

İnfeksiyon Hastalıkları ve Klinik Mikrobiyolojide Yapay Zekanın Kullanım Alanları

Ahmet Görkem Er, MD, PhD Candidate

Hacettepe Üniversitesi Tıp Fakültesi İnfeksiyon Hastalıkları ve Klinik Mikrobiyoloji Anabilim Dalı
Orta Doğu Teknik Üniversitesi Enformatik Enstitüsü Sağlık Bilişimi Anabilim Dalı

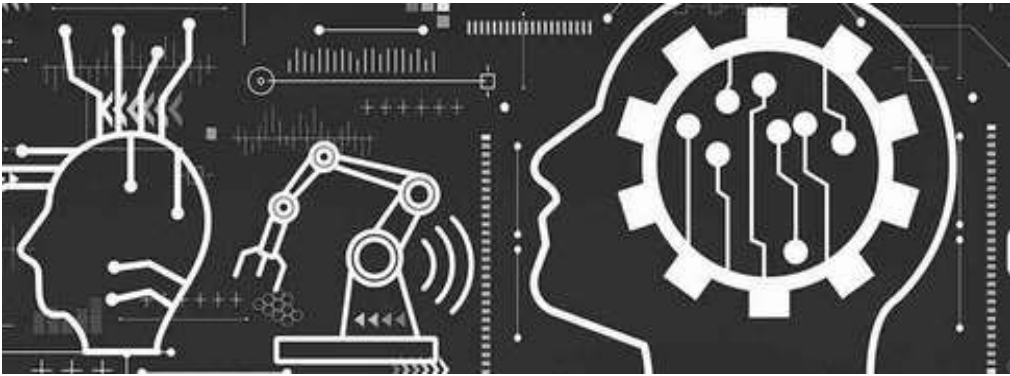
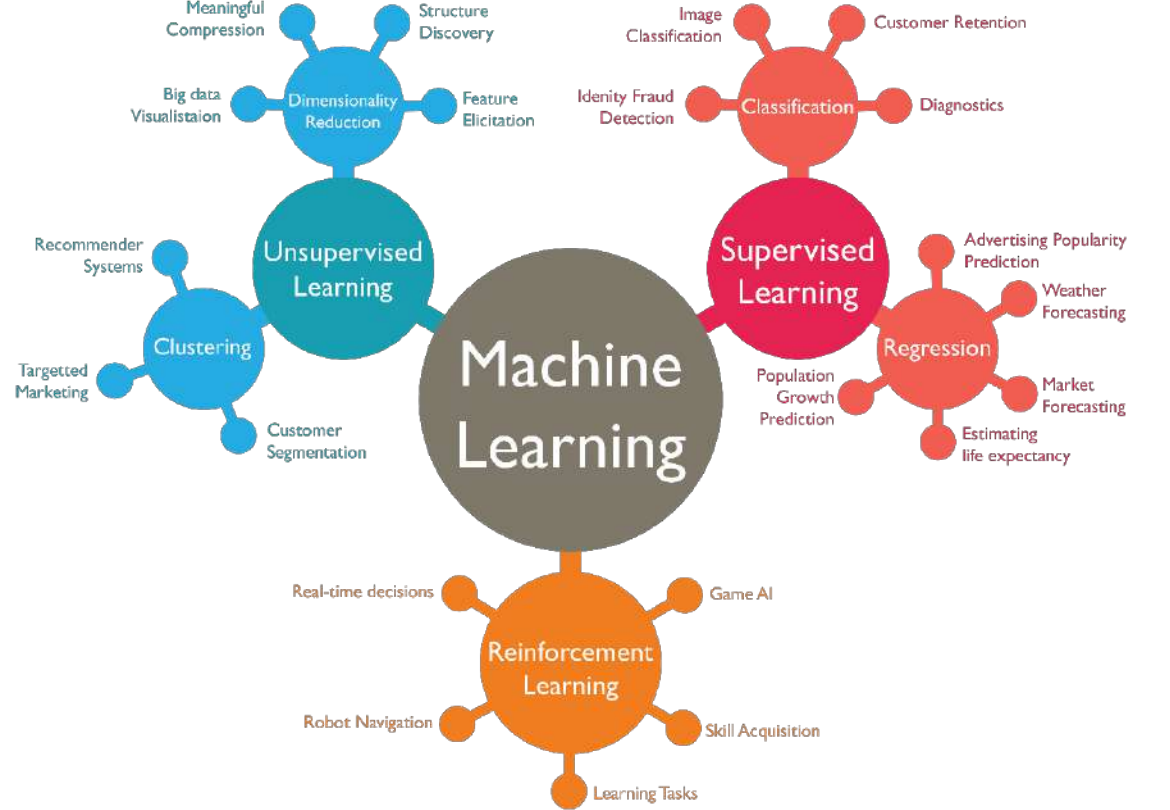
**AGUH 2024
Ankara**

Sunum Planı

- Yapay zeka tanımı
- Neden gerekli?
- Veri çeşitliliği
- Temel algoritma grupları
- İnfeksiyon hastalıklarında uygulama alanları
- Çalışmaları kliniğe uyarlamada zorluklar

Tanım ve Gruplandırma

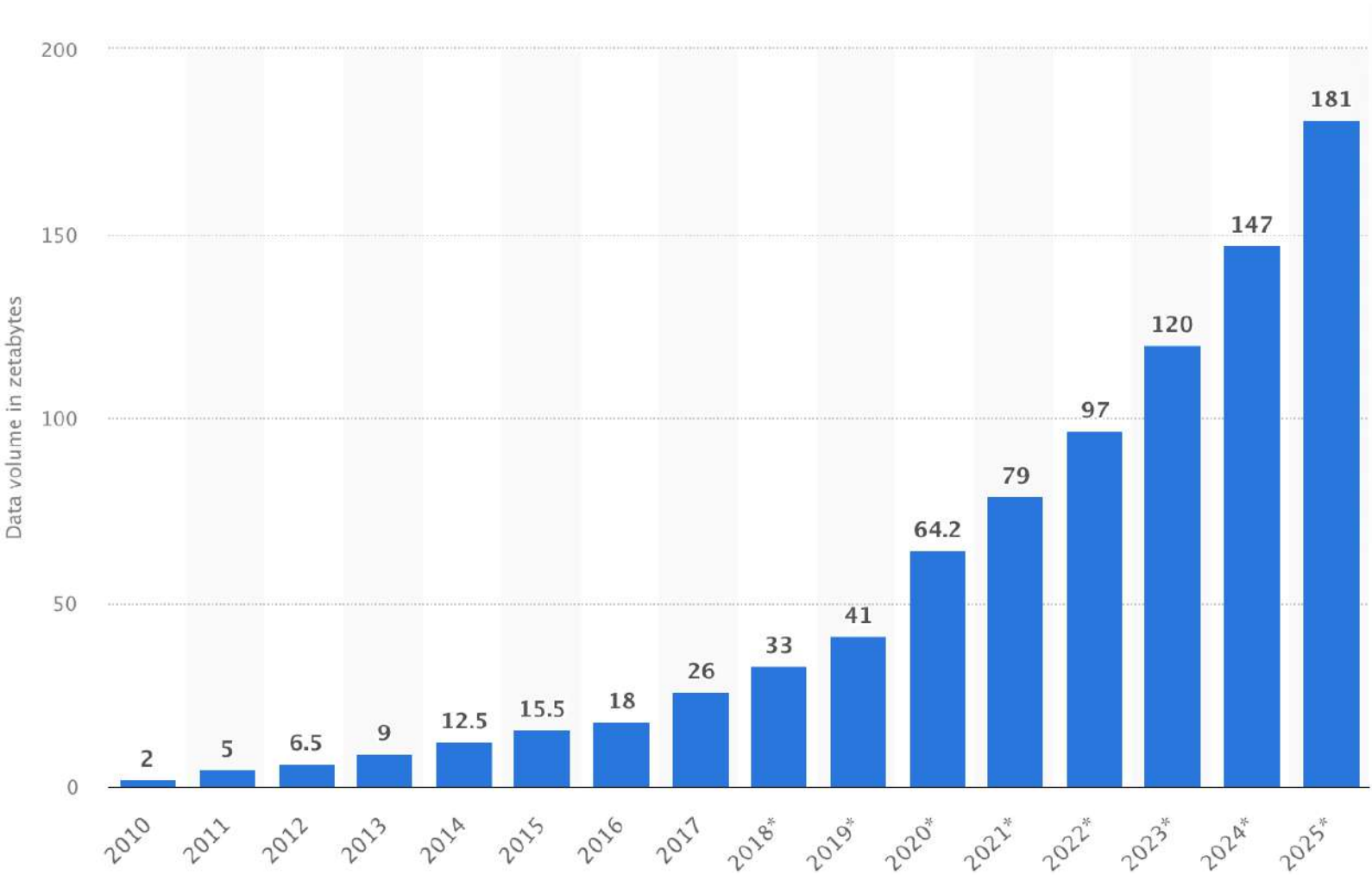
- **Yapay Zeka (Makine Öğrenmesi)**
 - Optimizasyon & Bilgisayar bilimleri
 - Açık programlama olmadan örüntü tespiti ve tahmin yapma
 - Karmaşık ve az tanımlanmış problemler için uygun



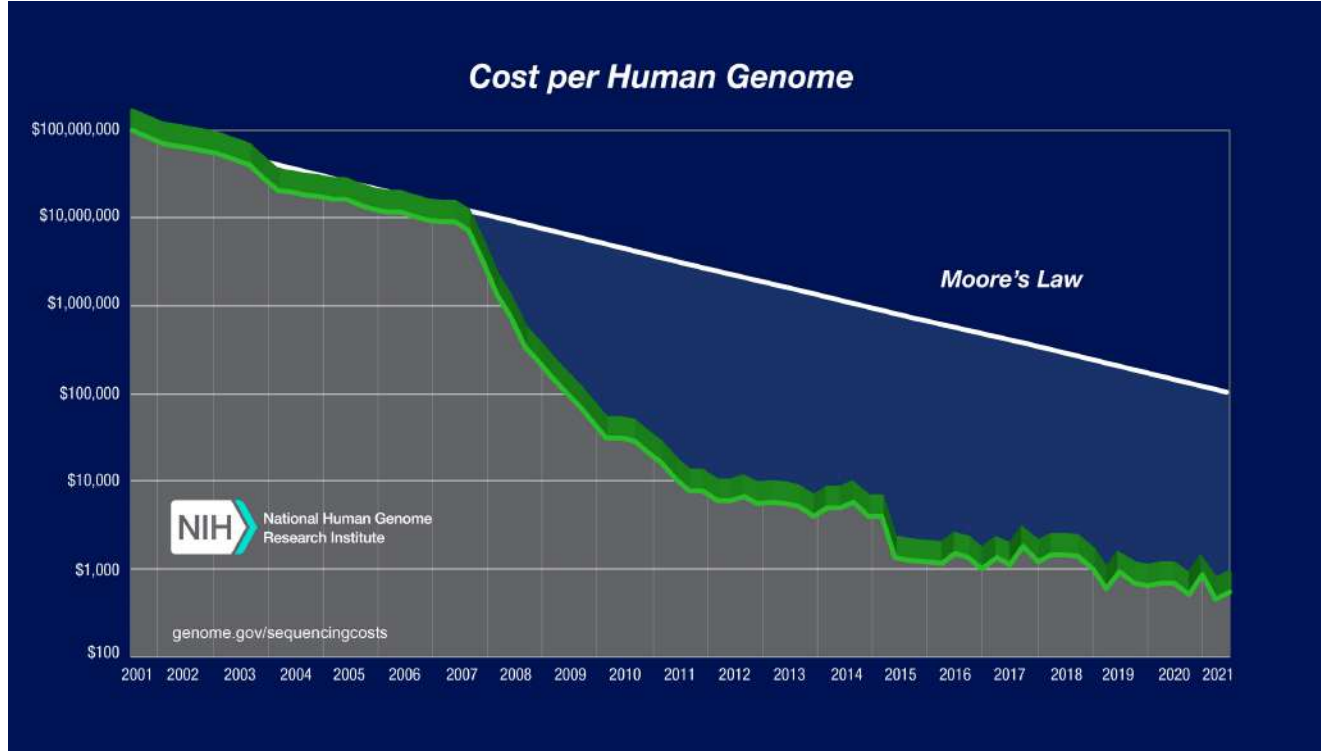
Multifaktöryel Hastalıklar

- Genetik ve çevresel faktörler
- Konak – patojen ilişkisi
- Klinik heterojenite
- Yüksek boyutlu (*high-dimensional*), çok ölçekli (*multi-scale*), çok modlu (*multi-modal*) biyomedikal veriler

Global Veri Hacmi



İnsan Genom Sekansı Başına Fiyatlandırma



- < \$1.000
- 20 yıl öncesine göre 100.000 kat daha az

Veri Kaynakları

- **Tüberküloz**



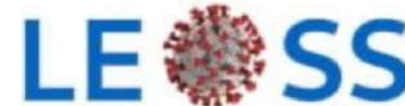
- **Sıtma**



- **HIV**



- **Influenza / COVID-19**



- **Komposit**



- **Yoğun Bakım**



- **Radyoloji**



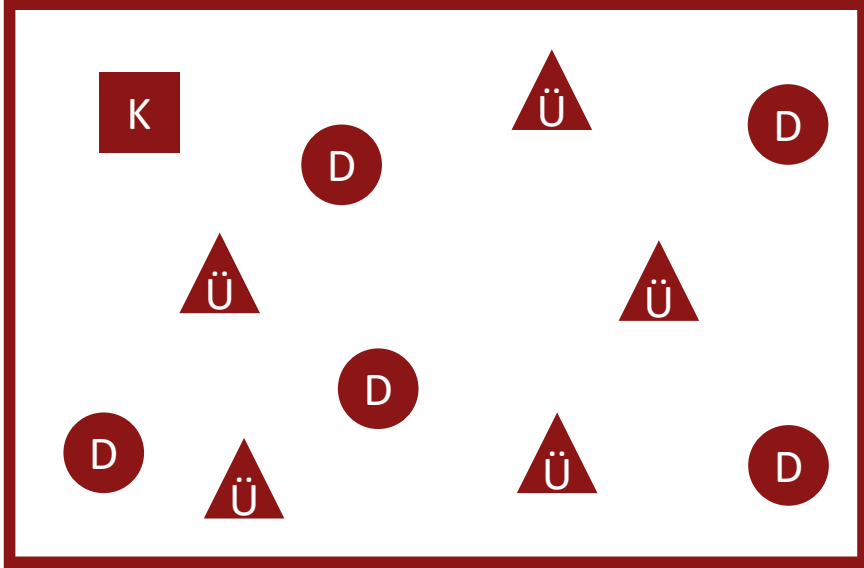
Veri Çeşitleri

- **Klinik veri**
 - Elektronik sağlık kayıtları
 - Laboratuvar sonuçları
 - Klinik notlar
- **Patoloji verisi**
 - Histopatoloji slaytları
 - İmmunofloresans boyalar
 - Tek hücre RNA-seq
 - Spatial transkriptomik veri
- **Moleküler veri**
 - Genomiks
 - Transkriptomiks
 - Proteomiks
 - Epigenetik
- **Görüntü verisi**
 - X-Ray
 - CT
 - MRI
- **Sensör verileri**
 - Giyilebilir cihazlar
 - **EKG**
 - **EEG**
 - **Mekanik ventilatör verileri**
 - **MALDI-TOF**

Veri matrisi, vektörizasyon vs.

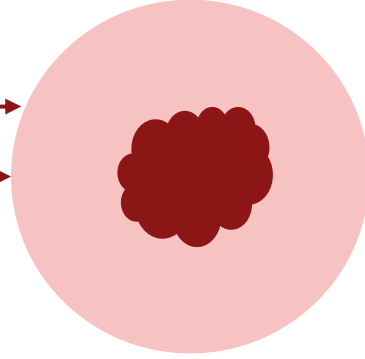
Supervised Learning (Denetimli Öğrenme)

Eğitim Seti

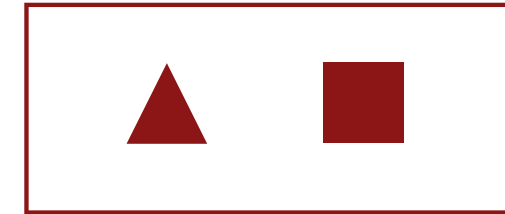
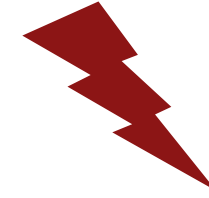


● Daire ▲ Üçgen ■ Kare

Lojistik regresyon
(Random forests,
derin öğrenme)



Sınıflandırma
(Regresyon)

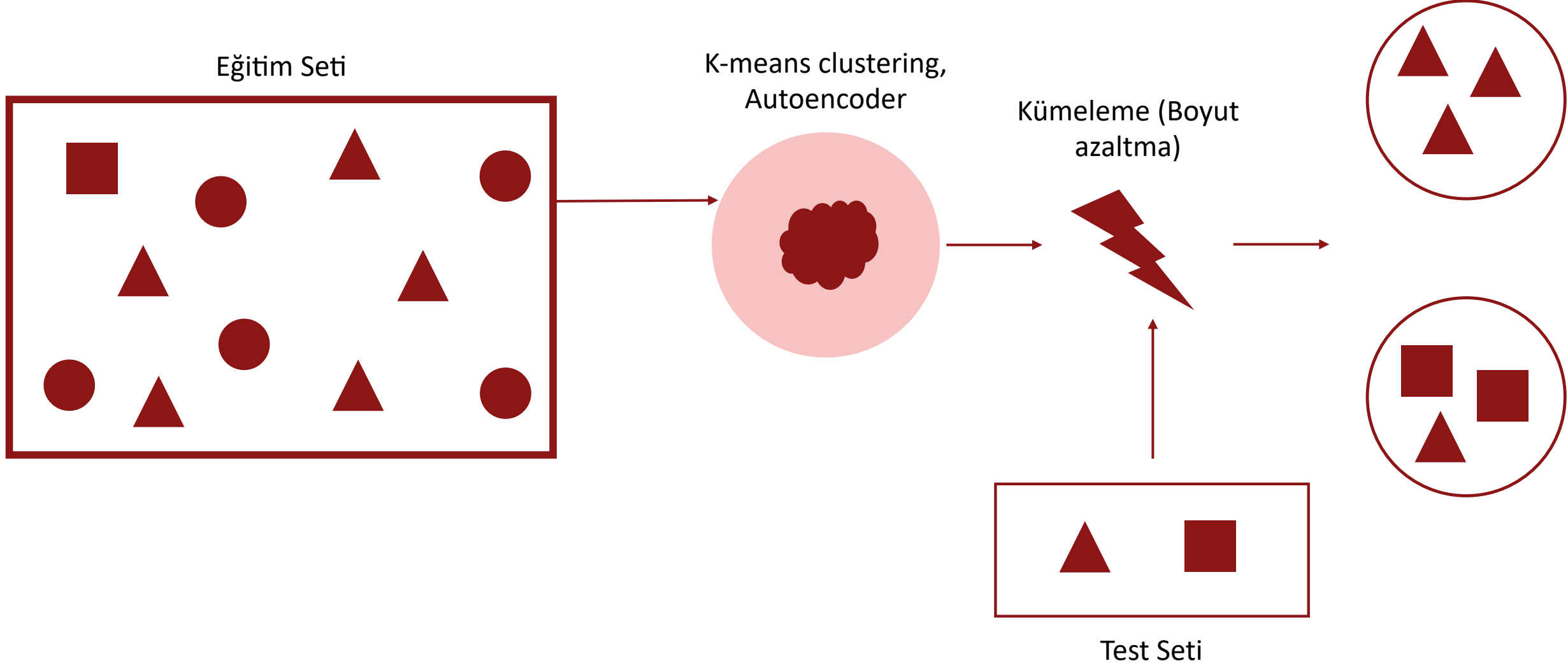


Test Seti

Üçgen

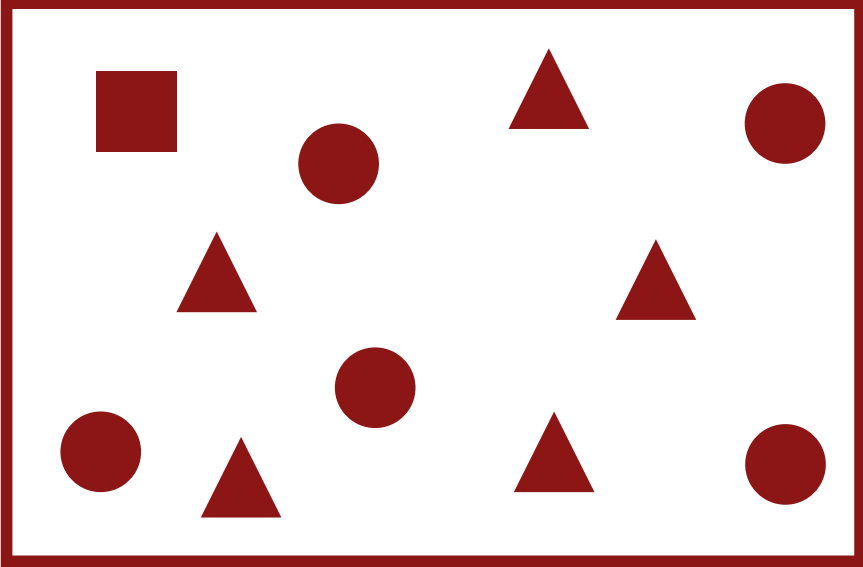
Kare

Unsupervised Learning (Denetimsiz Öğrenme)

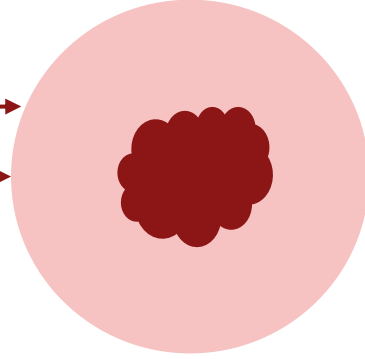


Semi-Supervised Learning (Yarı Denetimli Öğrenme)

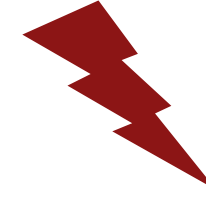
Etiketli olmayan geniş eğitim seti



DINO, CLIP,
Generative AI



Sınıflandırma
(Regresyon)

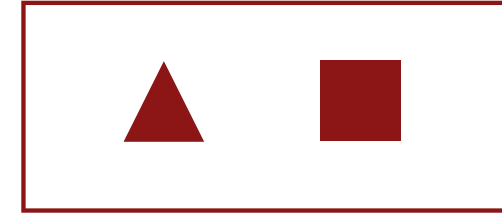


Üçgen

Kare

● Daire ▲ Üçgen ■ Kare

Az sayıda etiketli veri seti



Test Seti

SARS-CoV-2 Genomunun Tespiti

- Haftalar içerisinde viral genom sekans verisi üretildi ve dünyayla paylaşıldı

Article | [Open access](#) | [Published: 03 February 2020](#)

A new coronavirus associated with human respiratory disease in China

[Fan Wu](#), [Su Zhao](#), [Bin Yu](#), [Yan-Mei Chen](#), [Wen Wang](#), [Zhi-Gang Song](#), [Yi Hu](#), [Zhao-Wu Tao](#), [Jun-Hua Tian](#), [Yuan-Yuan Pei](#), [Ming-Li Yuan](#), [Yu-Ling Zhang](#), [Fa-Hui Dai](#), [Yi Liu](#), [Qi-Min Wang](#), [Jiao-Jiao Zheng](#), [Lin Xu](#), [Edward C. Holmes](#) & [Yong-Zhen Zhang](#) 

[Nature](#) **579**, 265–269 (2020) | [Cite this article](#)

Tanı Araçları

- **Mikroorganizma tespiti**

- Otomatize plak değerlendirme sistemi (Clever Culture Systems®, İsviçre)
- Plaka üzerindeki canlı kolonileri tespit etmek için dijital görüntü analizi
- %90'ın üzerinde hassasiyet ve özgüllük, saatte 200 plaka
- MRSA modülü: 100% NPV, 60% PPV



- **Antimikrobiyal direnç tespiti**

- 300,000'den fazla kütle spektrumu ve 750,000'den fazla AMR fenotipinden oluşan bir veri kümesi
- MALDI-TOF ile birleştirilmiş ML, *S. aureus* için AUROC 0.80, *E. coli* için 0.74 ve *K. pneumoniae* için 0.70

Theodosiou AA, Read RC. Artificial intelligence, machine learning and deep learning: Potential resources for the infection clinician. *J Infect.* 2023 Oct;87(4):287-294. doi: 10.1016/j.jinf.2023.07.006. Epub 2023 Jul 17. PMID: 37468046.

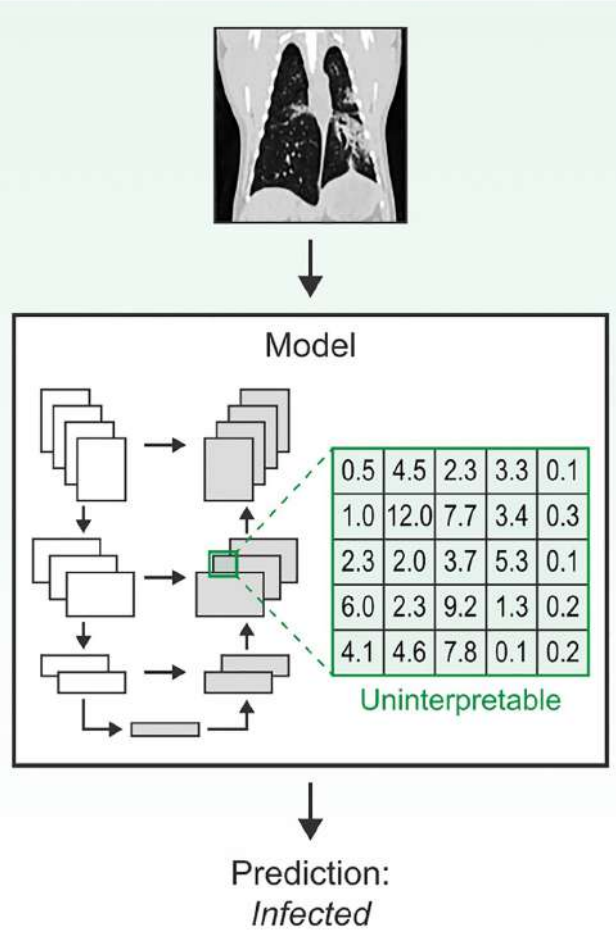
Weis, C., Cuénod, A., Rieck, B. *et al.* Direct antimicrobial resistance prediction from clinical MALDI-TOF mass spectra using machine learning. *Nat Med* **28**, 164–174 (2022). <https://doi.org/10.1038/s41591-021-01619-9>

Table 1. Performance of machine learning across different studies in predicting antibiotic susceptibility patterns.

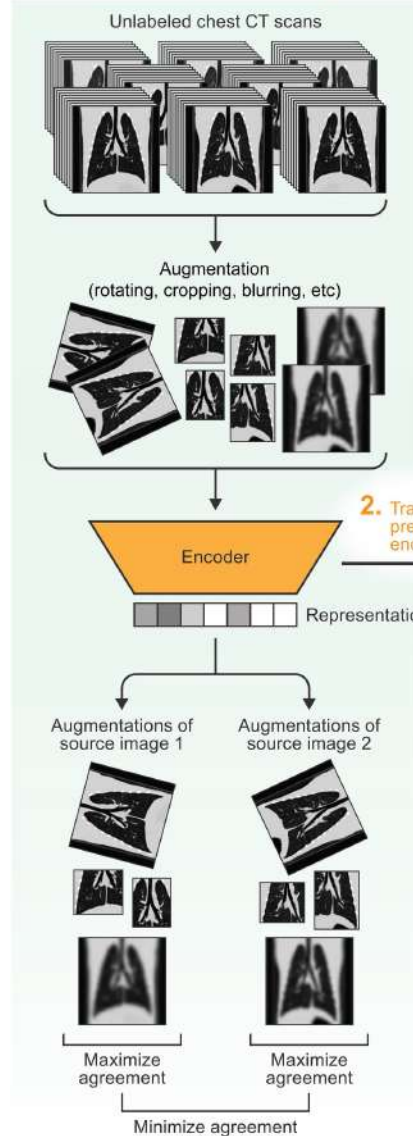
Authors	Year of Publication	Medical Setting	Geographical Setting	Input Data	ML Algorithms	Performance Evaluation	Bacterial Species
Goodman et al. [27]	2016	Hospital admissions	USA	Blood cultures/AST	Recursive partitioning, DT	PPV 0.908-NPV 0.919	<i>Escherichia coli</i> , <i>Klebsiella pneumoniae</i> , <i>Klebsiella oxytoca</i>
Vazquez-Guillamet et al. [29]	2017	Hospital admissions	USA	EHR data/Blood cultures/AST	Recursive partitioning, DT	AUC 0.61–0.80	GNB
Sousa et al. [28]	2019	Hospital admissions	Spain	Clinical/demographic data/Blood cultures/AST	DT	AUC 0.76	BL-GNB
Moran et al. [20]	2020	Hospital admissions and primary care	UK	Blood/urine cultures	XGBoost	AUC 0.70	<i>Escherichia coli</i> , <i>Klebsiella pneumoniae</i> and <i>Pseudomonas aeruginosa</i>
Feretzakis et al. [33]	2020	Medical wards	Greece	Demographics/Cultures/AST/Bacterial Gram stain/Type of sample	MLR	AUC 0.758	All isolated bacterial species
Feretzakis et al. [34]	2020	Intensive Care Unit	Greece	Demographics/Cultures/AST/Bacterial Gram stain/Type of sample	LR, RF, k-NN, J48, MLP	AUC 0.726	All isolated bacterial species
Feretzakis et al. [35]	2021	Intensive Care Unit	Greece	Demographics/Cultures/AST/Bacterial Gram stain/Type of sample	JRip, RF, MLP, Class. Regr, REPTree	F-measure 0.884, AUC 0.933	<i>Pseudomonas aeruginosa</i> , <i>Acinetobacter baumannii</i> , <i>Klebsiella pneumoniae</i>
Martínez-Agüero et al. [36]	2019	Intensive Care Unit	Spain	Demographics/Clinical data/Type of sample/Cultures/AST	LR, k-NN, DT, RF, MLP	Accuracy for quinolone resistance 88.1 ± 1.6	<i>Pseudomonas</i> , <i>Stenotrophomonas</i> , <i>Enterococcus</i>
McGuire et al. [5]	2021	Hospital admissions	USA	Demographic, medication, vital sign, laboratory, billing code, procedure, culture, and sensitivity data (67 features)	XGBoost	AUC 0.846	Bacterial isolates with CR
Pascual-Sánchez et al. [32]	2021	Intensive Care Unit	Spain	EHR data	LR, DT, RF, XGBoost, MLP	AUC 0.76	MDR bacteria
García-Vidal et al. [31]	2021	FN Hematological Patients	Spain	EHR data	RF, GBM, XGBoost, GLM	AUC 0.79	MDR- <i>Pseudomonas aeruginosa</i> / ESBL-E
Henderson et al. [30]	2022	HIV patients	USA	EHR data	PLR, naïve Bayes, gradient boosting, SVM, RF	AUC 0.70	MDR-E

Görüntüleme Araçları

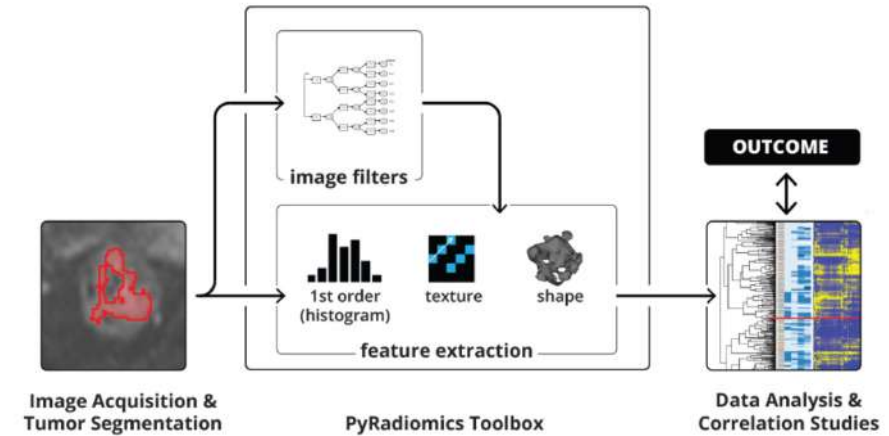
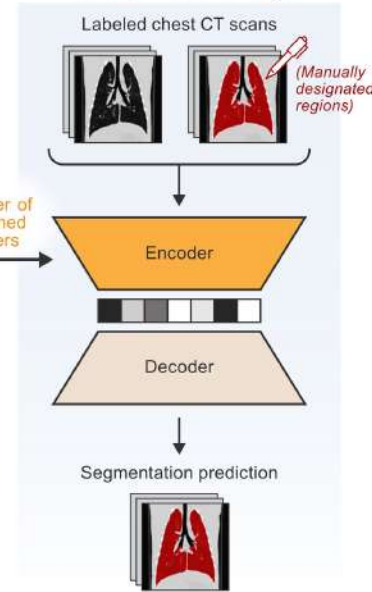
Standard deep learning model



1. Self-supervised learning



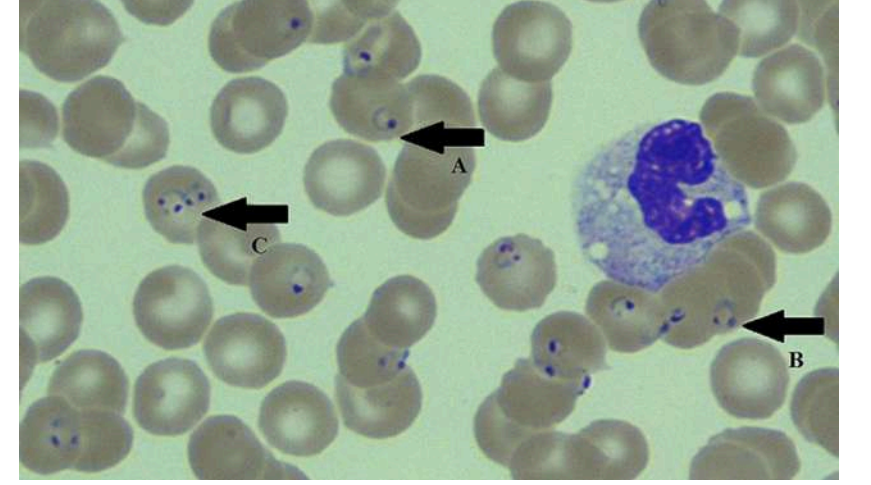
3. Supervised learning



Görüntüleme Araçları

- **Akciğer TB ve sıtma tanısı**

- 125,959 hasta, TB için %91 duyarlılık, %65 özgüllük
- Kan yaymalarında sıtma tespiti için %99 doğruluk



- **COVID-19 klinik fenotipleri ortaya koyma ve ağır hasta tespiti**

- 149 hasta, CT + viral genom sekansı + laboratuvar + klinik

Theodosiou AA, Read RC. Artificial intelligence, machine learning and deep learning: Potential resources for the infection clinician. *J Infect.* 2023 Oct;87(4):287-294. doi: 10.1016/j.jinf.2023.07.006. Epub 2023 Jul 17. PMID: 37468046.

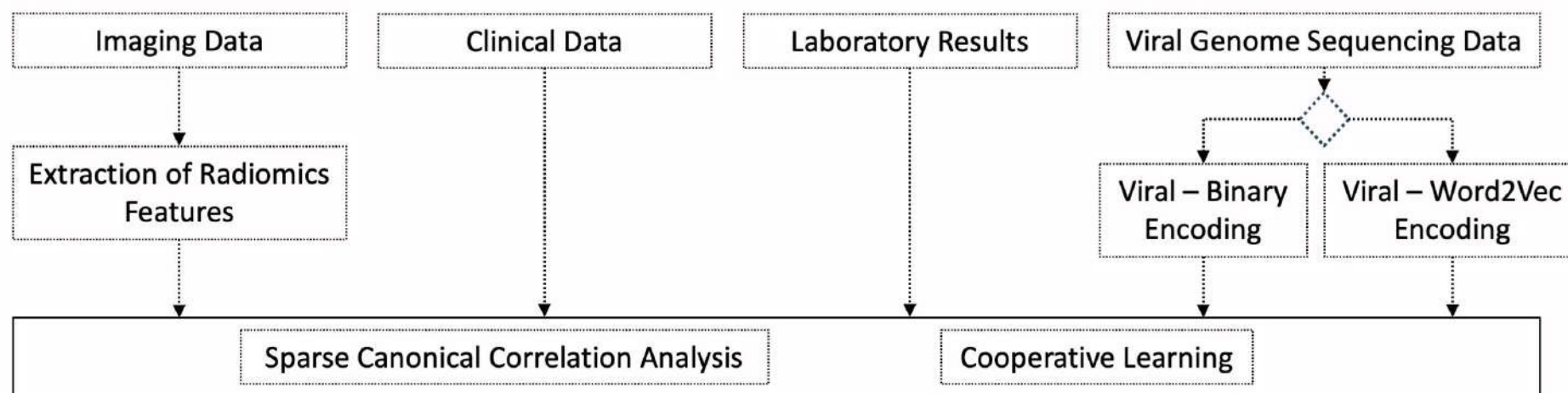
Er AG, Ding DY, Er B, et al. Multimodal data fusion using sparse canonical correlation analysis and cooperative learning: a COVID-19 cohort study. *NPJ Digit Med.* 2024;7(1):117. Published 2024 May 7. doi:10.1038/s41746-024-01128-2



Multimodal data fusion using sparse canonical correlation analysis and cooperative learning: a COVID-19 cohort study

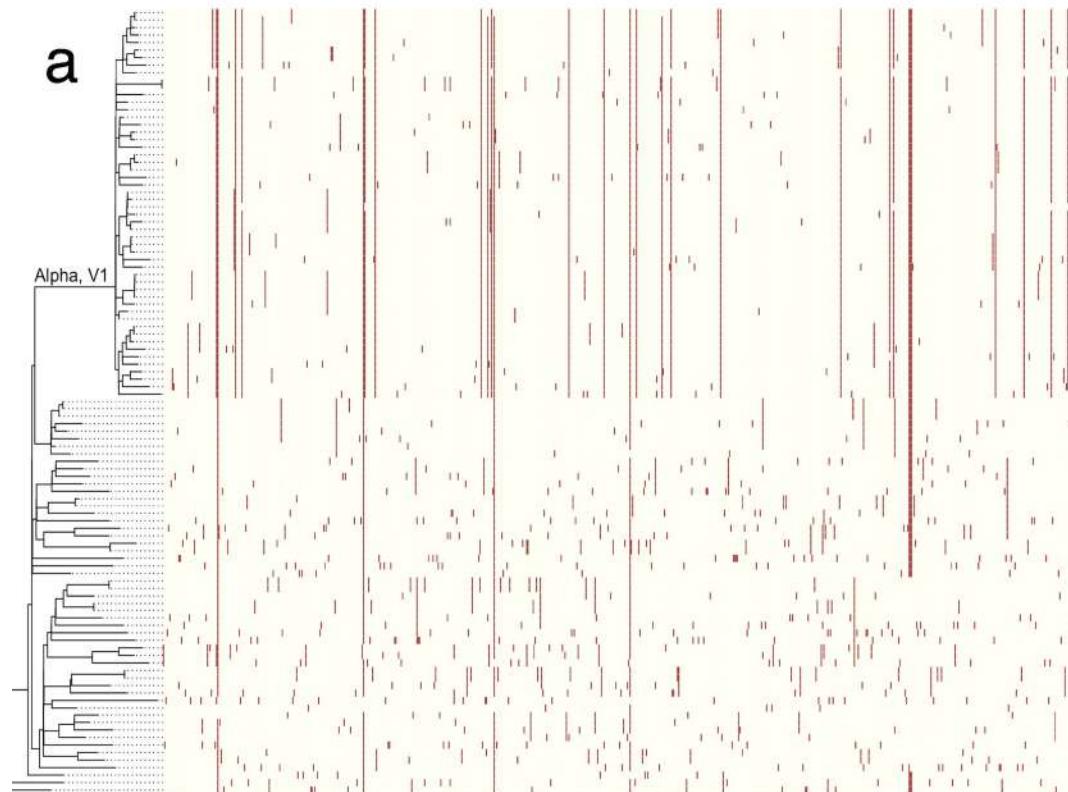
[Ahmet Gorkem Er](#) , [Daisy Yi Ding](#), [Berrin Er](#), [Mertcan Uzun](#), [Mehmet Cakmak](#), [Christoph Sadee](#), [Gamze Durhan](#), [Mustafa Nasuh Ozmen](#), [Mine Durusu Tanriover](#), [Arzu Topeli](#), [Yesim Aydin Son](#), [Robert Tibshirani](#), [Serhat Unal](#) & [Olivier Gevaert](#) 

- 149 hasta

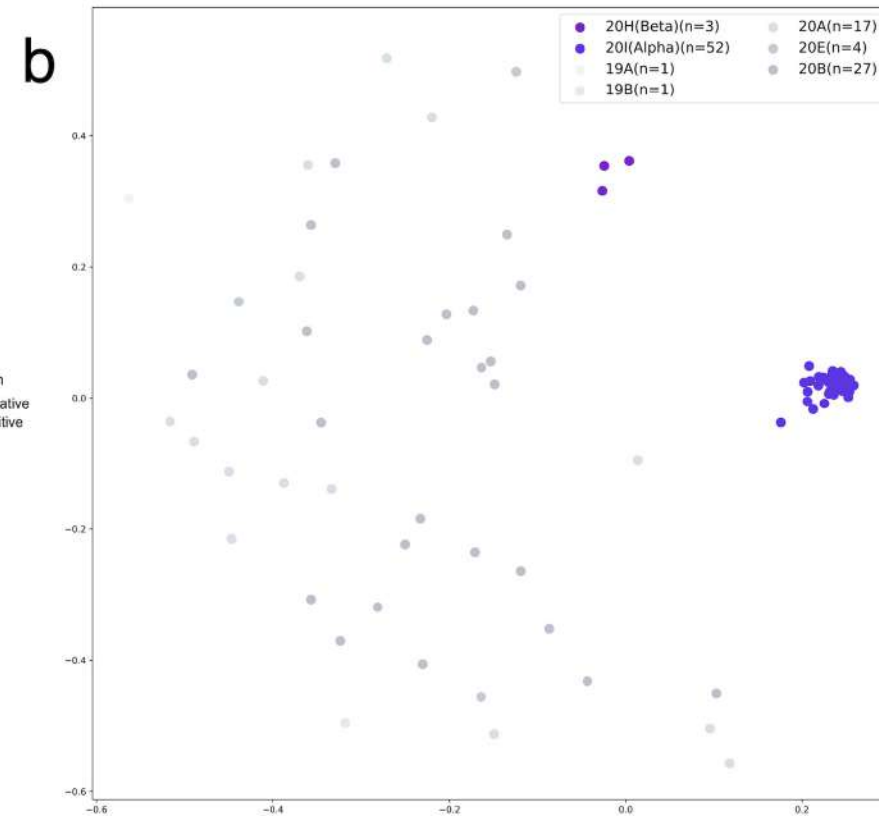


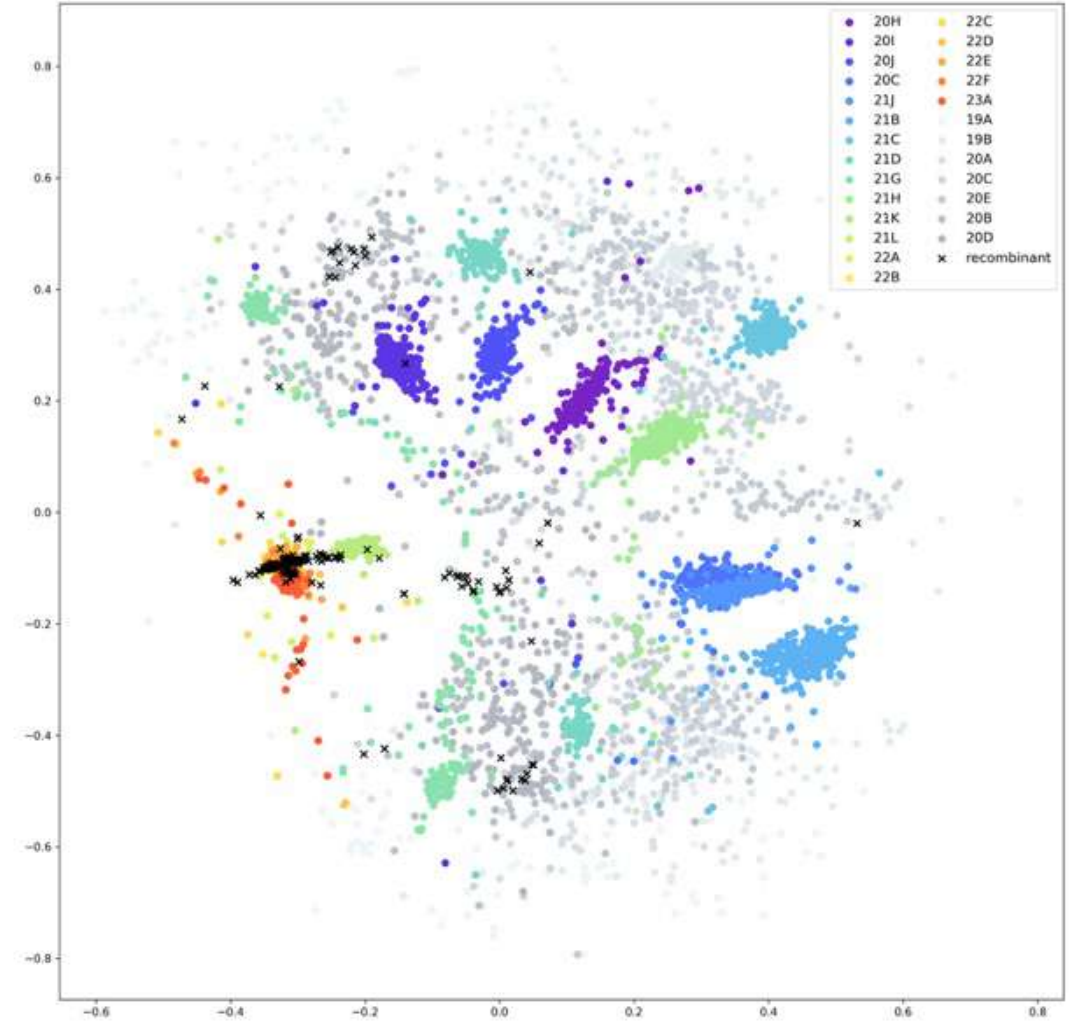
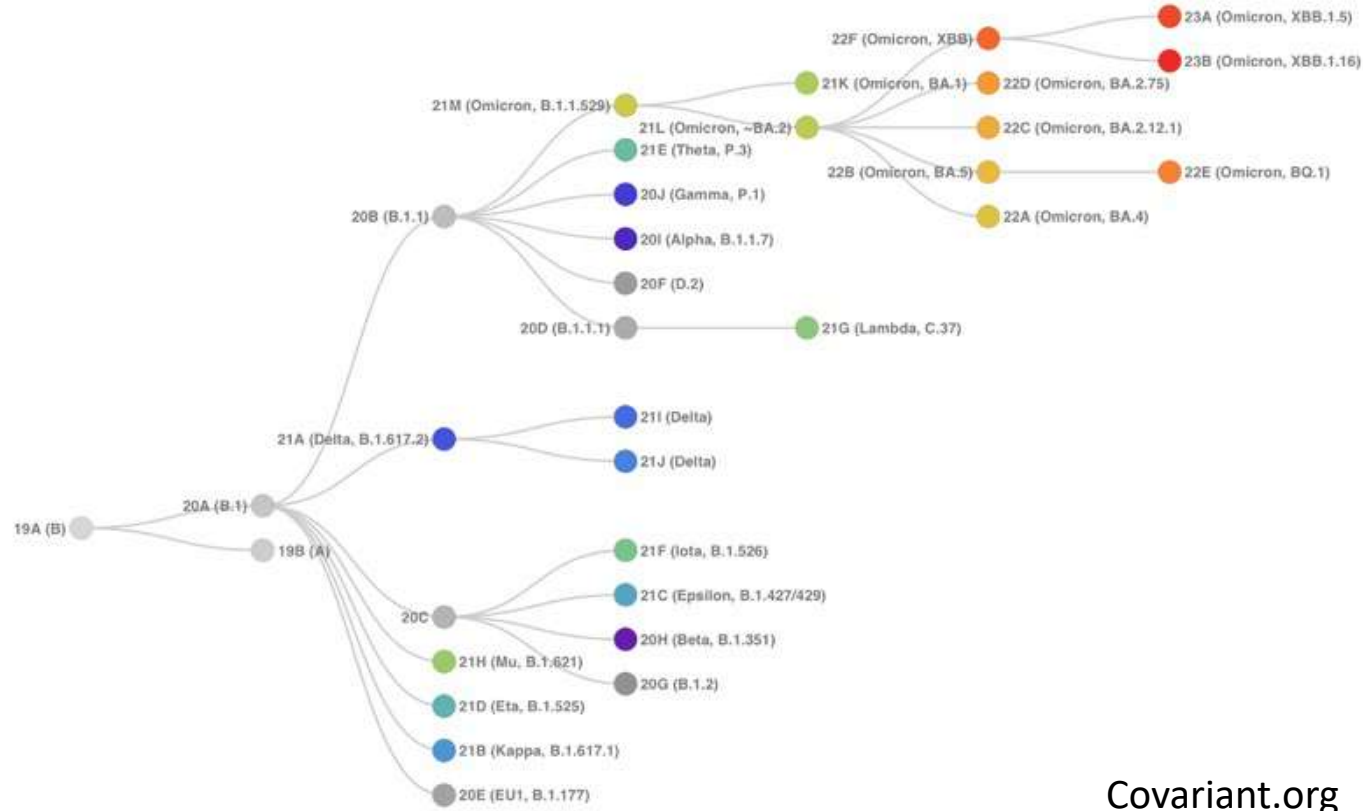


Viral-Binary Encoding



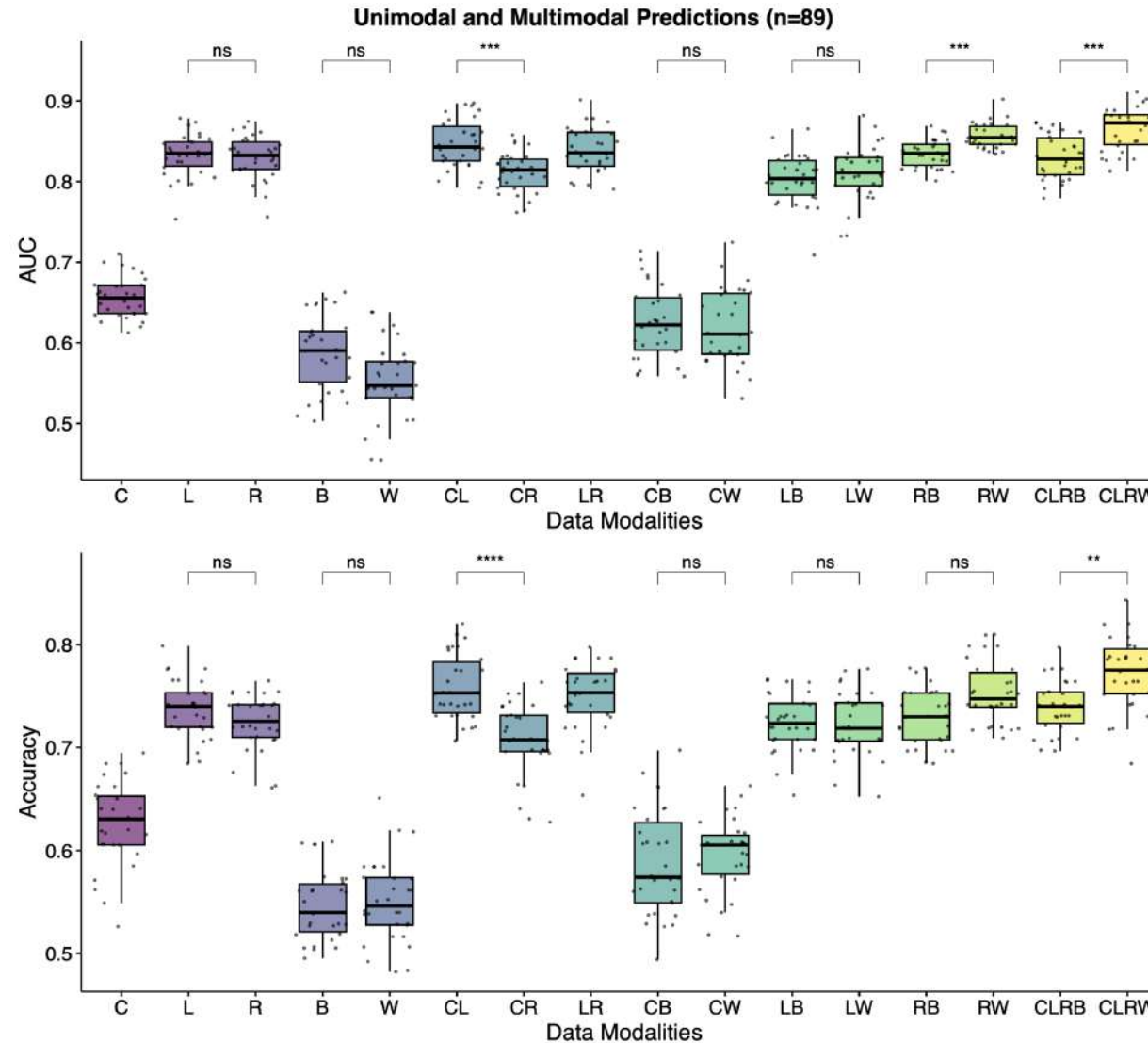
Viral-Word2Vec Encoding





- Major sınıflar başarıyla ayrışıyor
- Öncül varyantlar geniş bir dağılıma sahip
- Omicron diğer varyantlardan ayrışıyor
- Rekombinant suşlar

Kooperatif Öğrenme



Kooperatif Öğrenme



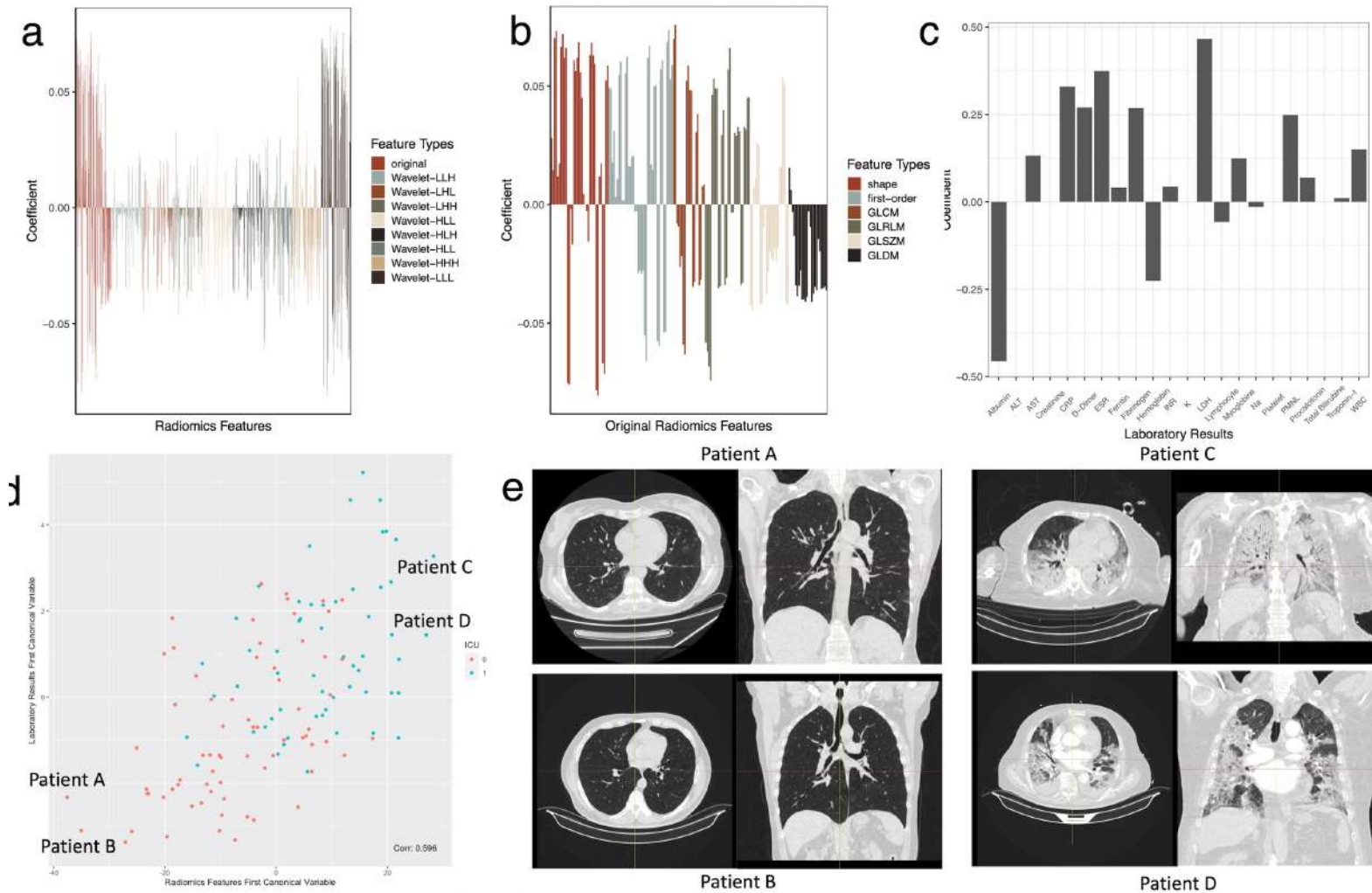
Supplementary Table 2: Unimodal and Multimodal Prediction Models Evaluated in 89 Patients with Post-hoc Analysis Results (n=89)

Model	Accuracy	Adjusted p-values															
		C	L	R	D	W	CL	CR	LR	CD	CW	LD	LW	RD	RW	CLRD	CLRW
C	0.63 ± 0.04	-	.000	.000	.000	.000	.000	.000	.000	.138	.167	.000	.000	.000	.000	.000	.000
L	0.74 ± 0.03	.000	-	.874	.000	.000	.366	.03	.958	.000	.000	.85	.824	.999	.774	1	.005
R	0.72 ± 0.03	.000	.874	-	.000	.000	.003	.731	.082	.000	.000	1	1	1	.019	.693	.000
B	0.55 ± 0.03	.000	.000	.000	-	1	.000	.000	.000	.036	.000	.000	.000	.000	.000	.000	.000
W	0.55 ± 0.04	.000	.000	.000	1	-	.000	.000	.000	.148	.003	.000	.000	.000	.000	.000	.000
CL	0.76 ± 0.03	.000	.366	.003	.000	.000	-	.000	1	.000	.000	.003	.004	.03	1	.431	.936
CR	0.71 ± 0.04	.000	.03	.731	.000	.000	.000	-	.000	.000	.000	.775	.906	.311	.000	.013	.000
LR	0.75 ± 0.03	.000	.958	.082	.000	.000	1	.000	-	.000	.000	.074	.08	.361	1	.981	.383
CB	0.59 ± 0.05	.138	.000	.000	.036	.148	.000	.000	.000	-	1	.000	.000	.000	.000	.000	.000
CW	0.60 ± 0.04	.167	.000	.000	.000	.003	.000	.000	.000	1	-	.000	.000	.000	.000	.000	.000
LB	0.72 ± 0.03	.000	.85	1	.000	.000	.003	.775	.074	.000	.000	-	1	1	.017	.658	.000
LW	0.72 ± 0.03	.000	.824	1	.000	.000	.004	.906	.08	.000	.000	1	-	1	.021	.642	.000
RB	0.73 ± 0.03	.000	.999	1	.000	.000	.03	.311	.361	.000	.000	1	1	-	.132	.987	.000
RW	0.75 ± 0.03	.000	.774	.019	.000	.000	1	.000	1	.000	.000	.017	.021	.132	-	.844	.567
CLRB	0.74 ± 0.03	.000	1	.693	.000	.000	.431	.013	.981	.000	.000	.658	.642	.987	.844	-	.006
CLRW	0.77 ± 0.03	.000	.005	.000	.000	.000	.936	.000	.383	.000	.000	.000	.000	.000	.567	.006	-

Model	AUC	Adjusted p-values															
		C	L	R	D	W	CL	CR	LR	CD	CW	LD	LW	RD	RW	CLRD	CLRW
C	0.66 ± 0.03	-	.000	.000	.000	.000	.000	.000	.000	.083	.03	.000	.000	.000	.000	.000	.000
L	0.83 ± 0.03	.000	-	1	.000	.000	.849	.134	1	.000	.000	.019	.329	1	.001	1	.000
R	0.83 ± 0.03	.000	1	-	.000	.000	.593	.354	.993	.000	.000	.064	.592	1	.002	1	.000
B	0.59 ± 0.05	.000	.000	.000	-	.24	.000	.000	.000	.17	.375	.000	.000	.000	.000	.000	.000
W	0.55 ± 0.05	.000	.000	.000	.24	-	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
CL	0.85 ± 0.03	.000	.849	.593	.000	.000	-	.000	.015	.000	.000	.000	.006	.862	.972	.665	.462
CR	0.81 ± 0.02	.000	.134	.354	.000	.000	.000	-	.999	.000	.000	.999	1	.008	.000	.304	.000
LR	0.84 ± 0.03	.000	1	.993	.000	.000	.015	.999	-	.000	.000	.002	.069	1	.219	.997	.03
CB	0.62 ± 0.04	.083	.000	.000	.17	.000	.000	.000	.000	-	1	.000	.000	.000	.000	.000	.000
CW	0.62 ± 0.05	.03	.000	.000	.375	.000	.000	.000	.000	1	-	.000	.000	.000	.000	.000	.000
LB	0.80 ± 0.03	.000	.019	.064	.000	.000	.000	.999	.002	.000	.000	-	1	.001	.000	.052	.000
LW	0.81 ± 0.03	.000	.329	.592	.000	.000	.006	1	.069	.000	.000	1	-	.085	.000	.539	.000
RB	0.83 ± 0.02	.000	1	1	.000	.000	.862	.008	1	.000	.000	.001	.085	-	.000	1	.000
RW	0.86 ± 0.02	.000	.001	.002	.000	.000	.972	.000	.219	.000	.000	.000	.000	.000	-	.003	.956
CLRB	0.83 ± 0.02	.000	1	1	.000	.000	.665	.304	.997	.000	.000	.052	.539	1	.003	-	.000
CLRW	0.87 ± 0.03	.000	.000	.000	.000	.000	.462	.000	.03	.000	.000	.000	.000	.000	.956	.000	-

Abbreviations: ICU, Intensive care unit; AUC, Area under the curve; C, Clinical data; L, Laboratory results; R, Radiomics; B, Viral-Binary encoding; W, Viral-Word2Vec encoding

Kanonikal Korelasyon Analizi



Karar Destek Sistemleri

- **Erken sepsis tespiti**
 - Yetişkin 17,758 hastayı kapsayan prospektif, açık etiketli, çok merkezli çalışma
 - Supervised gradient-boosted decision trees
 - Hastane içi mortalite, yatış süresi ve 30 günlük yeniden başvuruda iyileşme
- **Antibiyotik reçetesi**
 - Beş hafta boyunca 515 piperasilin-tazobaktam reçetesi, denetimli bir makine öğrenimi algoritması, hastanenin mevcut uzman kurallar sistemi ve eczacılar tarafından prospektif olarak değerlendirilmiştir.
 - İyileştirme mevcut

Burdick H, Pino E, Gabel-Comeau D, et al. Effect of a sepsis prediction algorithm on patient mortality, length of stay and readmission: a prospective multicentre clinical outcomes evaluation of real-world patient data from US hospitals. *BMJ Health Care Inform*. 2020;27(1):e100109. doi:10.1136/bmjhci-2019-100109

Chang Y, Yao Y, Cui Z, et al. Changing antibiotic prescribing practices in outpatient primary care settings in China: Study protocol for a health information system-based cluster-randomised crossover controlled trial. *PLoS One*. 2022;17(1):e0259065. Published 2022 Jan 7. doi:10.1371/journal.pone.0259065

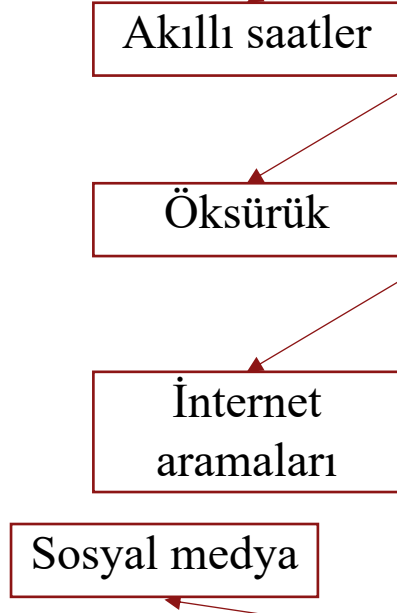
Sürveyans ve Halk Sağlığı





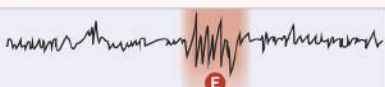

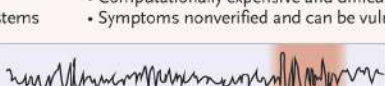
- Zamansal insidansın, hastalık risk faktörlerinin ve insanların mekansal hareketlerinin izlenmesi
 - Sosyal medya, GPS ve giyilebilir cihazlar
 - Salgınları belirlemek için çevrimiçi haber medyasının analizinde NLP modellerinin ve derlenmiş veri kaynaklarının kullanımı

- HealthMap



Sürveyans ve Halk Sağlığı



Individual event	Example of signal-generating method	Algorithm category	Signal of possible infectious disease in a population	Surveillance output
Biosignals passively measured by smartwatch	Gradient-boosting decision tree	Supervised classification	 A Change in biosignals	Early indication of possible outbreak
Method advantages <ul style="list-style-type: none">• Early warning can direct treatment and prevent spread• Continuously measured without requiring intervention		Method disadvantages <ul style="list-style-type: none">• Disease signal is nonspecific• Requires deployment of device before outbreak		
Cough detected by smart listening device	Regional proposal network	Artificial neural network	 B Cough begins	Spike in persons whose symptoms are detected early
Method advantages <ul style="list-style-type: none">• Passively monitor with already adopted devices• Can be used in homes or larger settings (e.g., waiting rooms)		Method disadvantages <ul style="list-style-type: none">• Requires advanced privacy protection schemes• Symptomatic person (i.e., who coughed) may be unknown		
Internet search query for viral testing site	Support vector regression	Supervised classification	 C Search query for testing	Hotspot of care-seeking behavior
Method advantages <ul style="list-style-type: none">• Can be inexpensive and centrally monitored• Captures behavior without requiring explicit participation		Method disadvantages <ul style="list-style-type: none">• Testing possibly unrelated to symptom status (e.g., for travel)• Searches may not lead to testing (e.g., resource constraints)		
Symptoms entered into website	Participatory surveillance	Human curated	 D Enters symptoms online	Real-time prevalence of possible cases
Method advantages <ul style="list-style-type: none">• Information can be disseminated without bureaucratic delay• Captures mild cases that may not formally test across settings		Method disadvantages <ul style="list-style-type: none">• Participants skew toward persons with high health literacy• Relies on syndromic definitions that may describe many causes		
Test result positive for virus	Traditional public health surveillance	Human curated	 E Positive test result returned	Official case counts
Method advantages <ul style="list-style-type: none">• Standard diagnostic accuracy• Mandatory reporting can capture rare and dangerous pathogens		Method disadvantages <ul style="list-style-type: none">• Verification can be slow and expensive• Requires resources that may not be available in certain settings		
Post on social media about diagnosis	Natural-language processing	Supervised classification	 F Post diagnosis on social media	Real-time prevalence of confirmed cases
Method advantages <ul style="list-style-type: none">• Rapid collection and dissemination of results• Wide array of users who may be missed by most other systems		Method disadvantages <ul style="list-style-type: none">• Computationally expensive and difficult to parse signal from noise• Symptoms nonverified and can be vulnerable to Internet trolls		
Mask wearing captured by CCTV	Convolutional neural network	Artificial neural network	 G Mask wearing starts	Nonpharmaceutical intervention levels
Method advantages <ul style="list-style-type: none">• Not vulnerable to desirability bias (i.e., captures true behavior)• High level of geographic specificity		Method disadvantages <ul style="list-style-type: none">• Highly invasive and susceptible to privacy abuse• Resource intensive, especially outside urban locales		

Sürveyans ve Halk Sağlığı

- CDC, influenza suşlarının tahmin etmek için bölgesel ve mevsimsel epidemiyolojik verilere dayalı ML modellerini kullanıyor
- SARS-CoV-2'nin hızlı ve ölçeklenebilir genomik sınıflandırması ve soy haritalaması

Reich NG, McGowan CJ, Yamana TK, et al. Accuracy of real-time multi-model ensemble forecasts for seasonal influenza in the U.S. *PLoS Comput Biol.* 2019;15(11):e1007486. Published 2019 Nov 22. doi:10.1371/journal.pcbi.1007486

Randhawa GS, Soltysiak MPM, El Roz H, de Souza CPE, Hill KA, Kari L. Machine learning using intrinsic genomic signatures for rapid classification of novel pathogens: COVID-19 case study. *PLoS One.* 2020 Apr 24;15(4):e0232391. doi: 10.1371/journal.pone.0232391. Erratum in: *PLoS One.* 2021 Jan 27;16(1):e0246465.

İlaç Keşfi

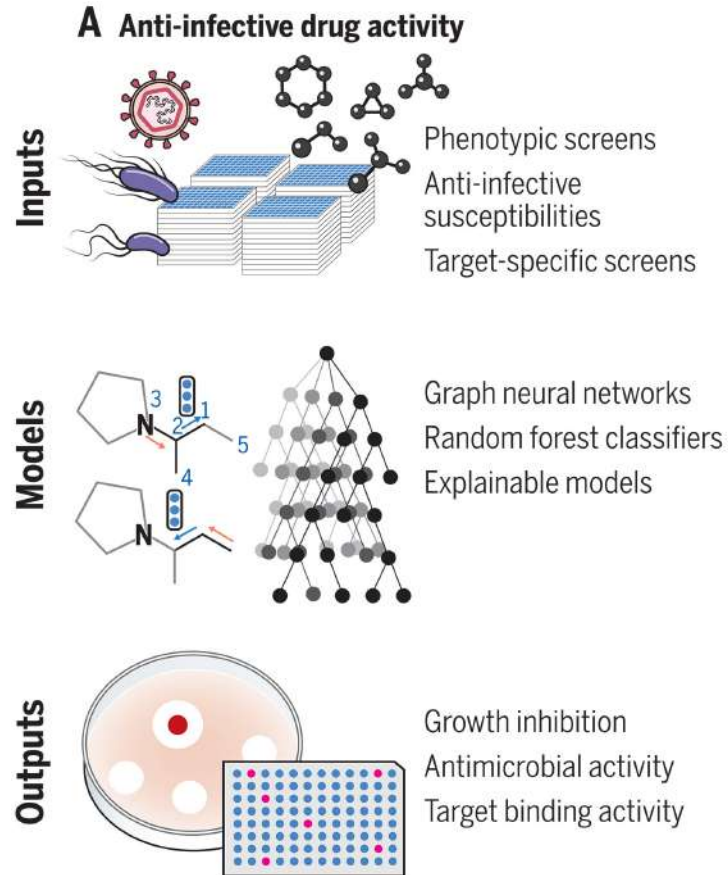
- **Halicin: In-silico derin öğrenme çalışması**
 - Drug Repurposing Hub Veritabanı'ndaki 6000'den fazla aday bileşik taranıyor
 - Hücre içerisinde demir birikimi üzerinden etkili
 - *C. difficile*, *A. baumannii*, *M. tuberculosis*
- **AlphaFold**
 - Amino asit sekans bilgisi üzerinden üç boyutlu protein yapısı tahmini

Stokes JM, Yang K, Swanson K, et al. A Deep Learning Approach to Antibiotic Discovery [published correction appears in *Cell*. 2020 Apr 16;181(2):475-483]. *Cell*. 2020;180(4):688-702.e13. doi:10.1016/j.cell.2020.01.021

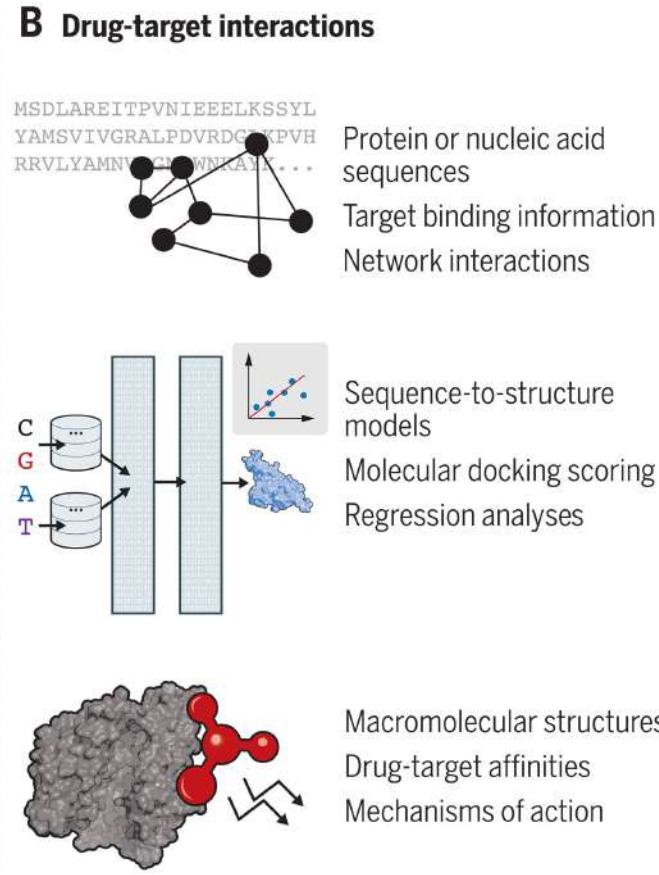
Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-589. doi:10.1038/s41586-021-03819-2

Bess A, Berglind F, Mukhopadhyay S, et al. Artificial intelligence for the discovery of novel antimicrobial agents for emerging infectious diseases. *Drug Discov Today*. 2022;27(4):1099-1107. doi:10.1016/j.drudis.2021.10.022

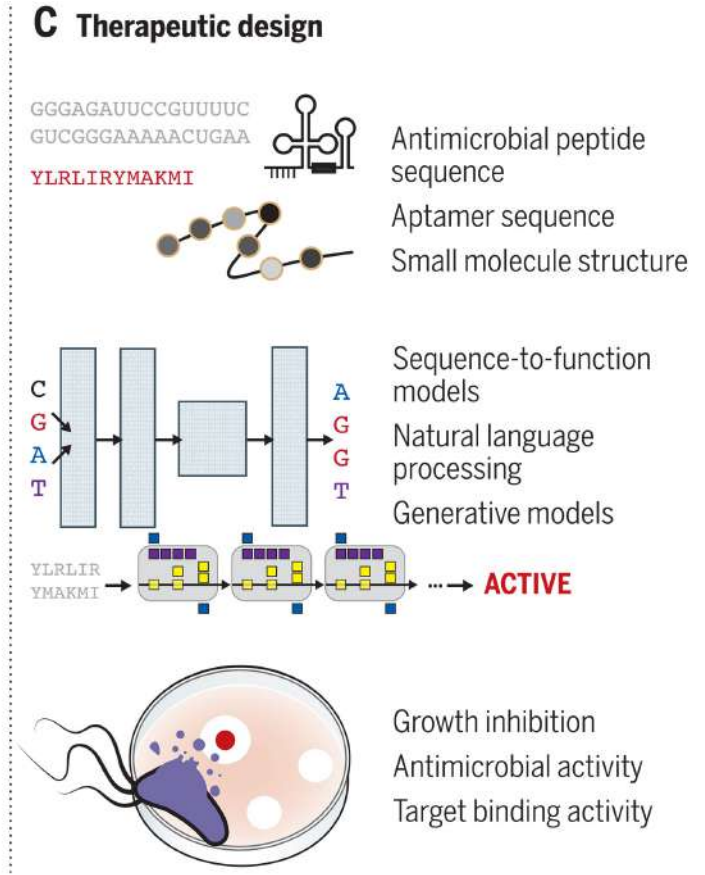
İlaç Keşfi



Halicin

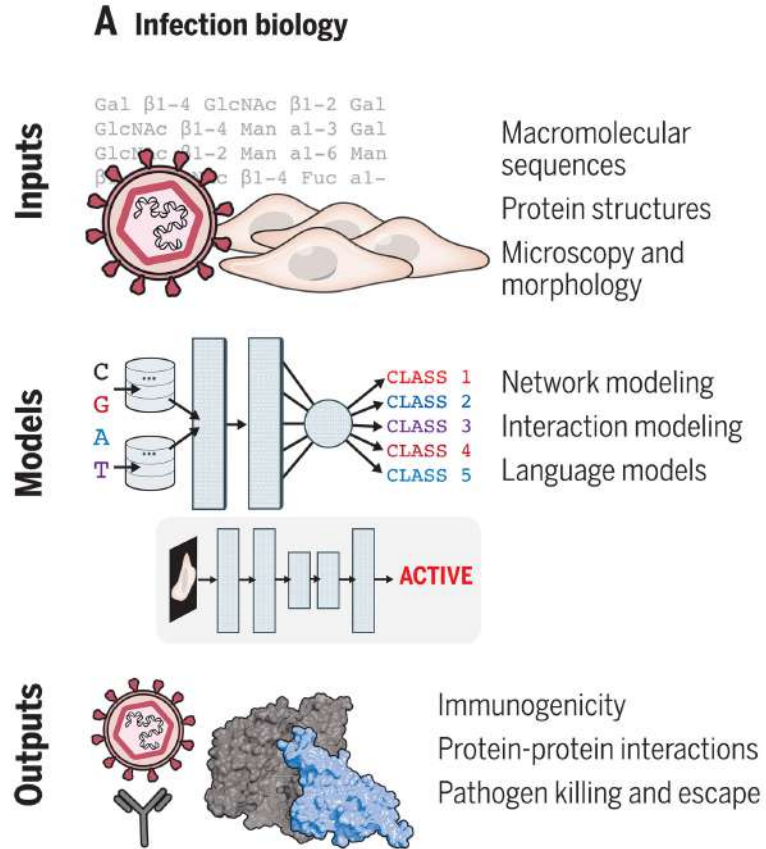


AlphaFold

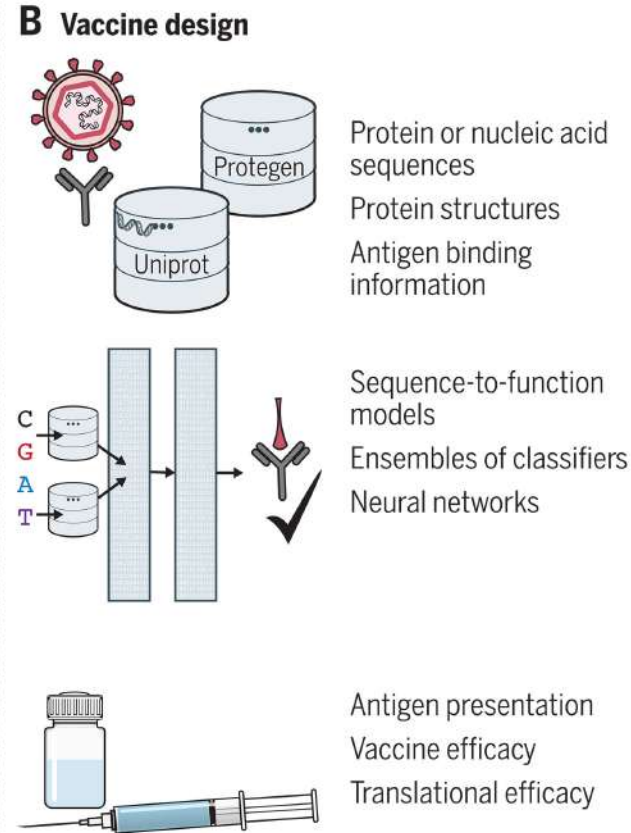


Peptid sekans tahmini

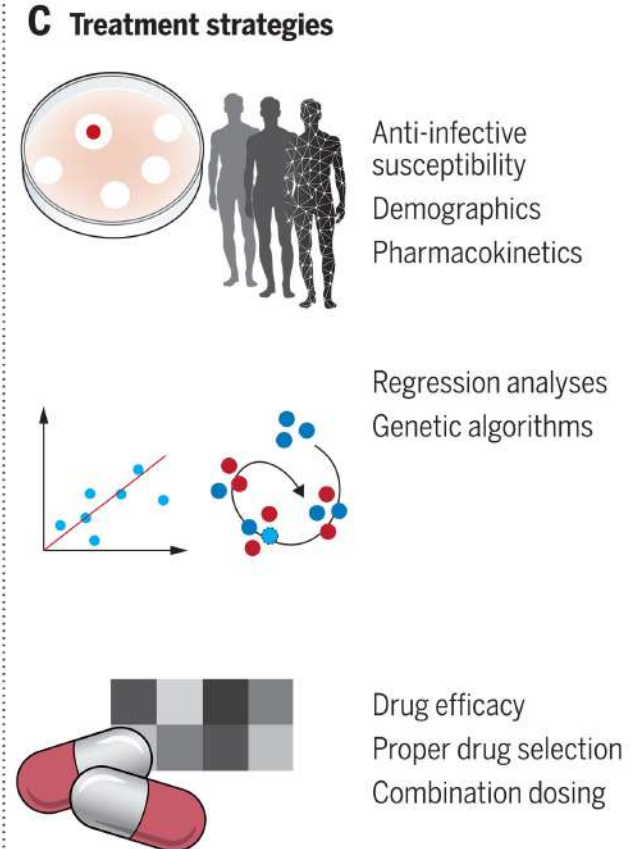
İlaç Keşfi



Moleküler yollar



Aşı etkinliği



Keskin tıp

Kliniđe Uyarlamada Zorluklar

- **Retrospektif vs. prospektif alıřmalar**
 - Byk miktarda veri ihtiyacı, retrospektif veriler
- **Klinikte RCT'ler altın standart**
 - İlgin olarak ML camiasında pre-print yazılar yaygın
- **Klinik veriler ođu zaman eksik**
 - Veri bilimciler tarafından yazılan alıřmalar
 - Demografik bilgiler bile ođu zaman eksik

Kliniğe Uyarlamada Zorluklar

- Klinik kullanılabilirlik açısından kullanılan metrikler çok uygun değil
 - Duyarlılık, özgüllük, AUROC
 - F1 score, cosine similarity, etc.
- Farklı algoritmaları karşılaştırmak zor
 - Farklı çalışmalarda, çeşitli popülasyonlarda farklı metodolojiler kullanılmakta
 - Adil karşılaştırmalar yapabilmek için, algoritmaların aynı performans ölçütleri kullanılarak aynı bağımsız test grupları üzerinde karşılaştırmaya tutulması gerekli

Kliniğe Uyarlamada Zorluklar

- **Overfitting veya underfitting**
 - Model ezberliyor mu? Genelleyebiliyor mu?
 - Eğitim veri seti yeterli büyüklükte mi?
- **Açıklanabilir olmayan modeller**
 - Klinisyenler nasıl çalıştığını anlayamadığı modellere nasıl güvensinler?
- **Veri setlerindeki değişiklikler**
 - Eğitim ve test verileri farklı alt popülasyonlardan oluşabilir, bunlar uyumlu olmalı
 - Mesela antibiyotik direnç oranları
 - Veri bilimciler alan bilgisi eksikliği nedeniyle bu durumu göz ardı etme eğiliminde

Kliniğe Uyarlamada Zorluklar

- **Karıştırıcı faktörler**
 - Cetveller, cerrahi işaretleyiciler
 - Modeller hile yapıyor!
- **Yeni popülasyonlara genellenebilme kabiliyeti**
 - Validasyon kohortlarında eksiklik
 - Lokal kohortlara kısıtlılık
 - Mesela Fitzpatrick skorlarında farklılık



Kliniğe Uyarlamada Zorluklar

- **Yanlış algoritmaların kullanımı**
 - Bazı modeller gürültülü veride başarısız
 - AdaBoost
- **AI sistemlerinin entegrasyonunda lojistik yetersizlikler**
 - Farklı veri setlerini birleştirmek zor
 - Teknik olarak farklı modaliteleri birleştirmek de zor
- **Kalite kontrol nasıl olmalı?**
 - Model başarısız olursa kim sorumlu? Klinisyen mi, veri bilimci mi?

Hipotez-merkezli araştırma

THE LANCET



Veri-merkezli araştırma

IEEE Xplore®



THE LANCET
Digital Health



NEJM
AI

Teşekkürler..