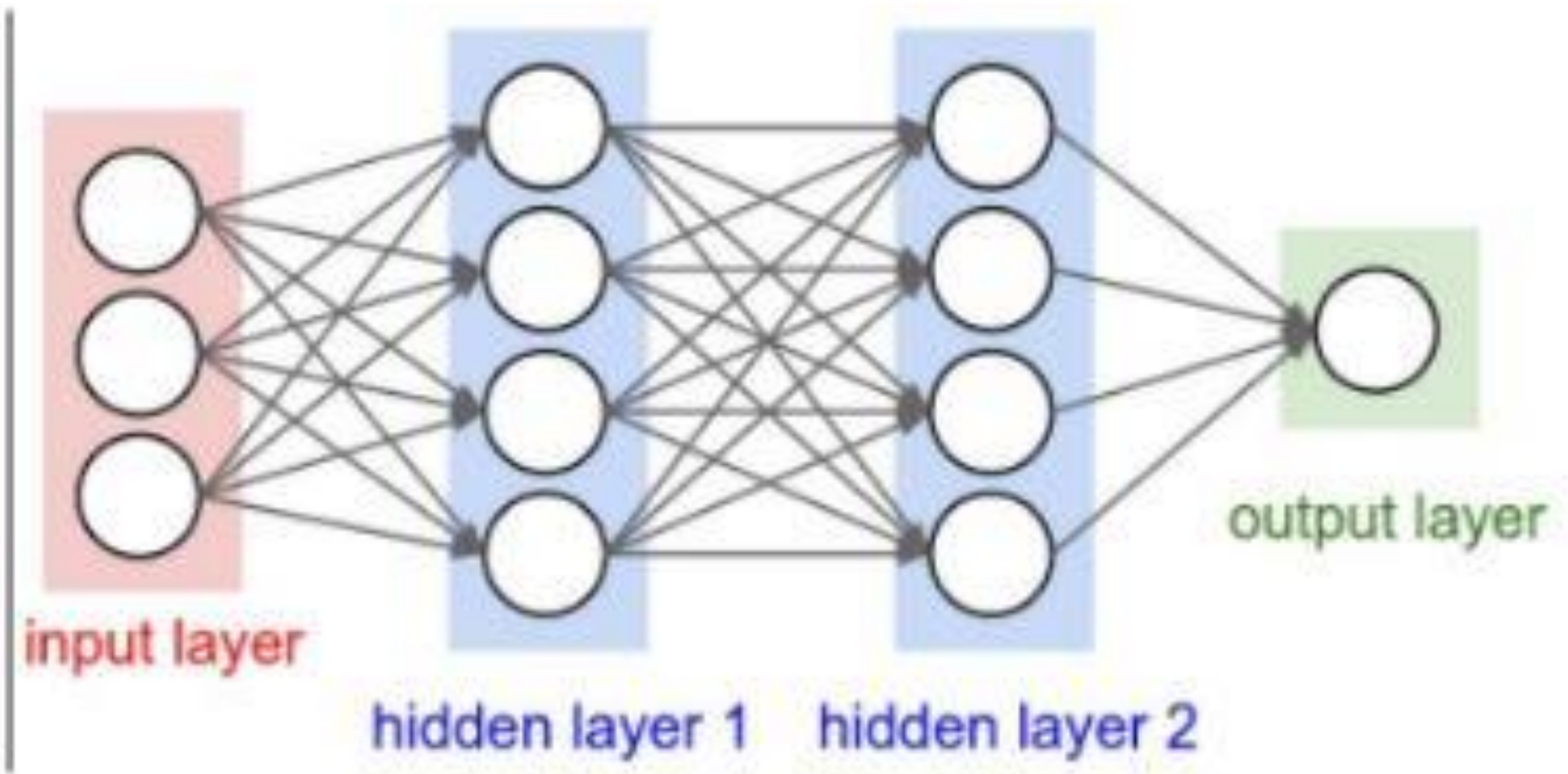
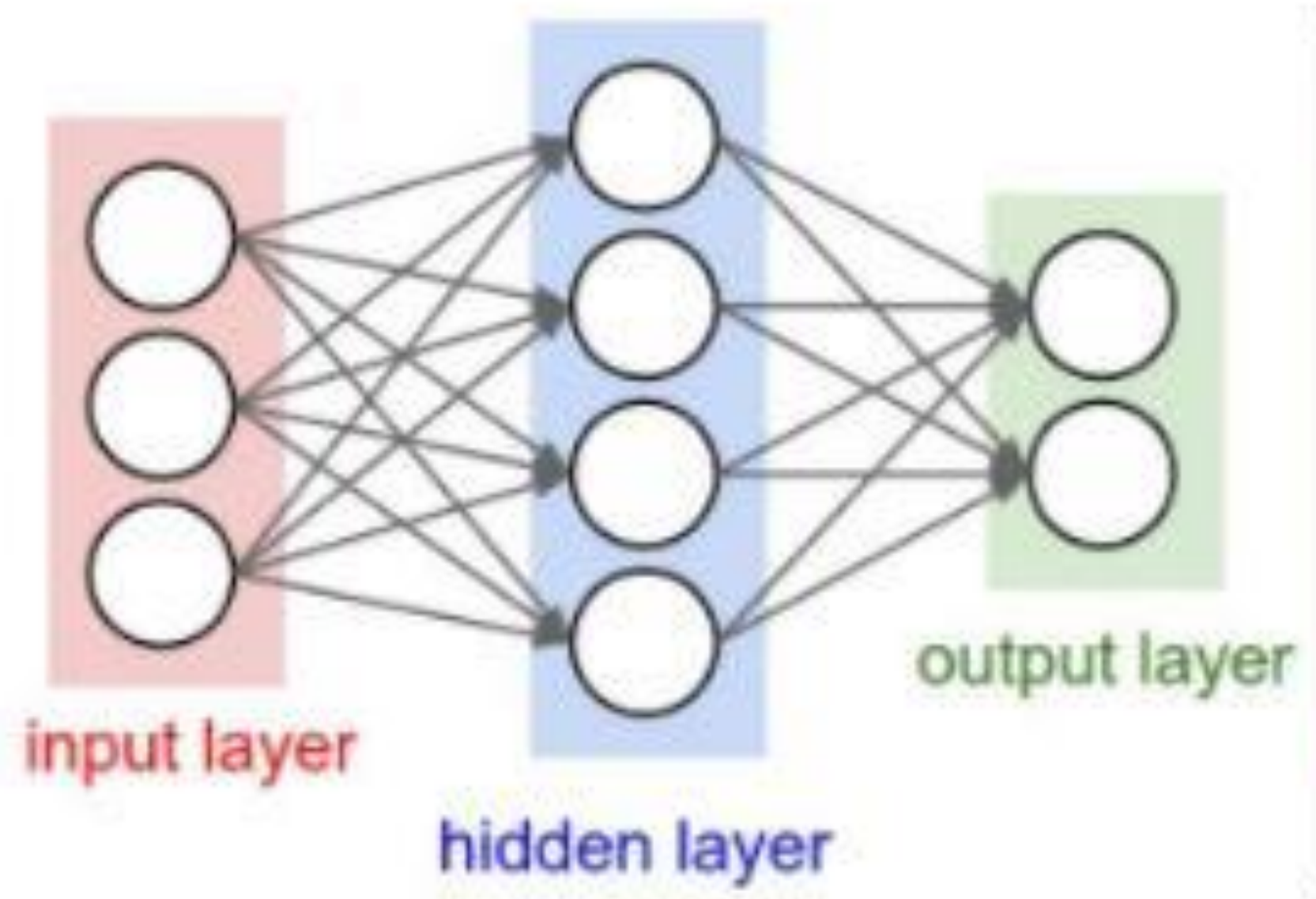
[Disorder of gastrointestinal tract co-occurrent with human immunodeficiency virus infection \(disorder\) {713300006, SNOMED-CT}](#)

Yapay öğrenme t yapısının modellen hiyerarşik bir gerçekleşmektedir. duyduğumuz, hisse bu yana öğrenmemizi, hatırlamamızı ve ezberlememizi sağlar.



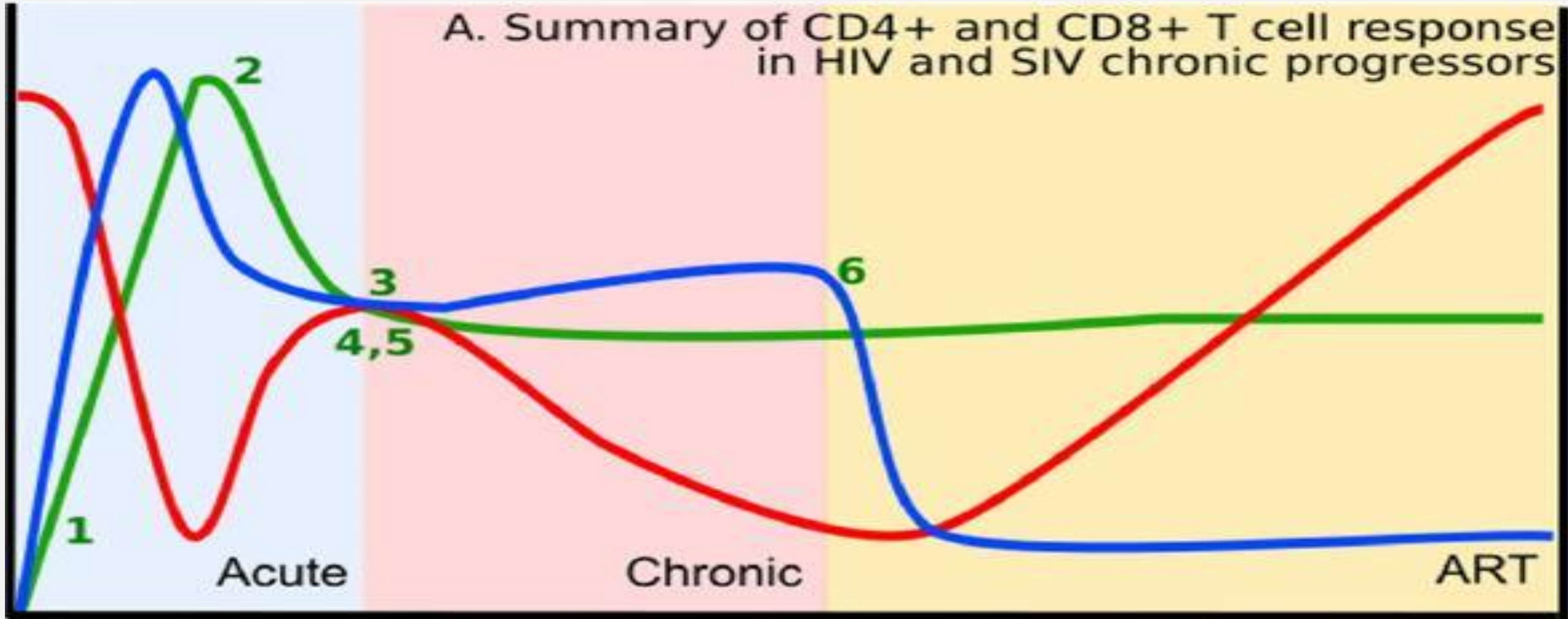
ara kimyasal taşıyıcılar -85mV ile polarizedir. - ol açar. -90mV (K+ dışarı): i bir eşik gerilim değerinin durumlarda hücre ti üretilmesine sinirsel





HIV YAPAY ZEKALARI

A. Summary of CD4+ and CD8+ T cell response in HIV and SIV chronic progressors



- HIV-RNA
- CD4+ T LENFOSİTLER
- CD8+ T LENFOSİTLER

- Güney Carolina'da HIV enfekte hastaların sadece % 66'sı tıbbi takip ve tedavi almaktadır.

- Sürekli...

- ABD UI...

- 4.9 m n...

- 18.998...

- Kamu p...

- ✓ Tak...

- çık...

- ✓ Çıkt...

belirlemek için kullanılacaktır.

Yüksek riskli bireylere yönelik hedefli yardım sağlamaya yardımcı olmaktır.
Algoritmalar, iyi performans sağlamak için yeni verilerin mevcudiyeti temelinde aralıklarla gözden geçirilecektir.



Figure 1 South Carolina Office of Revenue and Fiscal Affairs Integrated Data System.

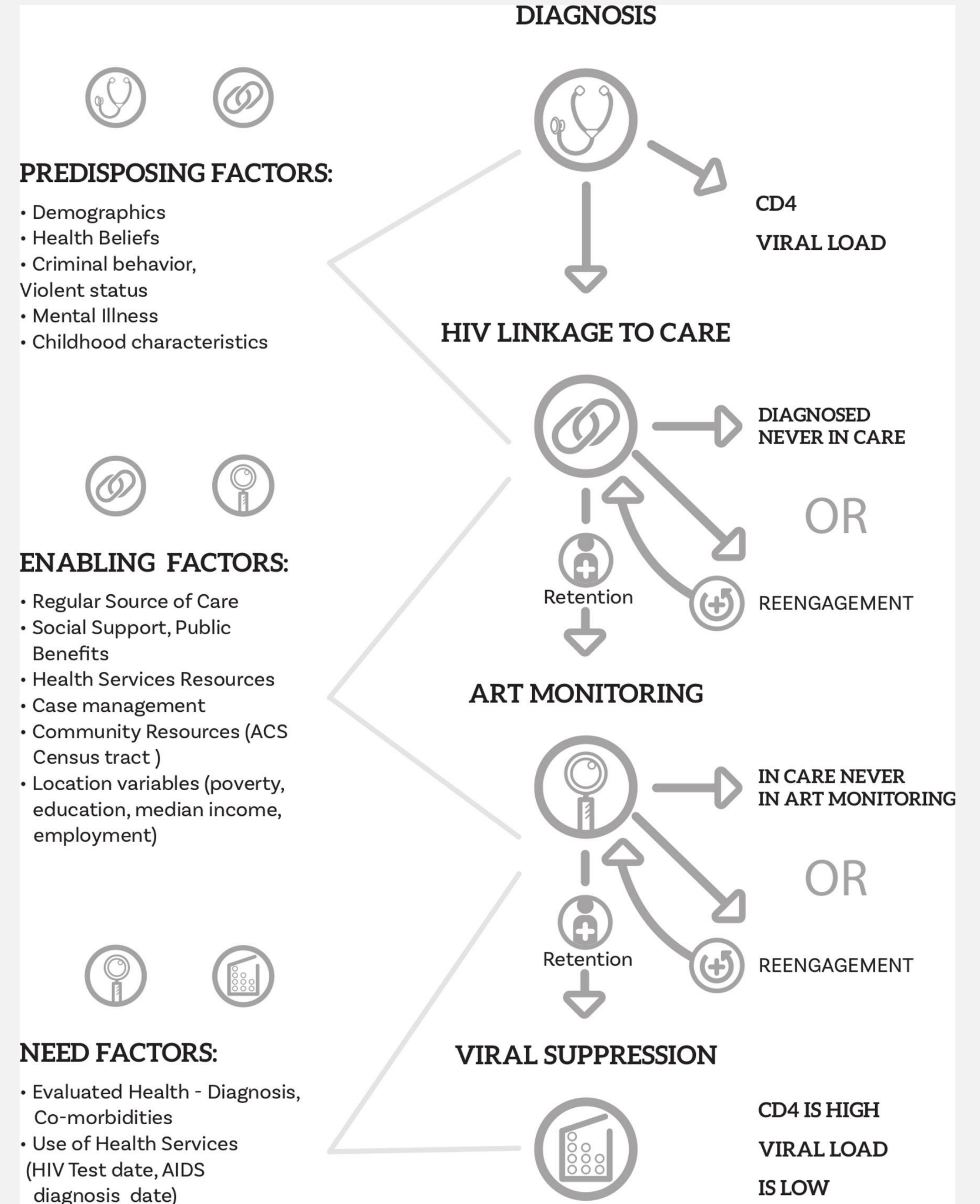
- Random forest- K nearest neighbour modeli makinel öğrenme modeli kullanılarak

- ✓ Klinik takibi umursamayanlar
- ✓ Klinik takibe düzensiz gelenler
- ✓ Klinik takipten çıkmışlar
- ✓ Klinik takipte olan vakalar

- HIV testi öncesi kaçırılan fırsatların nerede ortaya çıktığı

- HIV bakımı ve kaçırılan fırsatlar

- HIV tıbbi bakım arayışı davranışına dayanan hasta profilleri



- İlk örnekte gördüğümüz elektronik sağlık kaydı temelli tahmin modelleri enfekte bireyleri takipte
- Bundan bağımsız olarak HIV enfeksiyonu açısından yüksek risk taşıyan hastaları belirleyebilir.
- Kuzey California'da bir çalışma
- Populasyon: 3 750 664 hasta 2007–17 (3 143 963 öğrenme seti 606 701 uygulama seti)
- Ayaktan hastalardan en az 2 yıllık sağlık planı kaydı olanlardan alınmıştır.
- Temas öncesi profilaksi kullanmayan
 - ✓ Seksüel davranışları sisteme tanımlanmış.
 - ✓ Demografik özellikleri işlenmiş
 - ✓ Cinsel yolla bulaşan hastalıklar testi yaptıranlar, tanısını ve/veya tedavisi alanlar.
- Lojistik regresyon modeli makinel öğrenme

Development dataset, 2007-14 (n=3 143 963)

Validation dataset, 2015-17 (n=606 701)

Sex	Development dataset, 2007-14 (n=3 143 963)	Validation dataset, 2015-17 (n=606 701)
Female	1 681 510 (53.5%)	327 677 (54.0%)
Male	1 462 453 (46.5%)	297 024 (49.0%)

Sexual orientation*

- Data available
- Heterosexual
- Gay or lesbian
- Bisexual
- Unknown

Age (years)

Race or ethnicity

- Data available
- White†
- Hispanic†
- Asian†
- Black†
- Other†
- Unknown

Received care in three cities with high HIV incidence

Resided in one of eight urban ZIP codes with high HIV incidence‡

Incident HIV diagnosis within 3 years

Data are n (%) of

gender of sex partners, as recorded in the electronic health record. †Percentages are of participants with data available. ‡High HIV incidence was defined as being elevated relative to the state or city mean.

Table 1: Demographic characteristics and incident HIV diagnoses within 3 years among patients in the development and validation datasets

Demographics and social history

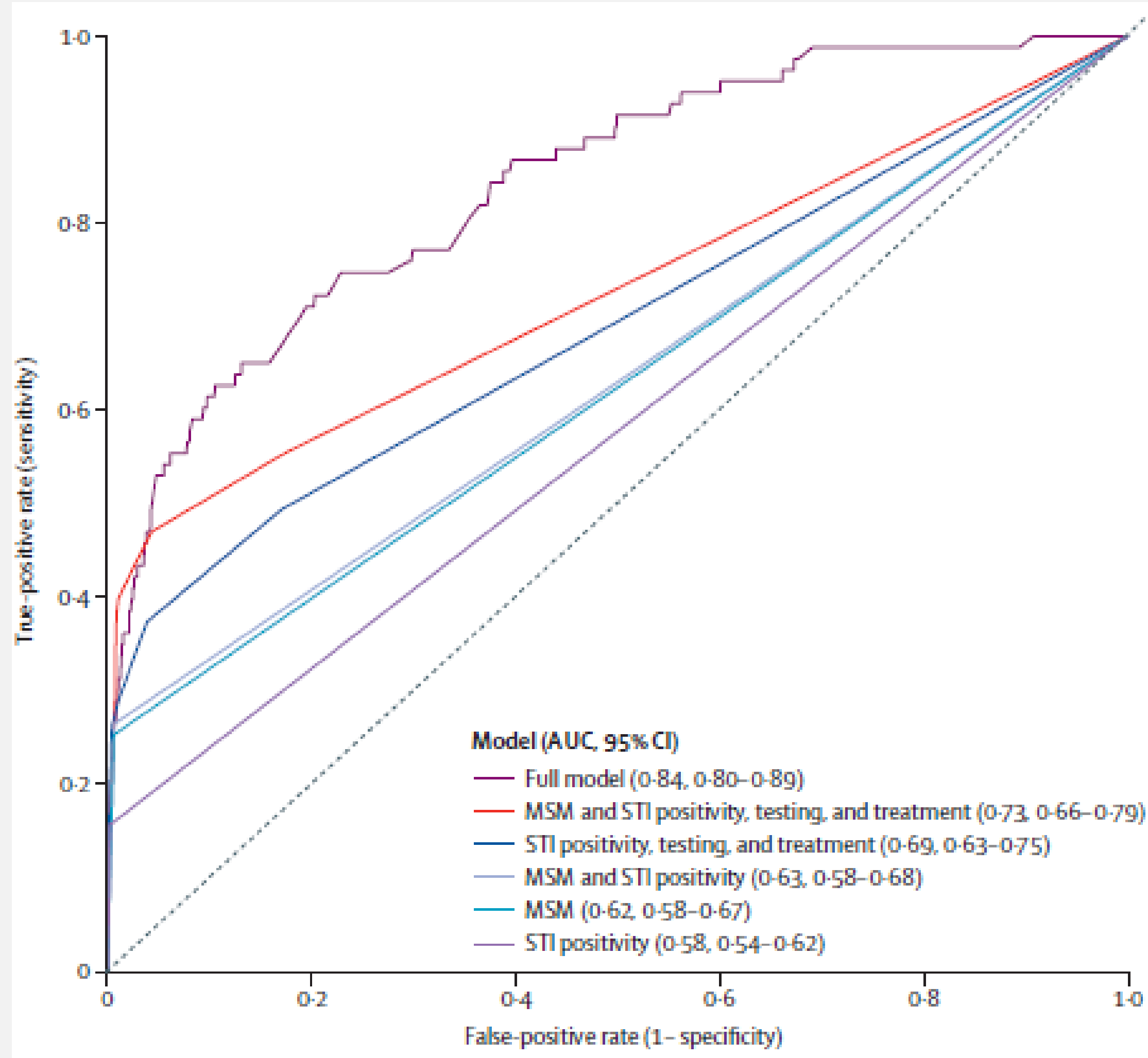
	Predictor prevalence		Adjusted OR
	No incident HIV (n=3 143 262)	Incident HIV (n=701)	
Male sex	1 461 821 (46.5%)	632 (90.2%)	7.2
Men who have sex with men	5 209 (0.2%)	41 (5.8%)	4.7
Sexually active	5 343 303 (17.0%)	77 (11.0%)	0.8
Age (years)			

	Number of patients flagged	Sensitivity for incident HIV cases					Sensitivity for incident PrEP users (n=640)	Specificity	Positive predictive value	Negative predictive value
		Overall (n=83)	Male (n=69)	Female (n=14)	Black (n=20)	White (n=20)				
Full LASSO model	13 463	38.6%	46.4%	0%	25.0%	25.0%	59.1%	97.8%	0.24%	>99.9%
MSM status and STI positivity, testing, and treatment	4 617	28.9%	34.8%	0%	15.0%	35.0%	42.7%	99.2%	0.52%	>99.9%
STI positivity, testing, and treatment	1 458	20.5%	24.6%	0%	10.0%	15.0%	16.2%	99.8%	1.17%	>99.9%
MSM status and STI positivity	3 849	25.3%	30.4%	0%	10.0%	30.0%	38.8%	99.4%	0.55%	>99.9%
MSM status	3 790	25.3%	30.4%	0%	10.0%	30.0%	38.1%	99.4%	0.55%	>99.9%
STI positivity	1 016	6.0%	7.2%	0%	10.0%	0%	6.4%	99.8%	0.49%	>99.9%

High risk scores were defined as a predicted probability of an incident HIV diagnosis within 3 years of 0.2% to <1.0%, and very high risk scores as a predicted probability of ≥1.0%. Specificity, positive predictive value, and negative predictive value refer to the model's ability to classify overall incident HIV cases. PrEP=pre-exposure prophylaxis. LASSO=least absolute shrinkage and selection operator. MSM=men who have sex with men. STI=sexually transmitted infection.

Table 3: Performance of full and simpler HIV risk prediction models among patients with high or very high risk scores in the validation dataset

Any psychiatric diagnosis§	135 317 (4.3%)	49 (7.0%)	1.0
Transgender-related diagnosis	911 (<0.1%)	3 (0.4%)	2.0
High-risk sexual behaviour (homosexual)	27 (<0.1%)	0	0.7
High-risk sexual behaviour (not specified)	453 (<0.1%)	3 (0.4%)	2.6
Exposure to HIV	304 (<0.1%)	4 (0.6%)	1.1
HIV counselling	2858 (0.1%)	7 (1.0%)	1.0
HIV education	6663 (0.2%)	15 (2.1%)	1.1



Modelin risk taşıyan hastaların 3 yıllık tahmini gerçek vaka sayılarını düşük oranda yakaladı.

Buna neden olarak veya çalışma kısıtlılığı olarak:

- Enfekte olup henüz teşhis edilmemiş olanlar yanlış sınıflama ve sapma artışına yol açmış olabilir
- 2005-14 arasında TÖP kullanan hasta sayısı günümüzden azdır.
- En az iki yıllık sağlık kaydı ve düzenli olanlar alındı. Sağlık erişimine sıkıntı yaşayanlar ve düzensiz takip olanlar hata riskini arttırarak çalışma başarısını düşürdü.



- New Mexico, Los Alamos Laboratory, CATNAP database
- 6179 antibody-virus matching results
- 3864 training - 2315 application
- Vaccine development study
- Artificial neural network model machine learning

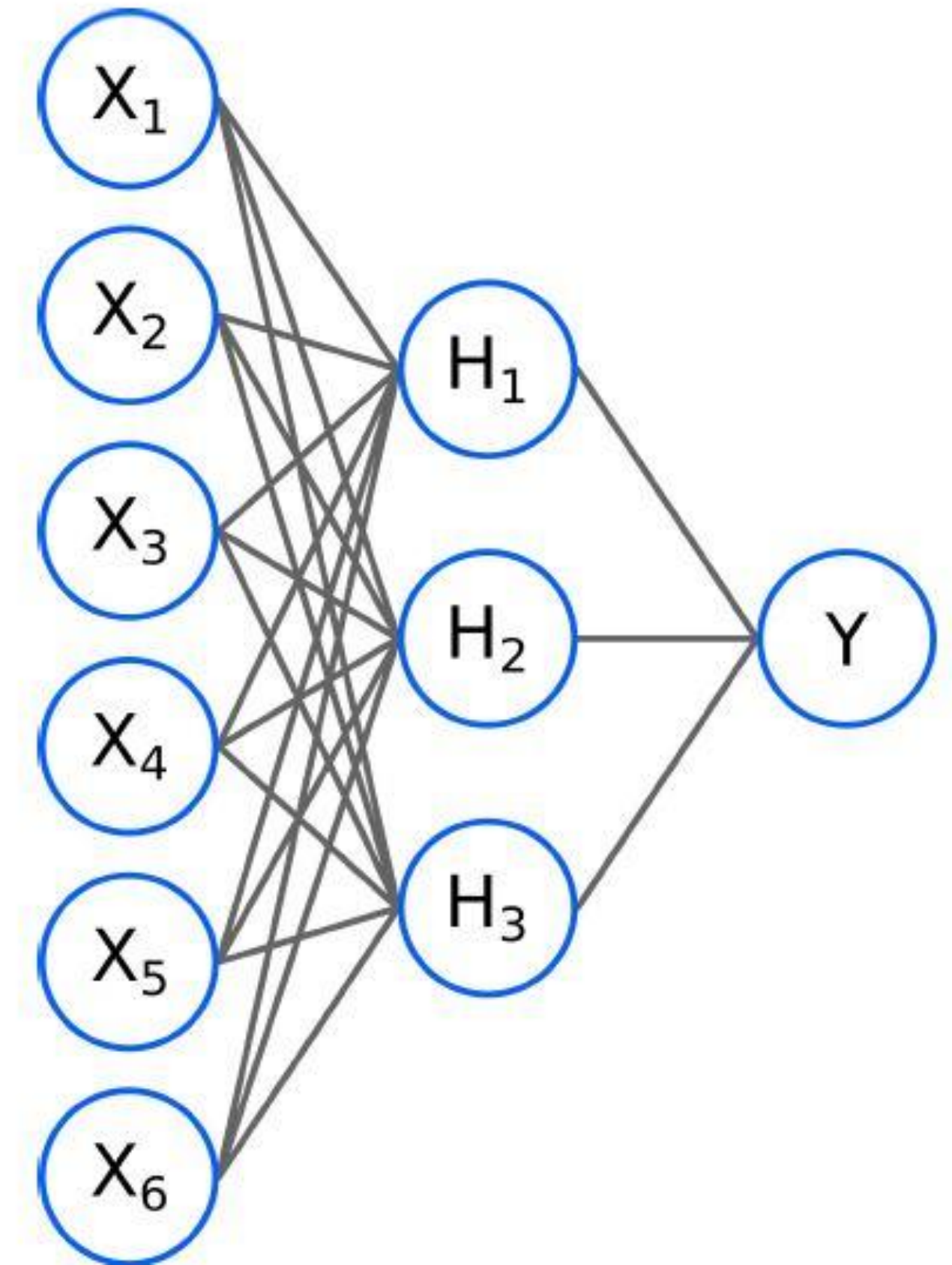


Fig 1. Generic structure of an Artificial Neural Network. It is composed of an input layer (the X_i nodes) that contains the descriptors developed for the system being studied, a hidden layer (H_i nodes) that combines non-linearly the descriptors in the input layer (through a logistic activation function), and an output layer (Y), that again combines non-linearly the nodes in the hidden layer and produces the output result.

<https://doi.org/10.1371/journal.pcbi.1006954.g001>

neural network used to predict the IC_{50} . The training is repeated 30 times using each time a randomly generated training and validation set, and the results are averaged over the 30 neural network results to estimate the error in the predictions.

Aşı geliştirme modelleri

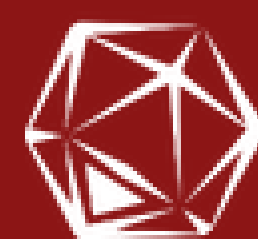
- «Elit Controller» hasta grubunda geniş ölçüde nötrleştirici antikolar adı verilen ve yüksek bir genişliğe sahip olan ve virüsün çeşitli varyantlarını nötrleştirebilen antikolar geliştirir.
- HIV aşı çalışmalarında bu durum önemli bir çalışma kolunu işaret etmektedir.
- Aşı çalışmalarında etkinliğin ölçüsü üretilecek antikoların nötrleştirme genişliğidir.

- Nötralizasyon kabiliyeti, antikoların yarı inhibitör konsantrasyonu (IC50) olarak ölçülmüştür.
- Her antikor / virüs kompleksi için, amino asit dizisi ve ilgili bir kompleksin bilinen bir yapısı kullanılarak model oluşturulur.

- Sisteme işlenen her bir antikorum “nötrleştirme genişliği”, virüsün farklı türevlerini tanıma ve nötrleştirme yeteneğinin bir ölçüsüdür. Yüksek genişliği olan bir antikor, birçok farklı varyantı etkili bir şekilde tanıyabilirken, düşük genişliği olan bir antikor daha spesifiktir. Deneysel olarak, bir antikorum nötralizasyon genişliği, hedef virüsün farklı bölümlerinden bir antijen panelini inhibe etme kabiliyeti test edilerek değerlendirilir

Tedavi esnasında direnci öngören bir model:

- USA/,Atlanta/Georgia State University
- Stanford üniversitesinin kamuya açık antiretroviral direnç sonuçlarına özel veri setleri üzerinde sistem eğitildi. (10.000 üzeri örnek)
- Genotip verilerden fenotipik direnci tahmin etmek için KNN ve RF modelleri kullanıldı.



Stanford University

HIV DRUG RESISTANCE DATABASE

A curated public database to represent, store and analyze HIV drug resistance data.

- [HOME](#)
- [GENOTYPE-RX](#)
- [GENOTYPE-PHENO](#)
- [GENOTYPE-CLINICAL](#)
- [HIVDB PROGRAM](#)
- [ABOUT HIVDB](#)
- [SUPPORT HIVDB!](#)

Genotype-Phenotype Datasets

Last updated on 2019-2-20

High Quality Filtered Datasets

This genotype-phenotype correlation dataset contains isolates on which *in vitro* susceptibility tests were performed using the PhenoSense assay (Monogram, South San Francisco, USA) (Zhang 2005). Redundant viruses obtained from the same individual that contained the same pattern of major drug resistance mutations (defined below) were excluded to minimize bias that would result from over-representing highly similar viruses. Viruses with sequences containing mixtures at these major drug resistance mutation positions were also excluded because the presence of mixtures at these positions may confound genotype-phenotype correlations.

PI Major Drug Resistance Positions: 30, 32, 47, 48, 50, 54, 76, 82, 84, 88

NRTI Major Drug Resistance Positions: 41, 65, 70, 74, 75, 151, 184, 210, 215

NNRTI Major Drug Resistance Positions: 100I, 101P, 103N, 106A/M, 181C/I/A, 188C/L/H, 190A/E/S/Q, 230L

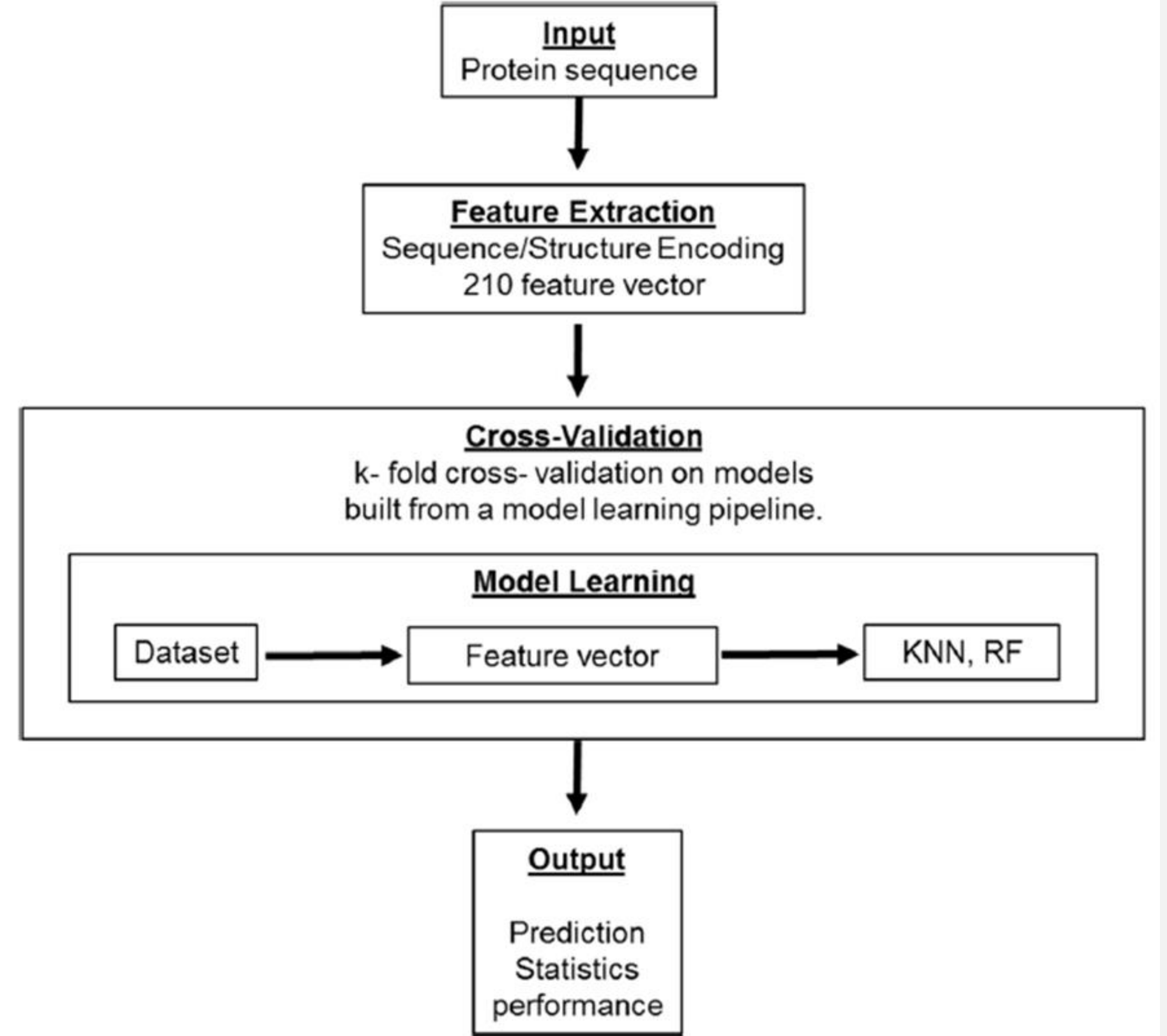
INI Major Drug Resistance Positions: 66A/I/K, 92Q, 118R, 143, 148H/R/K, 155H, 263K

In addition, the dataset can be read in directly over the web to [the R script provided here](#). This R script contains a function which runs least squares regression on this dataset in cross-validations and generates two output files by default: (1) the coefficient and the standard error of each input mutation estimated from cross-validation runs; (2) the mean square error (MSE) estimated in each cross-validation run. The input parameters and additional options to this R function are documented in the script.

To access high quality filtered datasets from HIVDB by drug class, click the links below:

Drug Class	Data
PI	11771 phenotype results from 1951 isolates
NRTI	9682 phenotype results from 1707 isolates
NNRTI	4177 phenotype results from 1812 isolates
INI	1519 phenotype results from 659 isolates

- Birleşik bir dizi ve protein yapısı kodlaması uygulayarak HIV proteaz ve reverse transkriptaz dizilerinden fenotipik sonuçları başarılı bir şekilde tahmin etme üzerine kurulu.
- Makine öğrenme algoritması, temel veritabanı daha fazla veri kazandıkça otomatik güncelleştiriyor.
- Kaynakların sınırlı olduğu ülkeler için
- Bu yaklaşım pratik bir web servisinde uygulanmıştır. Direnç tahmini için web sunucusu, <http://apollo.cs.gsu.edu/bshen/html/index.html> adresinde serbestçe kullanılabilir.



	AZT	DDI	D4T	3TC	ABC	TDF
Accuracy	0.988	0.989	0.991	0.992	0.990	0.986
stddev	0.002	0.001	0.001	0.001	0.001	0.002
Sensitivity	0.984	0.986	0.989	0.988	0.988	0.985
stddev	0.003	0.001	0.002	0.002	0.001	0.002
Specificity	0.991	0.991	0.993	0.995	0.991	0.986
stddev	0.002	0.001	0.001	0.001	0.002	0.003
Run time	98.5	142.7	144.7	143.1	166.3	56.1

KNN YÖNTEMİ

	SQV	LPV	FPV	DRV	ATV	NFV	TPV	IDV
Accuracy	0.973	0.979	0.971	0.989	0.982	0.981	0.985	0.979
stddev	0.003	0.003	0.005	0.003	0.002	0.001	0.002	0.002
Sensitivity	0.965	0.977	0.963	0.988	0.979	0.976	0.986	0.976
stddev	0.005	0.004	0.008	0.005	0.005	0.002	0.004	0.002
Specificity	0.980	0.981	0.980	0.990	0.986	0.985	0.984	0.982
stddev	0.004	0.003	0.005	0.004	0.002	0.002	0.003	0.005
Run time	17.2	18.3	21.0	5.1	18.5	31.8	8.8	26.4

	EFV	NVP	RPV	ETR
Accuracy	0.996	0.996	0.987	0.995
stddev	0.000	0.000	0.001	0.001
Sensitivity	0.996	0.995	0.983	0.992
stddev	0.000	0.001	0.003	0.002
Specificity	0.997	0.997	0.992	0.997
stddev	0.000	0.001	0.003	0.001
Run time	1199.8	1283.7	7.2	48.9

RF YÖNTEMİ

	AZT	DDI	D4T	3TC	ABC	TDF
Accuracy	0.994	0.993	0.994	0.997	0.994	0.992
stddev	0.001	0.001	0.001	0.001	0.000	0.001
Sensitivity	0.994	0.993	0.993	0.997	0.994	0.99
stddev	0.002	0.001	0.002	0.001	0.001	0.003
Specificity	0.995	0.993	0.994	0.997	0.994	0.993
stddev	0.001	0.002	0.001	0.001	0.001	0.002
Run time	8.9	13.6	12.2	9.7	10.7	6.6

	EFV	NVP	RPV	ETR
Accuracy	0.998	0.998	0.989	0.997
stddev	0.000	0.000	0.003	0.000
Sensitivity	0.998	0.998	0.985	0.995
stddev	0.000	0.001	0.006	0.001
Specificity	0.998	0.998	0.993	0.998
stddev	0.000	0.000	0.002	0.000
Run time	67.8	69.3	3.7	8.0

	SQV	LPV	FPV	DRV	ATV	NFV	TPV	IDV
Accuracy	0.984	0.988	0.981	0.992	0.986	0.988	0.988	0.989
stddev	0.002	0.003	0.003	0.004	0.002	0.002	0.004	0.001
Sensitivity	0.983	0.986	0.977	0.993	0.988	0.987	0.987	0.987
stddev	0.002	0.004	0.005	0.004	0.005	0.004	0.007	0.003
Specificity	0.986	0.989	0.984	0.992	0.984	0.99	0.988	0.99
stddev	0.003	0.004	0.001	0.004	0.002	0.003	0.002	0.002
Run time	3.6	3.8	4.0	2.2	4	4.6	2.9	4.3

RESEARCH ARTICLE

Open Access



A comparative study of logistic regression based machine learning techniques for prediction of early virological suppression in antiretroviral initiating HIV patients

Kuteesa R. Bisaso^{1,2,3*} , Susan A. Karungi^{2,3}, Agnes Kiragga¹, Jackson K. Mukonzo² and Barbara Castelnuovo¹

- Uganda'da bir analizde iki bağımsız grubun verileri kullanılmıştır.
- Enfeksiyon Hastalıkları Enstitüsü (IDI) kohort verileri, tahmin modelini eğitmek ve genellenebilirliğini test etmek için kullanılırken, efavirenz (EFV) kohortundan gelen veriler, modelin çalışılan popülasyonun dışında başka bir modelde tahmin etme yeteneğini test etmek için kullanıldı.
- IDI: 484 (öğrenme veriseti)
- EFV kohortu: 233 (eğitim uygulama veriseti)
- Amaç tedavi başlanırken virolojik süpresyon başarısını öngörmek?

- Enfeksiyöz Hastalıklar Enstitüsü (IDI) kohort verileri, tahmin modelini eğitmek ve genellenebilirliğini test etmek için kullanılırken, efavirenz (EFV) kohortundan gelen veriler, modelin çalışılan popülasyonun dışında tahmin etme yeteneğini test etmek için kullanıldı
- IDI verileri Nisan 2004 ve Nisan 2005 tarihleri arasında kaydedilen ardışık 484 HIV hastasından oluşmaktadır.
- Bu hastalar Mulago National referral hospital/kampala (155), Butabika hospital/kampala (60) and Bwera hospital/kasese (n = 47) in the years 2008 and 2009. ART başlanırken 158 hasta TB co-infected idi. ART sonrası sadece 235 hasta 6 ay sonra viral yük kontrolüne geldi
- IDI Stavudin / Lamivudin / Nevirapin (30/300/200 mg) veya (40/300/200 mg) olan üçlü ART rejim kollarından birinde başlatılmış. UGANDA Efavirenz / Zidovudin / Lamivudin (600 / 150/300 mg kolunda devam etmiştir.

Table 1 distribution of variables in the EFV and IDI cohort datasets

Variables	IDI cohort (n = 484)	EFV cohort (n = 233)
WT/kg (median [IQR])	55.0 [48.0–61.0]	51.0 [47.0–58.0]
AGE/ years (median [IQR])	35.0 [30.0–41.0]	33.0 [30.0–40.0]
CD4 / cell per ml (median [IQR])	100.0 [29.7–166.0]	109.0 [46.0–179]
VL*1000 copies per ml (median [IQR])	349 [116.5–595.2]	123.7 [42.7–253.7]
SEX (male %)	30.4	44.6
TB / %(n)	7	57.5
REGIMEN 1 d4T/3TC/NVP-30 (%)	49.5	0
REGIMEN 2 d4T/3TC/NVP-40 (%)	24.7	0
REGIMEN 3 AZT/3TC/EFV (%)	25.6	100

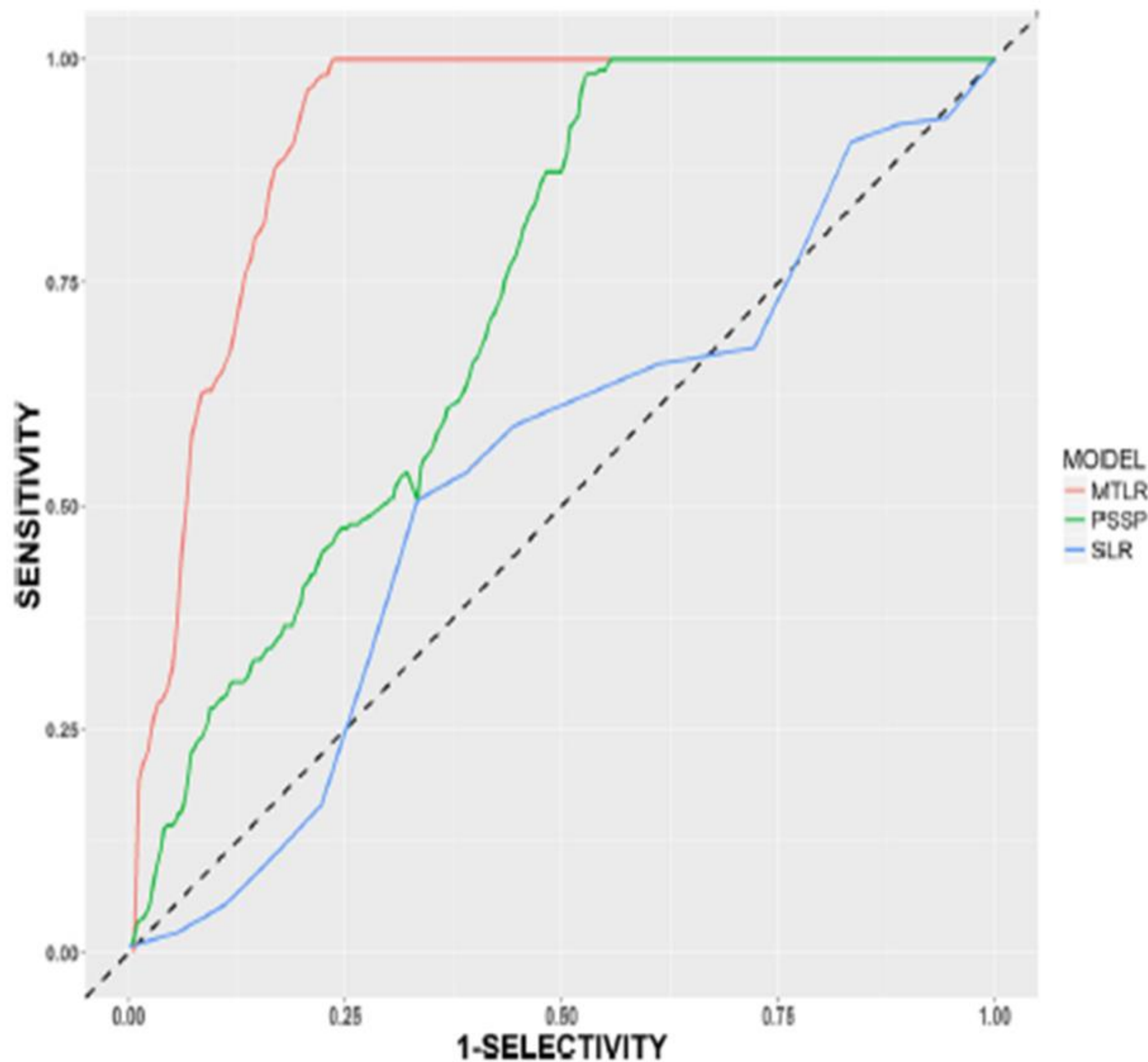


Fig. 2 A receiver operator curve (ROC) and precision recall curve (PRC) showing the model discrimination abilities of outcomes in

Simple logistic regression

In this approach, all data was aggregated together as if the outcome occurred at the same time point (6 months). The outcome variable was set to 1 if viral suppression was achieved or 0 if not achieved at 6 months of ART. Baseline predictors were used to predict the outcome using logistic regression. If y_i and x_i are the observation and its corresponding vector of predictors respectively, such that $y_i \in [0, 1]$ and θ is a vector of coefficients, the probability of virological suppression is given by

$$h_i = P(y_i = 1 | x_i, \theta) = \frac{1}{1 + e^{-\theta^T x_i}} \quad (1)$$

L2-regularization was applied to the model to reduce overfitting. This was accomplished by optimizing the following cost function.

$$\sum_{i=1}^N (y_i \log(h_i) + (1 - y_i) \log(1 - h_i)) + \lambda_1 \|\theta\|_2^2 \quad (2)$$

The hyperparameters λ_1 controls overfitting and N is the total number of individuals in the training dataset.

Multitask temporal logistic regression (MTLR)

Each clinic visitation day was assumed to be a unique learning task for which a logistic regression classification model was trained (fitted) and the task specific parameter (coefficients) and probability of virological suppression learned (estimated). Thus for any task $t \in [1, 2, \dots, M]$, if $y_{i,t}$ and $x_{i,t}$ are the observation and its corresponding feature vector respectively, such that $y_{i,t} \in [0, 1]$ and θ_t is a vector of task specific coefficients, the probability of virological suppression is given by

$$h_{i,t} = P(y_{i,t} = 1 | x_{i,t}, \theta_t) = \frac{1}{1 + e^{-\theta_t^T x_{i,t}}} \quad (3)$$

For each task, overfitting was reduced by explicitly controlling the complexity of the model using L2 regularization as described later. Additionally, the similarity between tasks was leveraged without concealing their uniqueness by applying the multitask learning approach. Specifically, all tasks were learned jointly such that the temporal relation between tasks was enforced. This was accomplished by optimizing the following cost function.

$$\sum_{i=1}^M \left[\sum_{j=1}^{N_i} (y_{i,j} \log(h_{i,j}) + (1 - y_{i,j}) \log(1 - h_{i,j})) + \lambda_1 \|\theta_i\|_2^2 \right] + \lambda_2 \sum_{i=1}^{M-1} \|\theta_{i+1} - \theta_i\|_2^2$$

The first term is likelihood of suppression across all tasks, the second term limits the generalization error via

temporal smoothness on weights from adjacent tasks. The hyperparameters λ_1 and λ_2 control overfitting and temporal smoothness, respectively.

Patient specific survival prediction modeling (PSSP)

In this approach, we formulated the problem as a survival one for each patient using the method developed and described by Yu et al. [33]. The aim was to predict whether or not suppression occurs within 168 days and the time at which it occurs for each patient. The dataset was restructured to include only the 4 most commonly shared observation times namely $t = \{0, 84, 98, 168\}$, also referred to as tasks, $t = \{1, \dots, M\}$, where $M = 4$. Patient outcomes, $y_t \in [0, 1]$ were recorded for each time point, for each patient, capturing the dependence between observations. Thus, if S is the time point at which undetectable viral load is first recorded for the n^{th} patient, then at all $t < S$, $y_t = 0$ while at all $t \geq S$, $y_t = 1$. The elements of the sequence $y = (y_1, y_2, \dots, y_M)$ of outcomes over all four time points were encoded as $y_{t,n}^{(s)}$ for the value at time t , where s is the survival time in the sequence. For our 4 time points, there are 5 possible sequences, including a sequence of all 0 s. The logistic regression method was extended to model the probability of observing the survival status sequence for the n^{th} patient as follows:

$$h_i = p(Y = (y_1, \dots, y_T) | x_n, \Theta) = \frac{e^{\left(\sum_{t=1}^T y_t (\theta_t^T x_n)\right)}}{\sum_{k=0}^M e^{f(x_n, k, \Theta)}} \quad (5)$$

Where Θ is the set of all parameter vectors $(\theta_1, \dots, \theta_M)$ and $f(x, k, \Theta) = \sum_{t=k+1}^M (\theta_t^T x)$ for $0 < k < M$ with viral load becoming undetectable ($y = 1$) in the interval $[t_k, t_{k+1}]$. In order to predict patient specific survival probabilities and times, we optimize the following cost function:

$$\sum_{n=1}^N \left[\sum_{t=1}^M y_{t,n}^{(s)} \log(\theta_t^T x_n) - \log \sum_{k=0}^M e^{f(x_n, k, \Theta)} \right] + \lambda_1 \sum_{t=1}^M \|\theta_t\|_2^2 + \lambda_2 \sum_{t=1}^{M-1} \|\theta_{t+1} - \theta_t\|_2^2 \quad (6)$$

The first term is the log-likelihood of observing a sequence given parameters $\theta = [\theta_1, \dots, \theta_M]$ and baseline predictor variables, x for all N patients. The second term is the L2 regularizer that prevents overfitting and the third term is a regularizer that enforces temporal smoothness on parameters from adjacent observation time points. The hyperparameters λ_1 and λ_2 control overfitting and temporal smoothness, respectively.

Data preparation and model building

Data preparation

The outcome of interest was viral suppression. This was



Teşekkür ederim