



دانشگاه علم و صنعت ایران

کارگاه آشنایی با هوش مصنوعی و کاربردهای آن در آنکولوژی

ارایه دهنده: سیدموسی حسینی
دکتری مهندسی برق، دانشگاه علم و صنعت ایران

فهرست مطالب

◦ مقدمه‌ای بر هوش مصنوعی

◦ کاربردهای هوش مصنوعی در آنکولوژی

◦ جمع‌بندی

مقدمه‌ای بر هوش مصنوعی

○ هوش مصنوعی چیست؟

○ تاریخچه هوش مصنوعی

○ اهداف

○ روشها

○ کاربردها

هوش مصنوعی چیست؟

○ هوش مصنوعی (ai)

حوزه ای از علوم ریاضی، کامپیوتر و مهندسی است با هدف طراحی و ساخت ماشینهای با هوش انسانی و یا فراتر.

○ کاربردهای وسیعی در زندگی امروز

○ موتورهای جستجوی اینترنتی

○ سامانه پیشنهاد محتوا

○ درک گفتار انسانی

○ ابزارهای مولد و خلاق



تاریخچه کوتاه

○ اصطلاح هوش مصنوعی توسط جان مک کارتی در کارگاهی که تابستان ۱۹۵۶ در دارتموث برگزار شد پیشنهاد گردید

○ آلن تورینگ: مبانی ریاضی هوش ماشینی

○ شکوفایی سریع:

○ Arthur Samuel's "Checkers playing program"

○ شبکه های عصبی مصنوعی

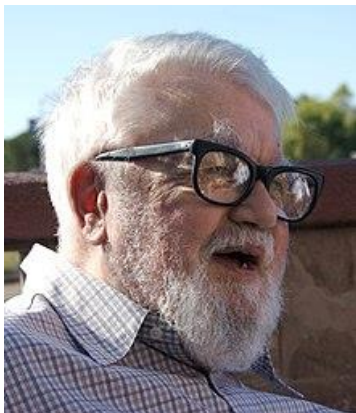
○ توسعه رباتیک

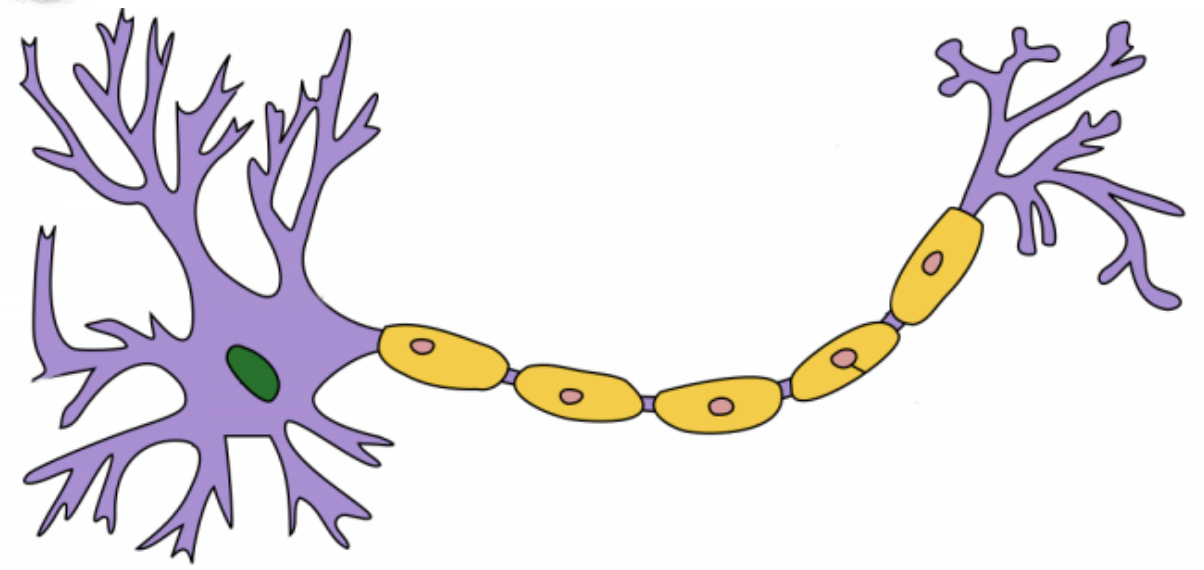
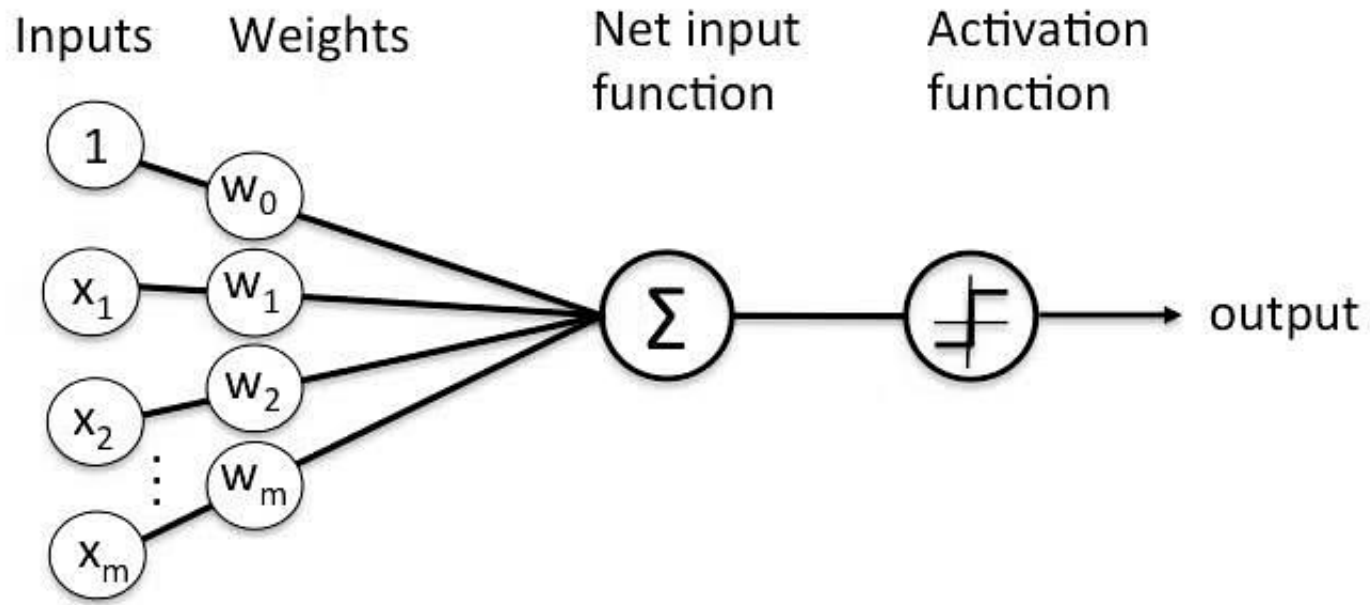
○ شروع افول

○ مسایل پیچیده تر قابل حل نبود

○ کاهش حمایت مالی

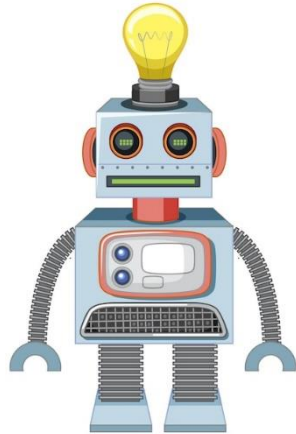
○ زمستان هوش مصنوعی (۱۹۸۷-۲۰۰۰)





زمستان هوش مصنوعی

the spirit is willing but the flesh is weak



the vodka is good but the meat is rotten

- ۱۹۶۶: شکست ترجمه ماشینی
- ۱۹۷۱: انتقادات به مدل اولیه پرسپترون
- ۱۹۷۳: دارپا تحقیقات هوش مصنوعی را معلق میکند
- ۱۹۸۸: کاهش سرمایه گذاری دولت امریکا در هوش مصنوعی
- ۱۹۹۰: تعلیق پروژه های expert systems



○ دوران بازسازی

- مسایل دنیای واقعی
- دانش زیاد
- روشهای اکتشافی حل مساله

○ مطالعه عمیقتر عملکردهای شناختی

○ تقسیم هوش مصنوعی به حوزه های مختلف

○ توسعه expert systems

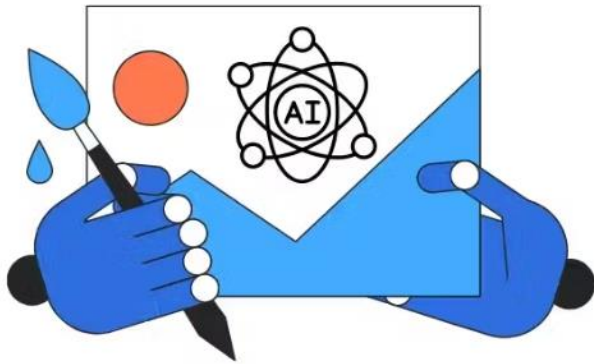
○ بهره مندی از منابع اینترنت و پردازنده های جدید

○ توسعه AI مبتنی بر داده (دسترسی به داده های بزرگ)

شکوفایی (از ۲۰۱۲ تا کنون)

◦ پروژه ImageNet و برتری deep learning

◦ افزایش سرمایه گذاری توسط شرکتهای بزرگ نظیر گوگل، مایکروسافت ...

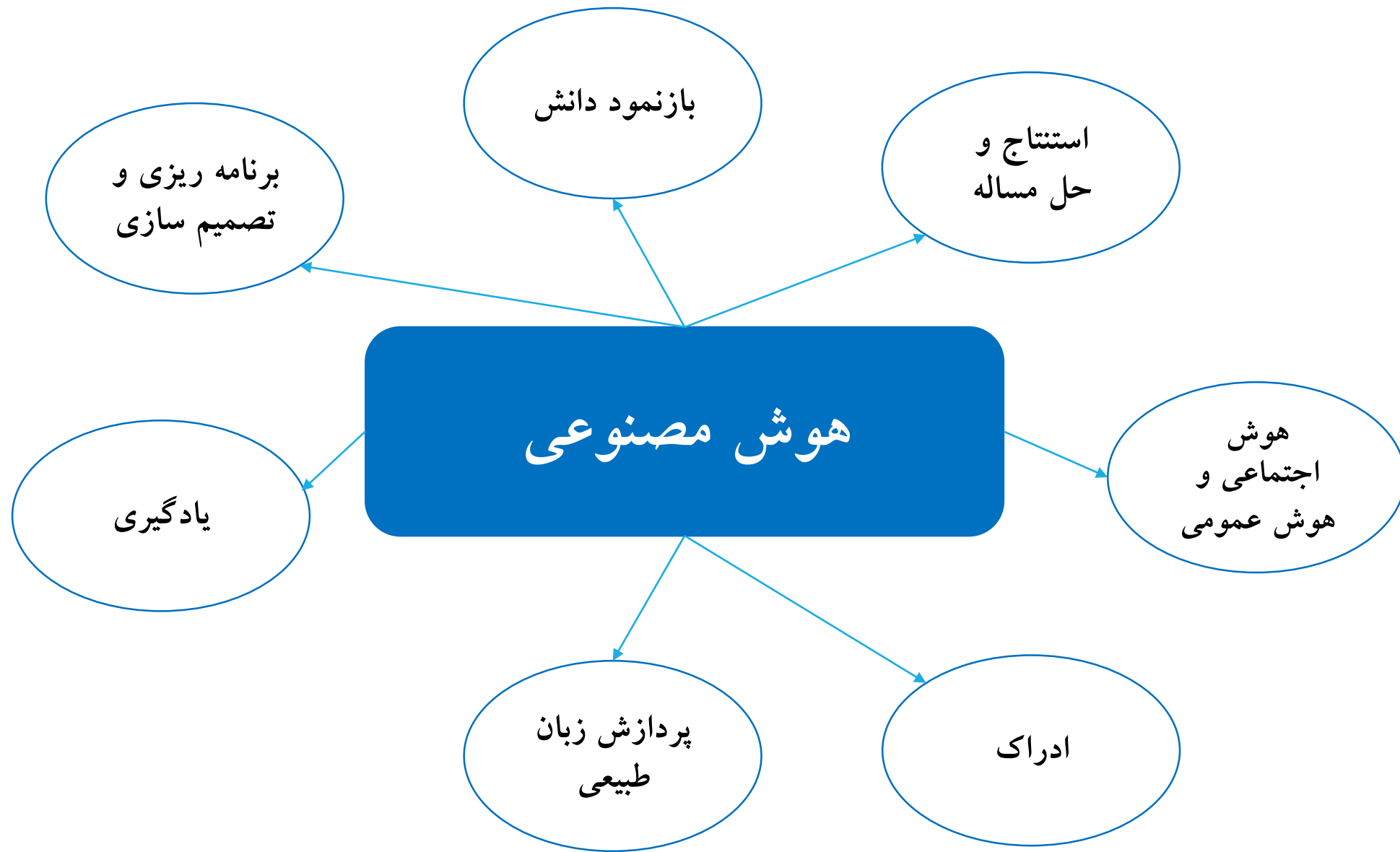


◦ ابداع شبکه های generative

◦ جذب بیشترین منابع مالی دولتی

◦ بیشترین تعداد مقالات و پژوهشها

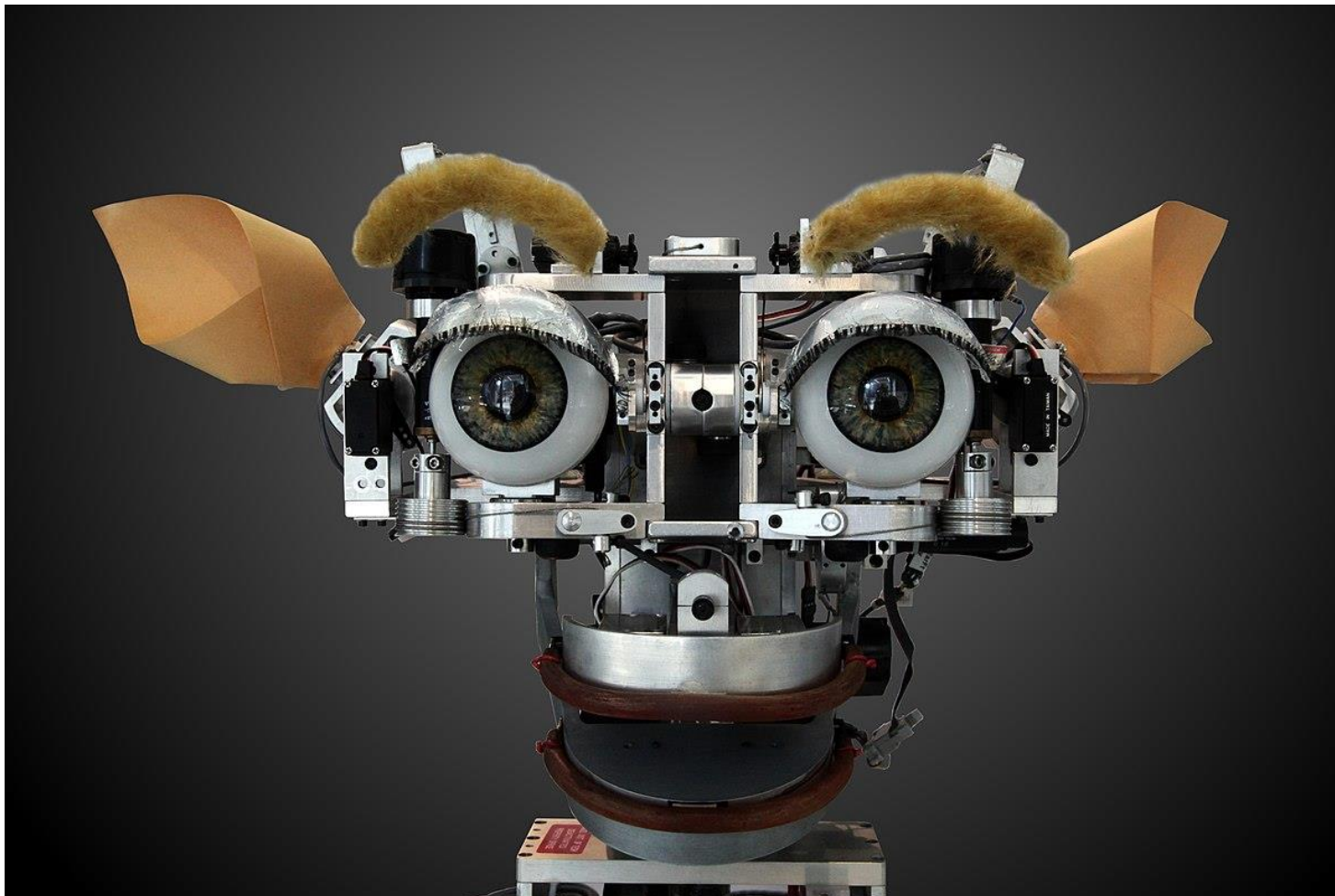
◦ ابزار رقابت ابرقدرتهای جهانی



اهداف اصلی هوش مصنوعی

- **استنتاج و حل مساله:** عموماً راه حل‌های مرحله به مرحله
- **بازنمود دانش (knowledge representation)** توانایی پاسخ دادن هوشمند به سوالات در مورد حقایق دنیای واقعی. در دستیارهای هوشمند، ابزار تشخیصی کاربرد دارد.
- **برنامه‌ریزی و تصمیم‌سازی:** برنامه‌ریزی برای نیل به یک هدف خاص، تصمیم‌سازی هوشمند براساس موقعیت و شرایط. ترجیح دادن برخی تصمیم‌ها و اجتناب از برخی دیگر. ماشین هوشمند باید آثار هرگونه عمل خود را بداند. حوزه ریاضی تئوری بازی‌ها.
- **یادگیری:** ماشینها میتوانند عملکرد خود را در انجام یک وظیفه به صورت خودکار بهبود دهند.
 - یادگیری نظارت شده
 - یادگیری بدون نظارت
 - یادگیری تقویتی
 - یادگیری انتقالی

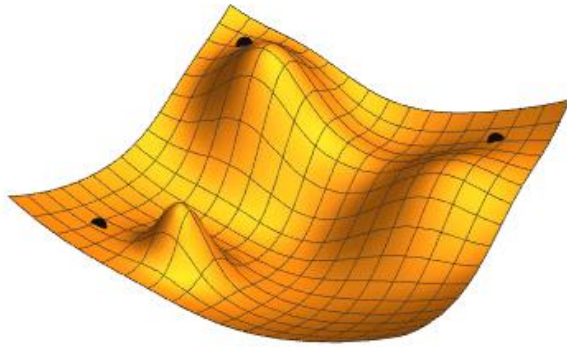
- پردازش زبان طبیعی: توانایی خواندن نوشتن و ارتباط از طریق زبانهای انسانی
- ادراک ماشین: توانایی استفاده از حسگرهای ورودی (دوربین، میکروفون ...) در درک دنیای پیرامونی.
 - تشخیص گفتار
 - تشخیص چهره
 - تشخیص اشیا
- هوش اجتماعی: توانایی تشخیص، تفسیر، پردازش و شبیه‌سازی احساسات انسانی
- هوش جامع: توانایی جامع در حل هوشمندانه طیف وسیعی از مسایل مشابه هوش انسانی



Kismet

در دهه ۱۹۹۰ در دانشگاه MIT
رباتی برای شناسایی و شبیه‌سازی عواطف
انسانی

روش‌ها



○ جستجو و بهینه سازی

○ منطق

○ منطق گزاره ای

○ منطق فازی

○ روشهای آماری

○ استنتاج بیزی

○ مدل های مارکوف

○ تئوری بازی ها

○ تئوری اطلاعات

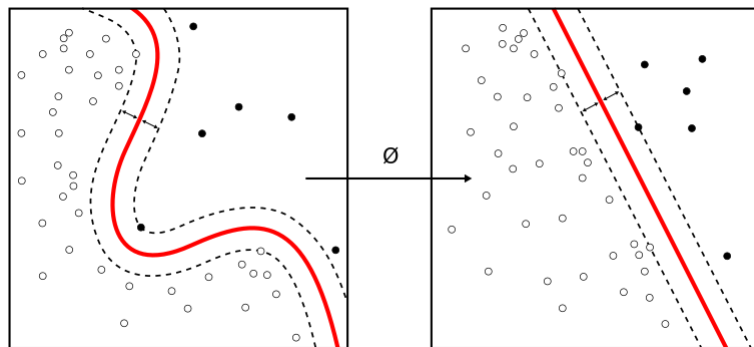
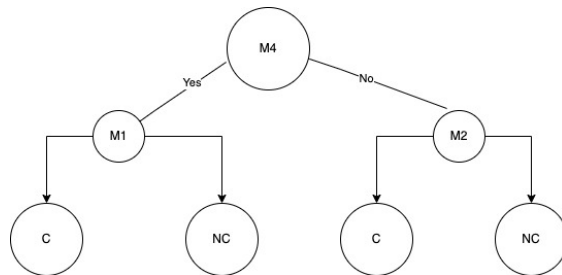
○ طبقه بندها و روشهای آماری یادگیری

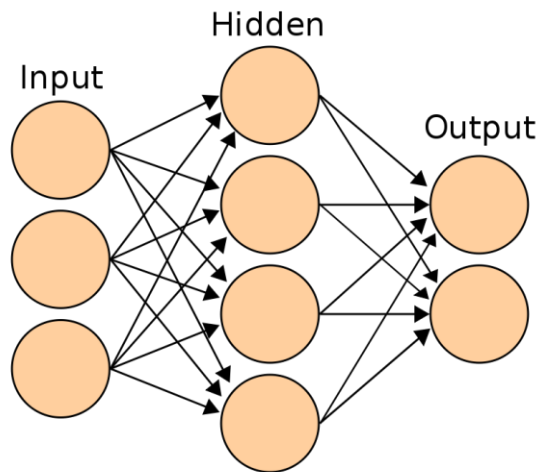
○ شناسایی الگو

○ درخت تصمیم گیری

○ روش های مبتنی بر کرنل

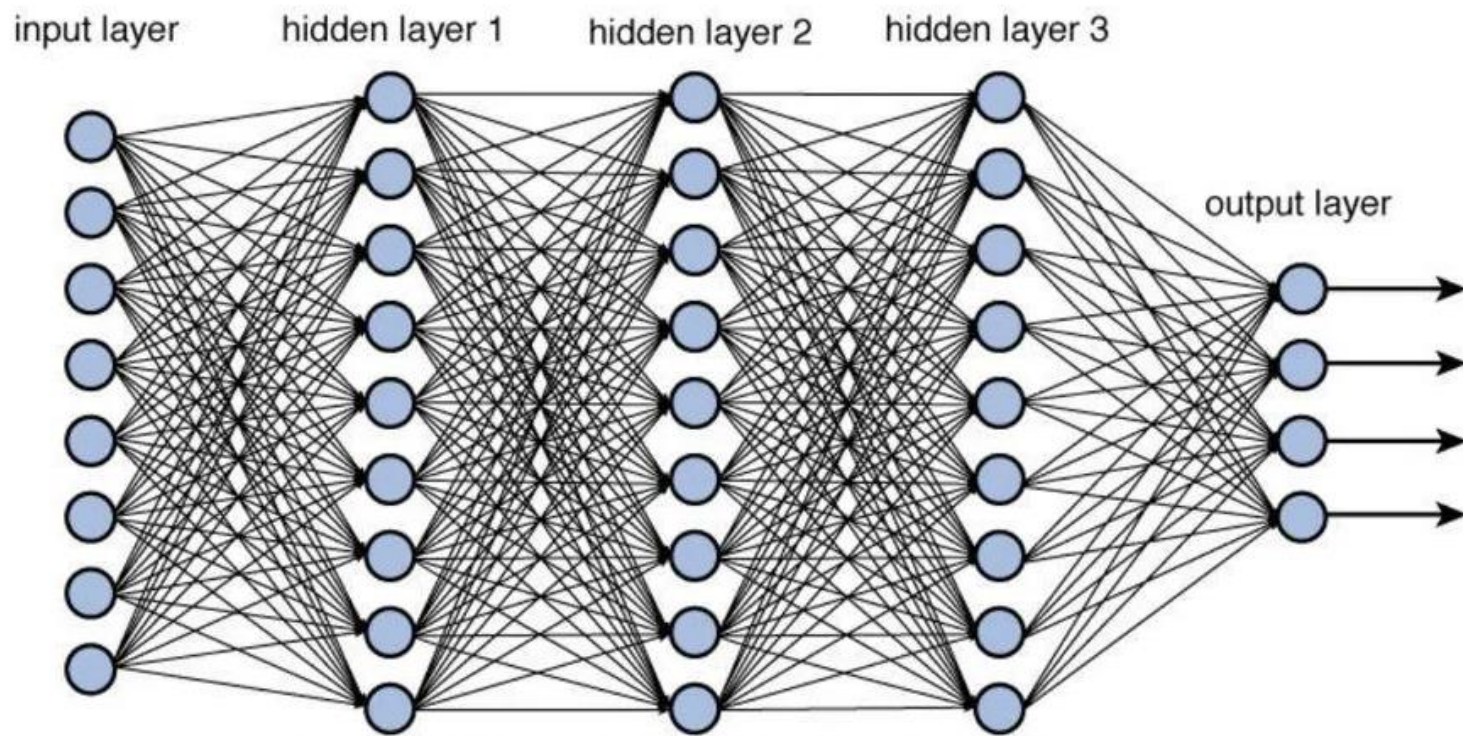
$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}$$





○ شبکه های عصبی مصنوعی

- مدل ساده ای از مغز
- محدود به مسایل ساده

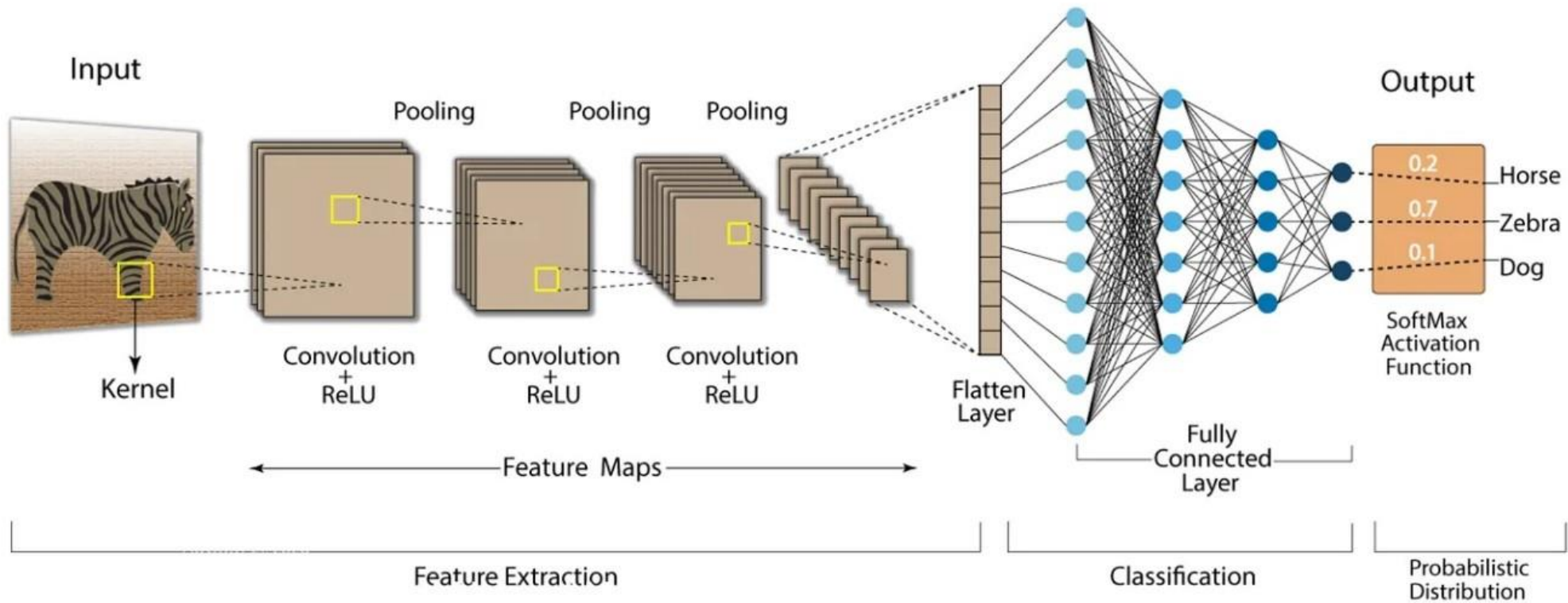


○ شبکه های عمیق

- سه لایه پنهان و بیشتر
- حل مسایل پیچیده

شبکه کانولوشنی CNN

Convolution Neural Network (CNN)



GPT

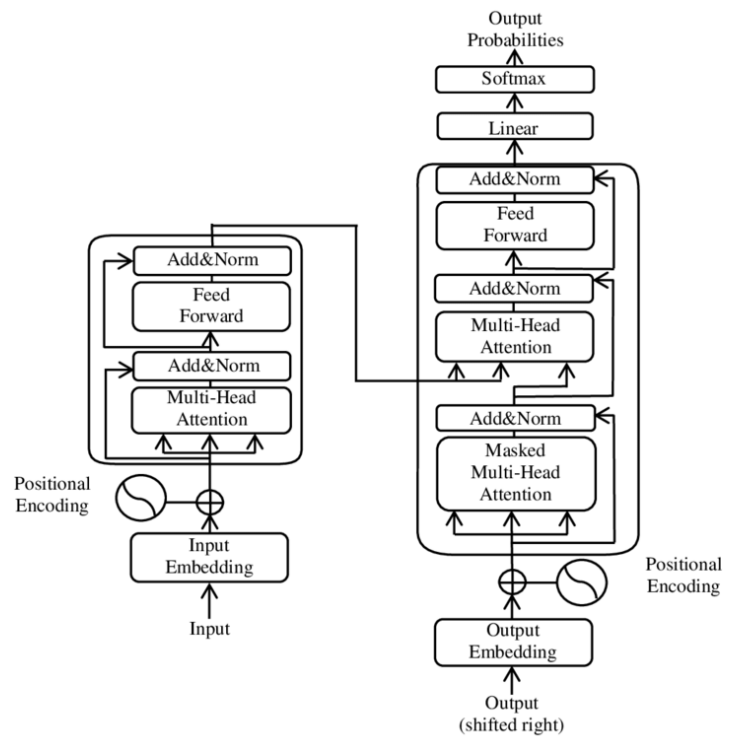
Generative pre-trained transformers

توانایی مکالمه با انسان و تولید متن و محتوا

سخت افزار و نرم افزارهای تخصصی

واحدهای پردازش گرافیکی مختص هوش مصنوعی (GPU)

زبانهای برنامه نویسی نظیر Python, Lisp, Prolog



کاربردهای هوش مصنوعی



Google Ads

○ حوزه اینترنت

○ موتورهای جستجوگر

○ ترجمه آنلاین

○ سامانه های اشتراک و پیشنهاد محتوا

○ تبلیغات آنلاین

○ دستیارهای هوشمند



Hey Siri





کاربردهای هوش مصنوعی

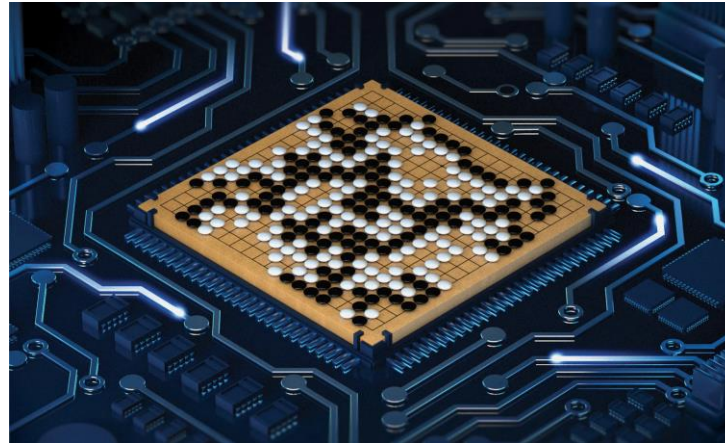
○ حوزه بازی

○ آزمون پیشرفته ترین AI در حوزه بازی

○ DEEP BLUE

○ AlphaGo

○ Pluribus





کاربردهای هوش مصنوعی

◦ هوش مصنوعی مولد

◦ Generative adversarial network

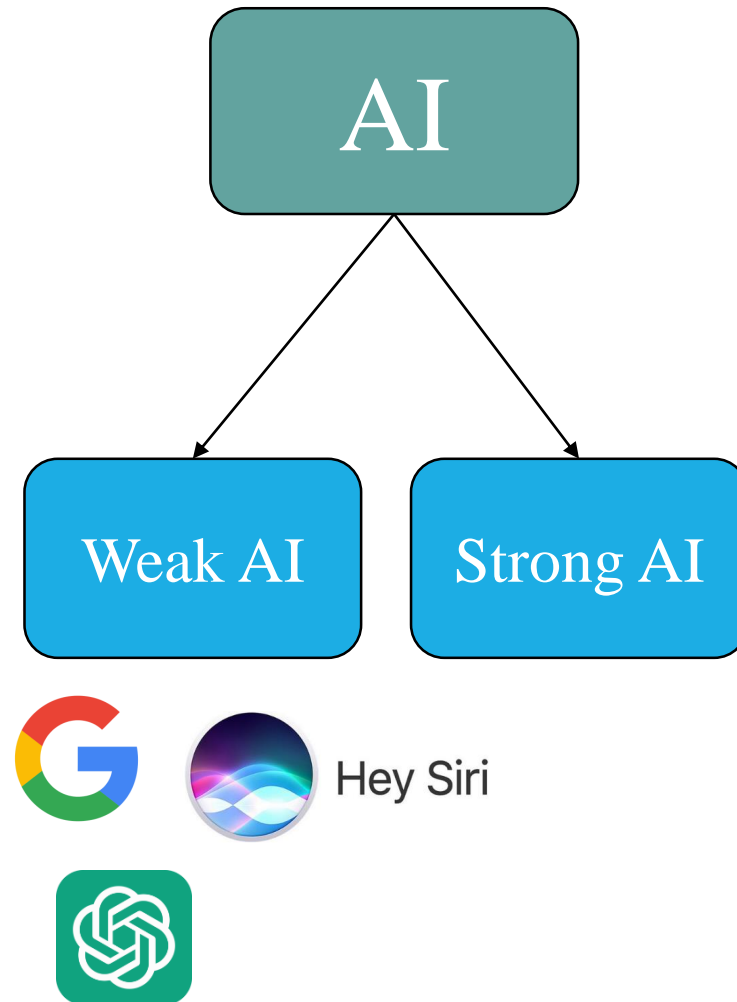
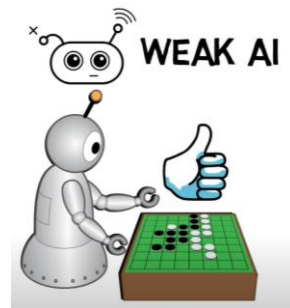
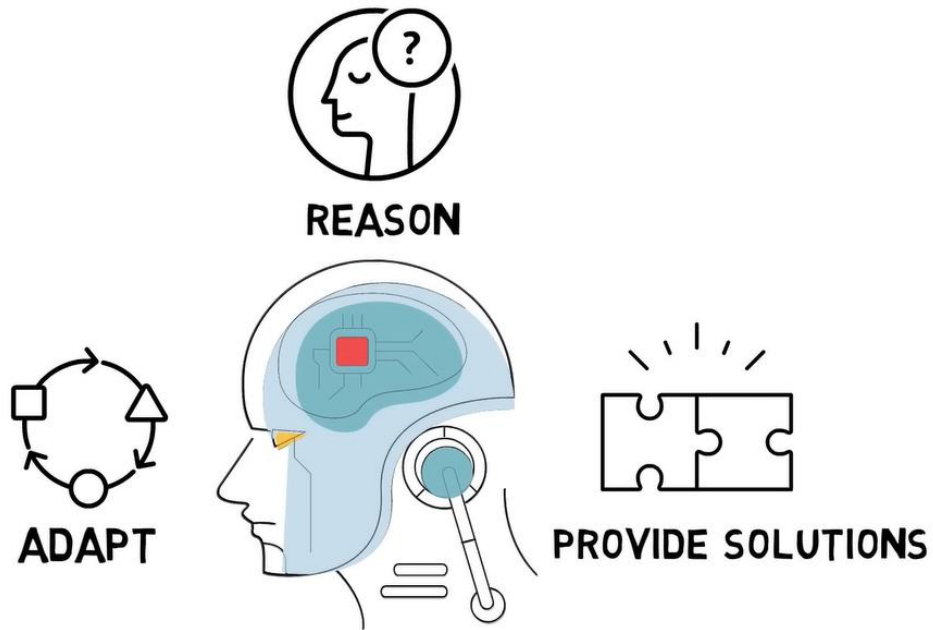
◦ Ian Goodfellow 2014

◦ آموزش بر اساس متن، تصاویر، صداها

◦ تولید محتوای جدید

◦ تشخیص و تولید سبک

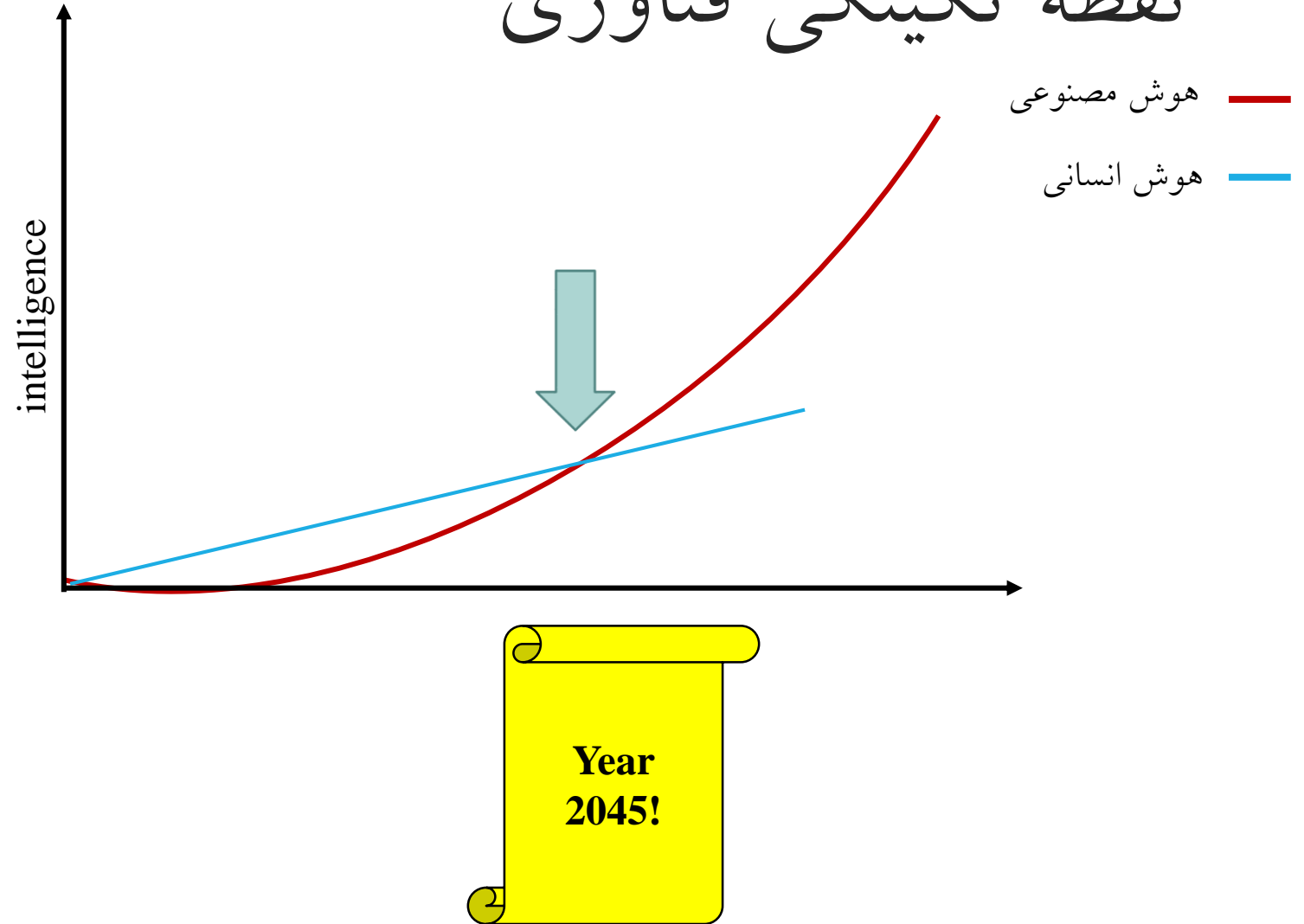
انواع هوش مصنوعی



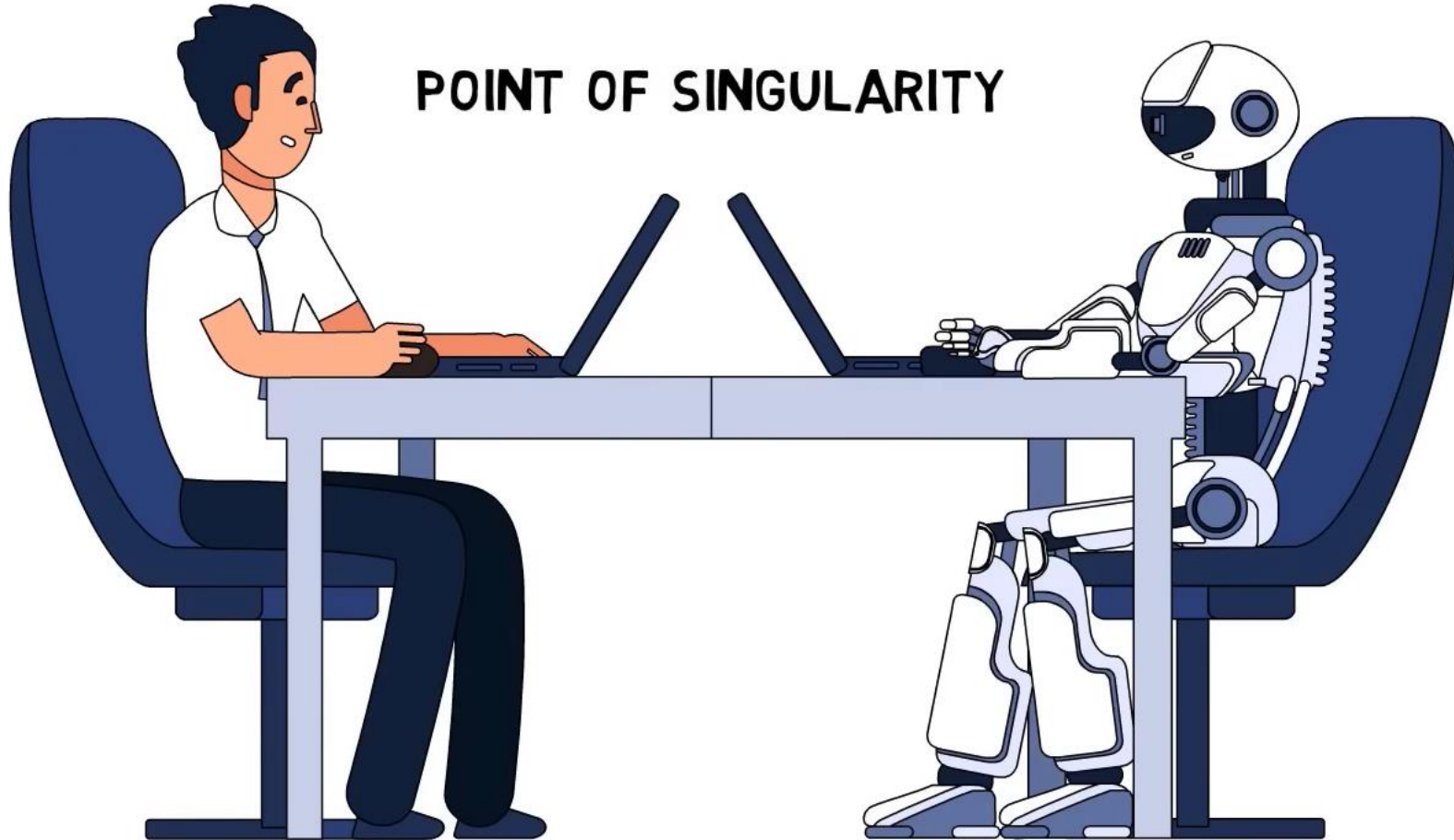


Ray Kurzweil

نقطه تکینگی فناوری



POINT OF SINGULARITY





Artificial Intelligence:

Mimicking the intelligence or behavioural pattern of humans or any other living entity.

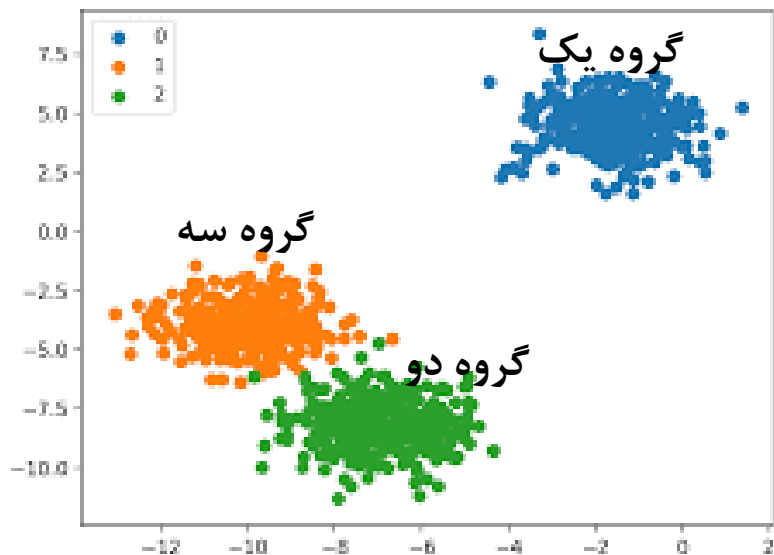
Machine Learning:

A technique by which a computer can "learn" from data, without using a complex set of different rules. This approach is mainly based on training a model from datasets.

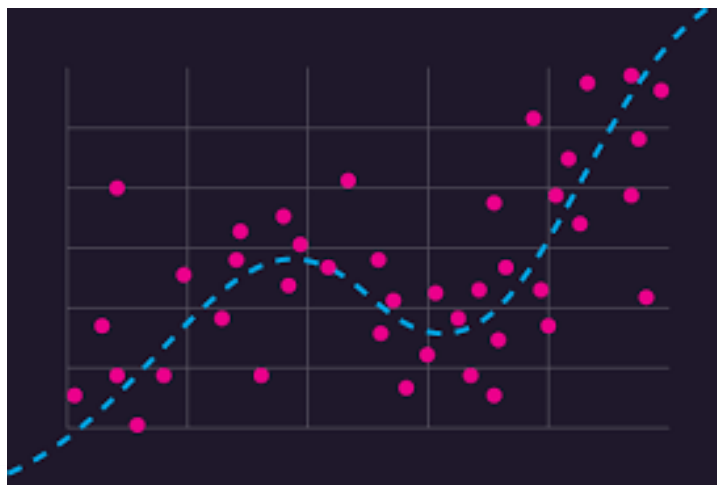
Deep Learning:

A technique to perform machine learning inspired by our brain's own network of neurons.

انواع مسایل



- کلاسبندی
- خروجی گسسته
- تشخیص درست طبقه بندی (دو یا چند کلاسه)
- بیمار بودن یا نبودن (نوع subtype یک بیماری)



- رگرسیون
- تخمین خروجی پیوسته
- طول عمر
- غلظت دارو در خون

هوش مصنوعی در آنکولوژی

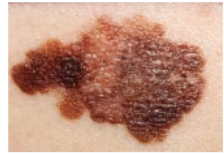
هوش مصنوعی در سرطان پوست

طبقه بندی سرطان پوست با استفاده از شبکه های یادگیری عمیق

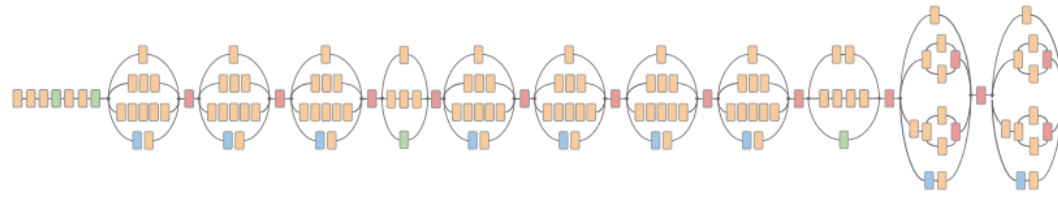
- سرطان پوست، شایع ترین در جوامع انسانی (یک نفر از هر پنج نفر در امریکا در طول حیات)
- تشخیص: بررسی بالینی، درموسکوپی، بیوپسی و هیستوپاتولوژی
- مشابهت ظاهر ضایعات پوستی تشخیص ماشینی و خودکار را دشوار میکند
- شبکه CNN برای تشخیص بدخیمی (Google's Inception v3 CNN)
- تعداد ۱۲۹۴۵۰ تصویر
- مساله اول: شناسایی رایج ترین سرطان، جدا کردن keratinocyte carcinomas از benign seborrheic keratosis
- مساله دوم: شناسایی مرگبارترین سرطان، جدا کردن malignant melanomas از benign nevi
- نتیجه: شبکه عصبی به دقتی در حد متخصص درماتولوژی یا بهتر میرسد
- ابزار تشخیصی کارآمد قابل استفاده بدون حضور متخصص

48 layers, 129,450 clinical images, 2,032 different diseases

Skin lesion image



Deep convolutional neural network (Inception v3)



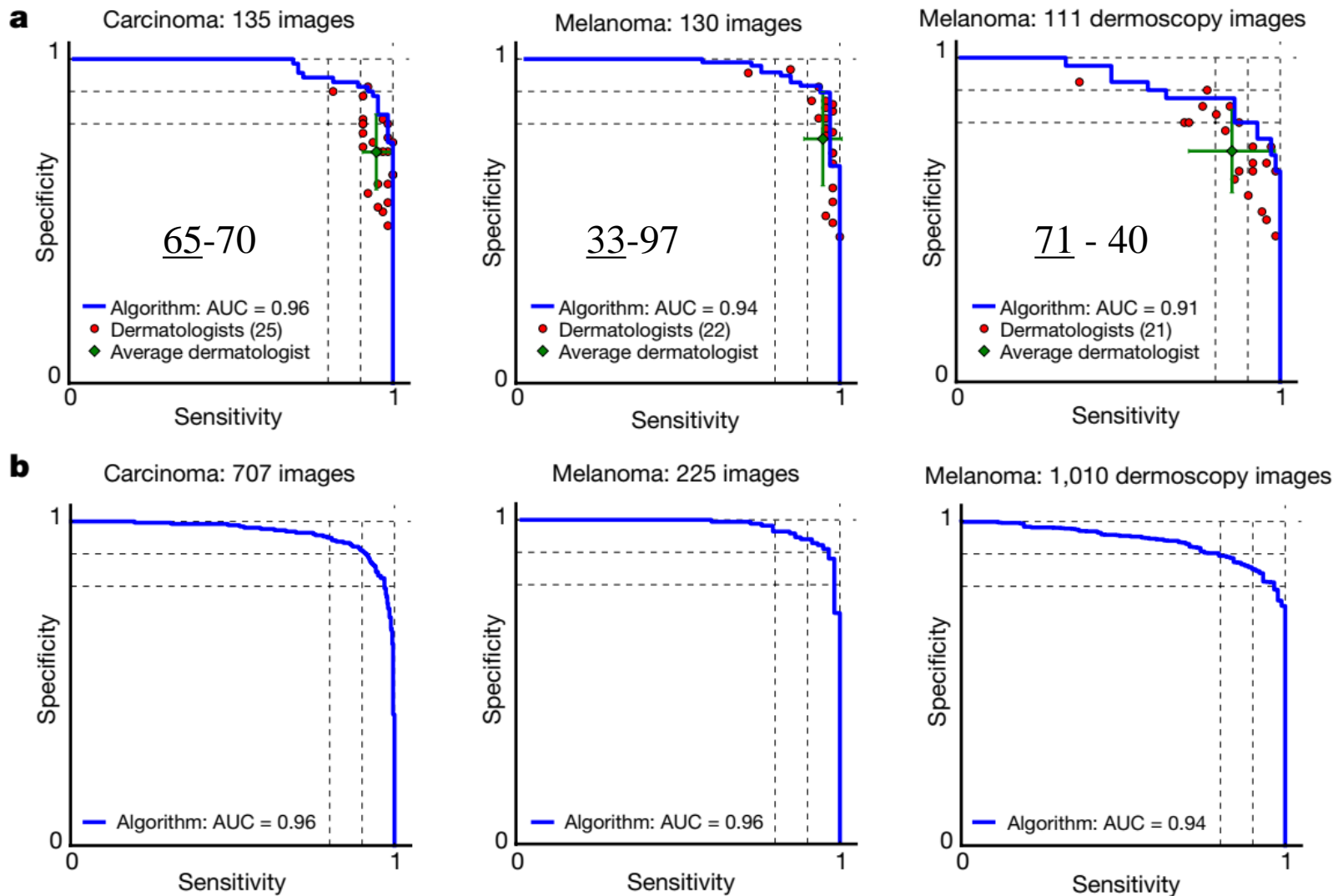
- Convolution
- AvgPool
- MaxPool
- Concat
- Dropout
- Fully connected
- Softmax

Training classes (757)

- Acral-lentiginous melanoma
- Amelanotic melanoma
- Lentigo melanoma
- ...
- Blue nevus
- Halo nevus
- Mongolian spot
- ...
-
-
-

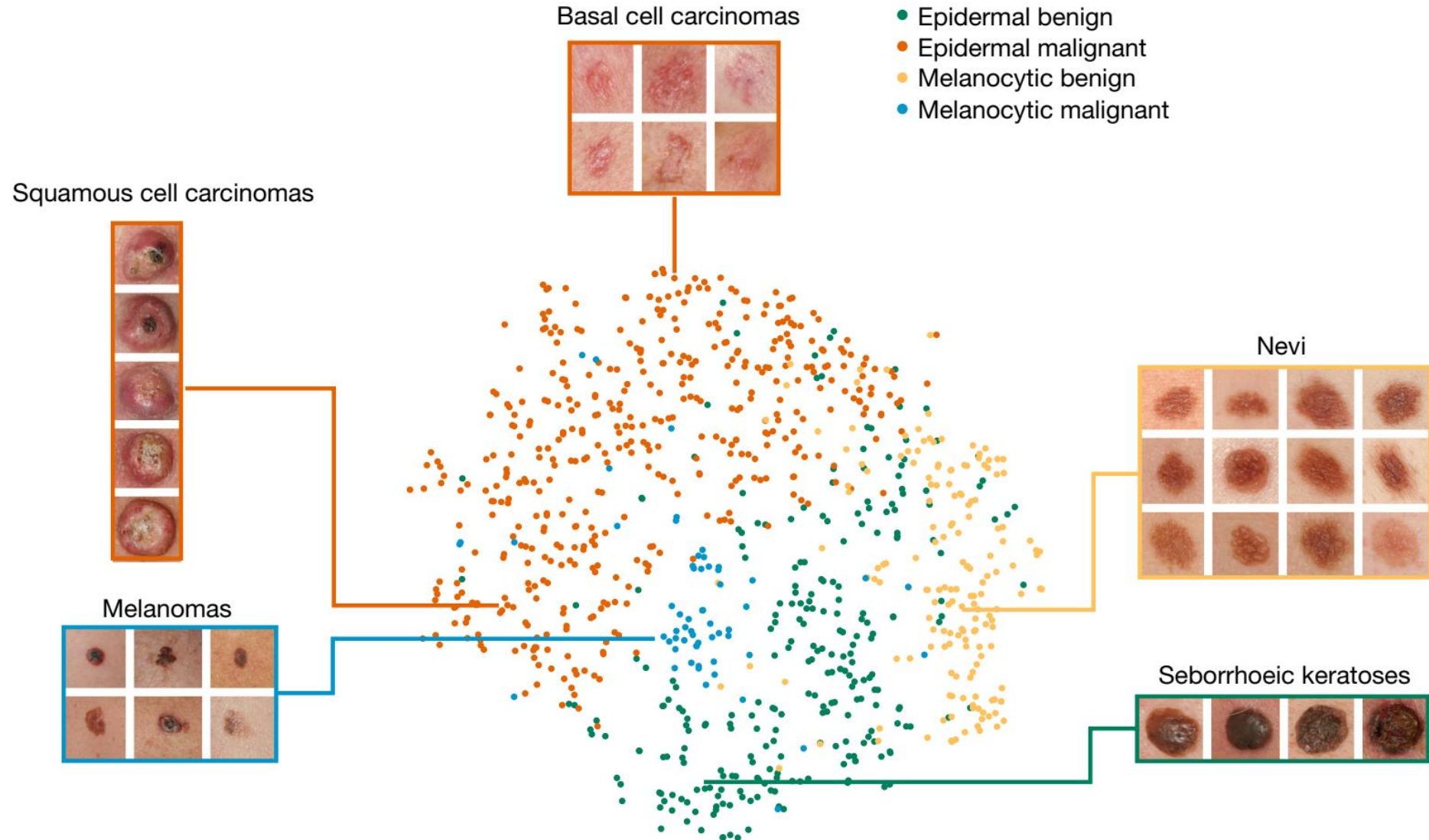
Inference classes (varies by task)

- 92% malignant melanocytic lesion
- 8% benign melanocytic lesion



Larger datasets

فضای ویژگی ها



Extended Data Table 2 | General validation results

a. Classifier Three-way accuracy

Classifier	Three-way accuracy
Dermatologist 1	65.6%
Dermatologist 2	66.0%
CNN	69.4 ± 0.8%
CNN - PA	72.1 ± 0.9%

c. Disease classes: three-way classification

0. Benign single lesions
1. Malignant single lesions
2. Non-neoplastic lesions

b. Classifier Nine-way accuracy

Classifier	Nine-way accuracy
Dermatologist 1	53.3%
Dermatologist 2	55.0%
CNN	48.9 ± 1.9%
CNN - PA	55.4 ± 1.7%

d. Disease classes: nine-way classification

0. Cutaneous lymphoma and lymphoid infiltrates
1. Benign dermal tumors, cysts, sinuses
2. Malignant dermal tumor
3. Benign epidermal tumors, hamartomas, milia, and growths
4. Malignant and premalignant epidermal tumors
5. Genodermatoses and supernumerary growths
6. Inflammatory conditions
7. Benign melanocytic lesions
8. Malignant Melanoma

تشخیص سرطان پوست با استفاده از اپلیکیشن و دوربین موبایل

- توسعه مدل CNN با نام DenseNet169
- طبقه بندی هفت نوع ضایعه پوستی
- آموزش دیده بر اساس دیتاست HAM10000
- نتایج $ACC = 92\%$, $sensitivity = 93\%$
- مدل سبک تر برای اندروید طراحی شد (دو کلاس: خوشخیم، بدخیم)

هفت کلاس

Melanoma (MEL), Basal Cell Carcinoma (BCC), Actinic Keratosis and Intraepithelial Carcinoma (AKIEC), Melanocytic nevus (NV), Benign Keratosis Lesions (BKL), Dermatofibroma (DF), Vascular lesion (VASC),

Lesion	Characteristics			
	<i>Non-Melanocyte Lesions</i>	<i>Melanocyte Lesions</i>	<i>Benign</i>	<i>Malignant</i>
MEL		X		X
BCC	X			X
AKIEC	X			X
NV		X	X	
BKL	X		X	
DF	X		X	
VASC	X		X	

یازده شبکه مورد استفاده قرار گرفت

Model Name	Size (MB)	Parameters	Depth
CNN	90	28.6 M	4
AlexNet	233	61 M	8
VGG16	528	138.4 M	16
VGG19	549	143.7 M	19
ResNet50	98	25.6 M	107
InceptionV3	92	23.9 M	189
MobileNetV2	14	3.5 M	105
InceptionResNetV2	215	55.9 M	449
DenseNet121	33	8.1 M	242
DenseNet169	57	14.3 M	338
DenseNet201	80	20.2 M	402

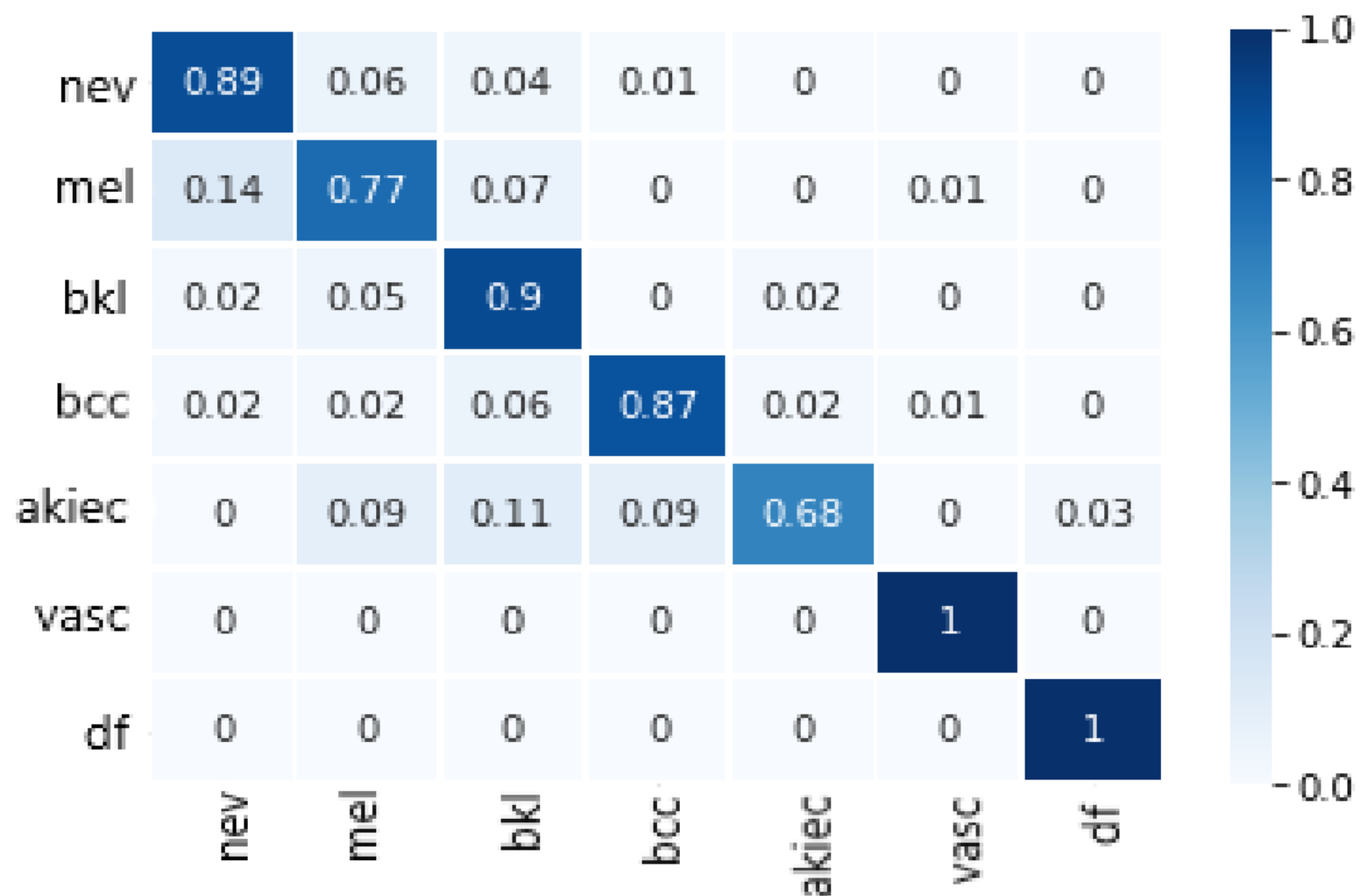


Figure 3. Confusion matrix for DenseNet169.

نتایج

Model Name	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN	82.15	85.73	88.53	87.11
AlexNet	80.74	86.25	88.14	87.19
VGG16	90.15	91.71	92.82	92.26
VGG19	85.55	88.44	89.34	88.89
ResNet50	90.36	86.49	84.04	85.25
InceptionV3	91.13	91.34	92.56	91.95
MobileNetV2	80.69	84.98	86.65	85.81
InceptionResNetV2	90.29	91.77	92.87	92.32
DenseNet121	91.20	92.12	93.08	92.60
DenseNet169	92.25	92.95	93.59	93.27
DenseNet201	90.14	91.20	92.85	92.02

نتایج مدل دو کلاسه

Table 5. Metrics for the two-class DenseNet169 mapping model.

Model	Accuracy	Specificity	Sensitivity
DenseNet169—two classes	91.10	95.67	82.49

- مدت زمان مجاز در معرض آفتاب قرار گرفتن بر اساس اندیس UV محل
- تجویز میزان مناسب ضد آفتاب برای نوع پوست کاربر
- تشخیص ضایعه خوش خیم از بدخیم بر اساس عکس دوربین موبایل

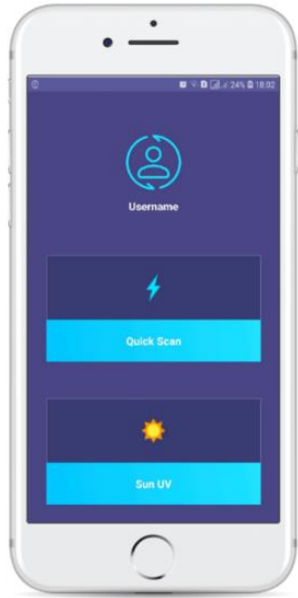
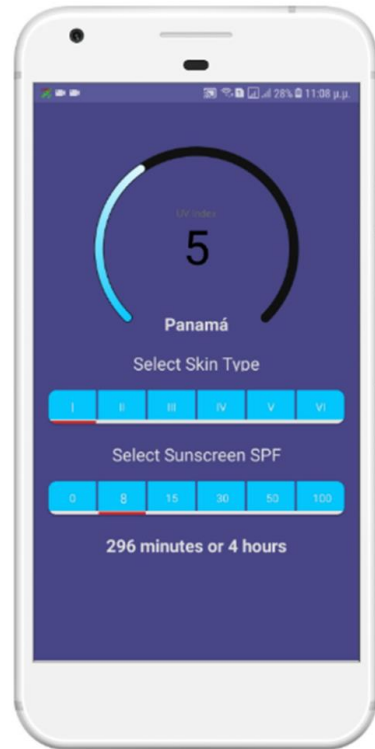
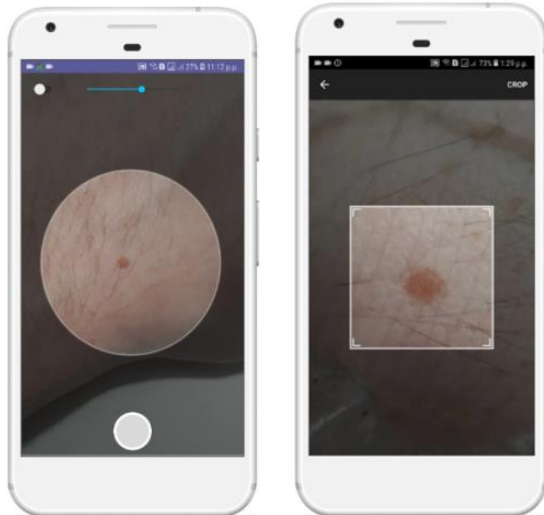


Figure 4. Application screenshot: main menu.

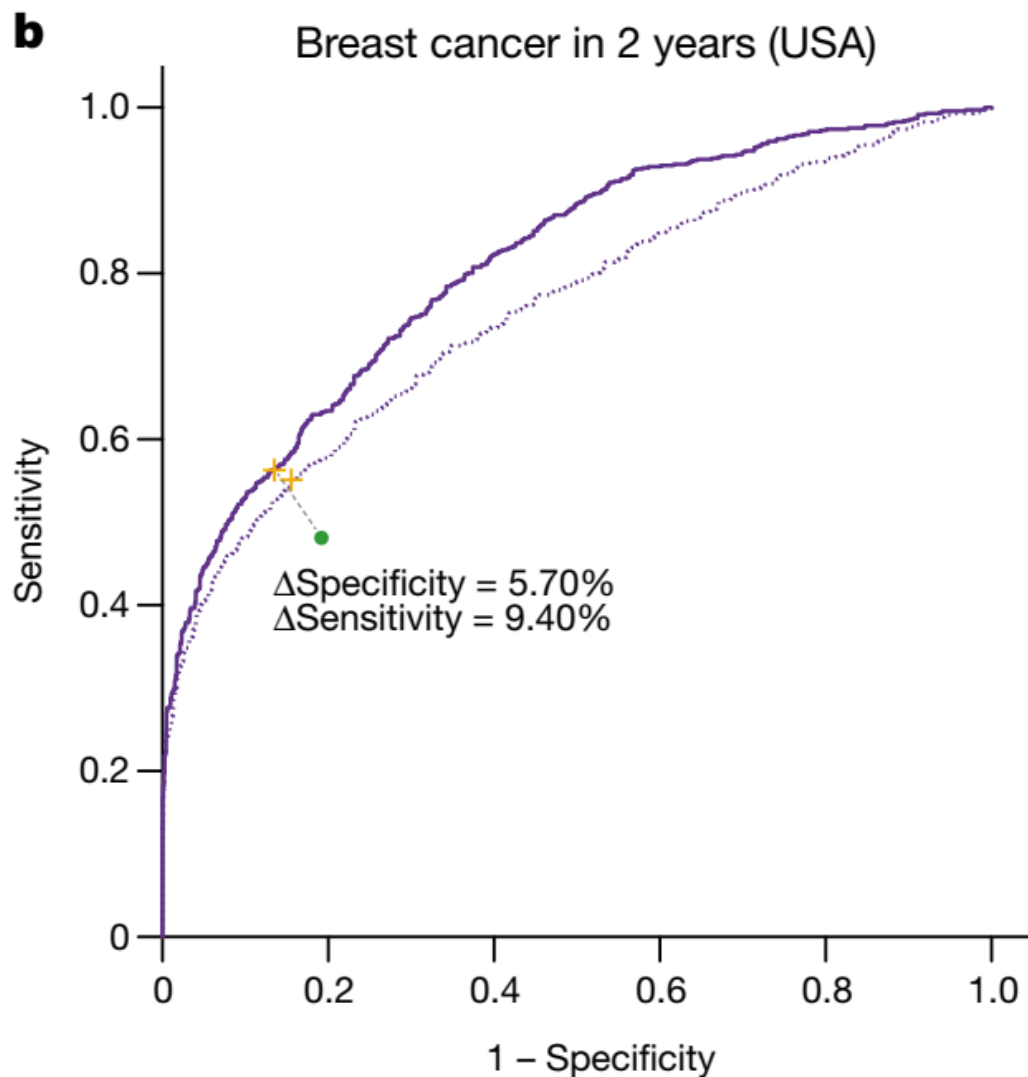


هوش مصنوعی در سرطان پستان

سرطان پستان

- یکی از شایع ترین انواع سرطان
 - بالاترین نرخ مرگ و میر بعد از سرطان ریه
 - نرخ مرگ و میر سرطان پستان طی دو دهه اخیر کاهش پیدا کرده: غربالگری و درمان های نوین
 - غربالگری؛ بهترین رویکرد
 - رایج ترین روش غربالگری سرطان پستان بر اساس ماموگرام
 - محدودیت: مثبت کاذب و منفی کاذب زیاد (**high false positive and false negative**)
- (rates)
- هوش مصنوعی با کاهش ۶۹ درصدی در موارد مثبت کاذب و افزایش حساسیت از ۸۴ تا ۹۱ درصد به تشخیص کمک می کند.

○ روش رایج در اروپا **double reader**، ارزیابی توسط دو رادیولوژیست (برای کاهش خطا)



○ استفاده از هوش مصنوعی به جای رادیولوژیست دوم

○ حجم کار تا ۸۸ درصد کاهش پیدا کرد

○ بهبود **sensitivity** و **specificity**

○ کاهش ۵.۷ درصدی **false positive**

○ کاهش ۹.۴ درصدی **false negative**

○ تایید وجود بیماری با بویپسی تا ۳۹ ماه بعد از غربالگری

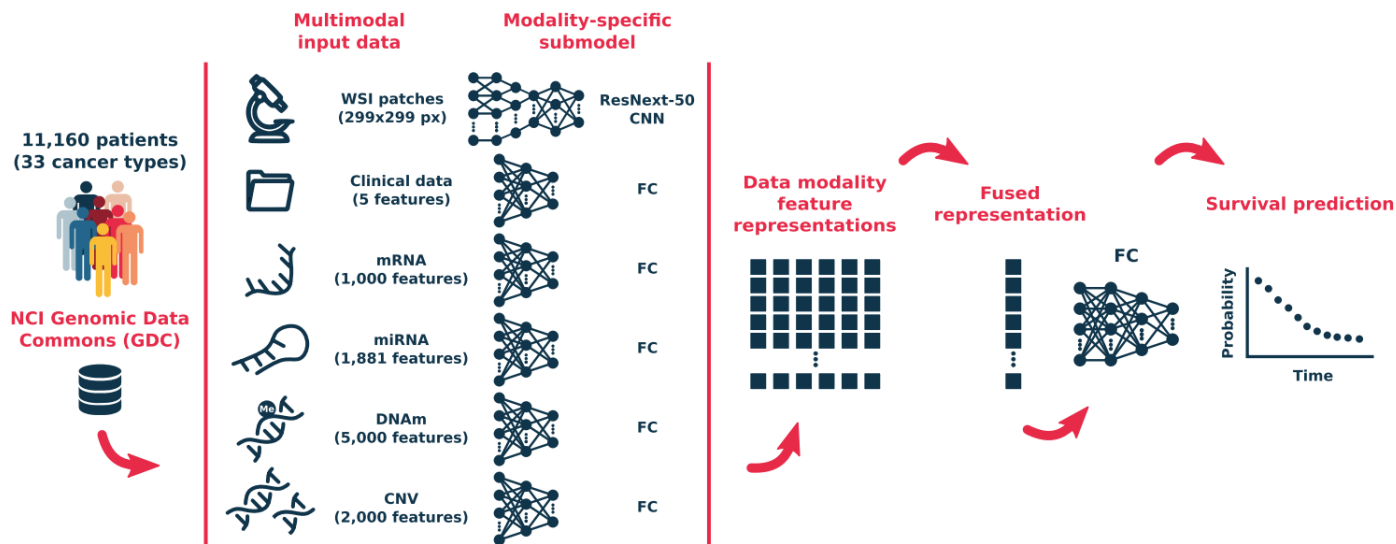
پیش بینی نرخ بقا سرطان پستان

○ پیش‌بینی نرخ بقا بر اساس عواملی مانند سن، اندازه تومور، وضعیت غده های زیر بغل، نوع بافت شناسی، تعداد میتوزی، پلئومورفیسم هسته ای

○ استفاده از داده های multi-modal

○ bi-phase: Deep learning + random forest

○ بهبود ۵ درصدی sensitivity



یادگیری عمیق با استفاده از داده های **multi-modal**: ژنتیک، هیستوپاتولوژی، داده های کلینیکی

مدل های دوفازی

Sigmoid gated attention neural network

Random forest classifier

پنج درصد بهبود پیشبینی نرخ بقا (پنج ساله)

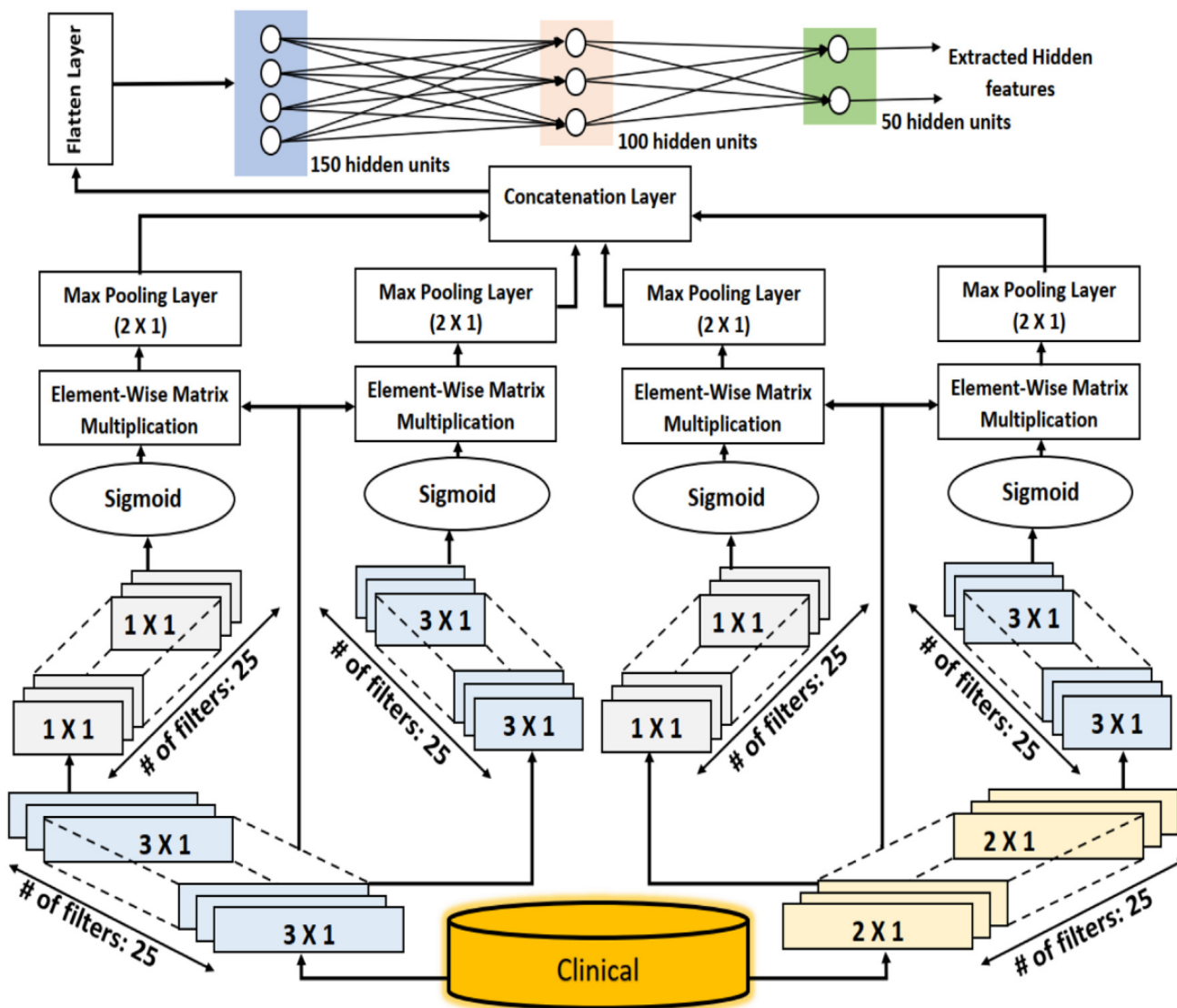


Fig. 1. SiGaAtCNN as feature extractor from clinical uni-modal data.

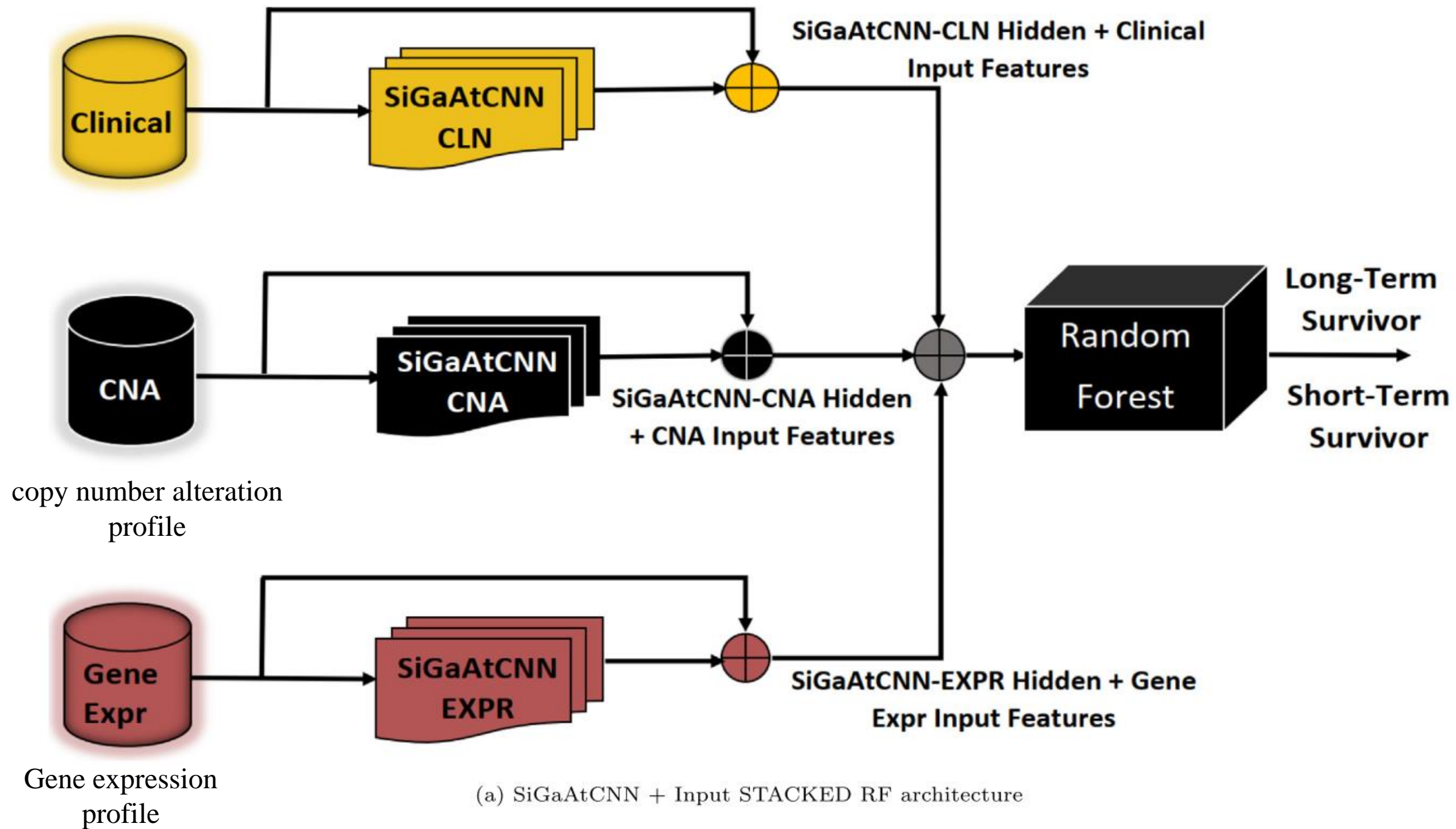


Table 5

Performance indicators of various models for breast cancer survival prediction.

Model	AUC	Acc	Pre	Sn	Mcc
(Sp = 95%)					
SiGaAtCNN + Input STACKED RF	0.950	0.912	0.841	0.798	0.762
SiGaAtCNN STACKED RF	0.943	0.911	0.841	0.790	0.757
SiGaAtCNN Bi-Attention	0.841	0.820	0.723	0.446	0.467
SiGaAtCNN Bi-Attention STACKED RF	0.828	0.814	0.726	0.401	0.441
MDNNMD	0.845	0.826	0.749	0.450	0.486
STACKED RF	0.930	0.902	0.841	<u>0.747</u>	0.730

Survival Analysis Learning with Multi-Omics Neural Networks (SALMON)

- نیاز به الگوریتمهای پیچیده برای تحلیل فضای پیچیده دیتای ورودی
- بررسی predictor های احتمالی
- یادگیری عمیق ابزاری مناسب برای آنالیز بقا

◦ **Deep Learning-based** networks to determine how **gene expression data** predicts **Cox regression survival** in breast cancer.

◦ روش پیشنهادی ویژگی های gene exp و بایومارکرهای سرطان را ترکیب و ساده سازی میکند تا به prognosis برسد.

◦ نقش مهم omics data در بهبود مدل در مقایسه با تنها اطلاعات بیان ژنی

OMICS data

"Omics" به رشته ای از مطالعات در علوم زیستی اشاره دارد که شامل تجزیه و تحلیل جامع مولکول های بیولوژیکی است که به ساختار، عملکرد و پویایی موجودات زنده کمک می کند. اصطلاح "omics" معمولاً به عنوان پسوند برای توصیف شاخه های مختلف زیست شناسی استفاده می شود که بر انواع خاصی از مولکول ها تمرکز می کنند و به محققان این امکان را می دهد که به جای تمرکز بر اجزای منفرد، بینشی در مورد سیستم های بیولوژیکی به عنوان یک کل به دست آورند.

در اینجا برخی از زمینه های کلیدی تحقیق omics آورده شده است:

Genomics: این شامل مطالعه مجموعه کامل ژن ها (ژنوم) در یک موجود زنده است. هدف ژنومیکس درک ساختار، عملکرد و تعاملات ژن است. اغلب شامل تعیین توالی و تجزیه و تحلیل DNA است.

Transcriptomics: این حوزه بر مطالعه مولکول های RNA تولید شده از ژنوم، به ویژه رونوشت، که از تمام رونوشت های RNA پیام رسان mRNA تشکیل شده است، تمرکز دارد. ترنسکریپتومیکس به درک الگوهای بیان ژن و کمک می کند.

Proteomics: مطالعه جامع کل مجموعه پروتئین ها (پروتئوم) تولید شده در یک موجود زنده است. پروتئومیکس عملکردها، ساختارها، تعاملات و تغییرات پروتئین را بررسی می کند و بینش هایی را در مورد فرآیندهای سلولی ارائه می دهد.

Metabolomics: این رشته مجموعه کامل متابولیت ها (متابولوم) را در یک نمونه بیولوژیکی، مانند یک ارگانسیم یا بافت بررسی می کند. تجزیه و تحلیل متابولیت ها امکان ارزیابی فرآیندها و مسیرهای متابولیک را فراهم می کند.

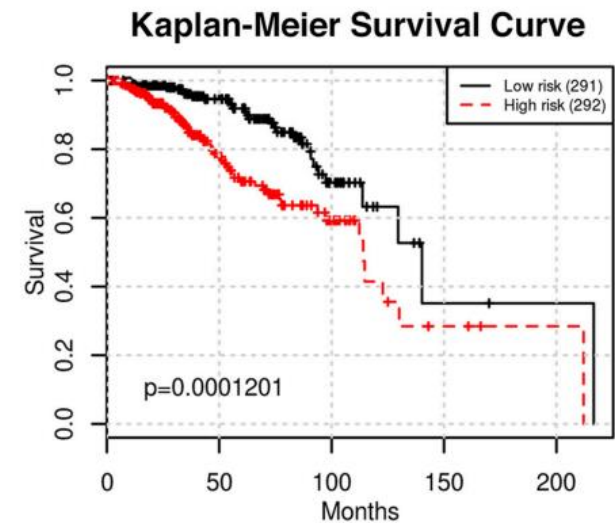
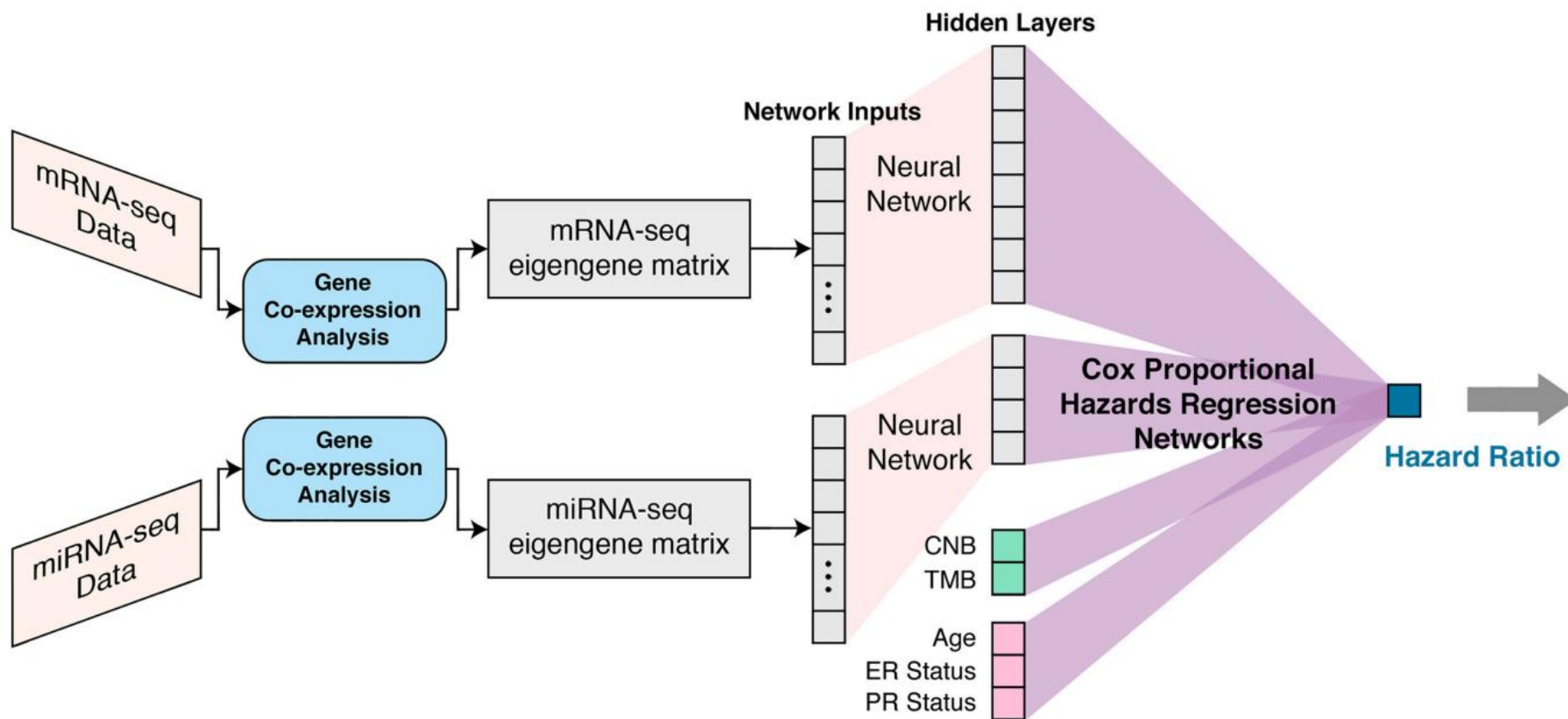
Lipidomics: این حوزه بر تجزیه و تحلیل جامع لیپیدها (اسیدهای چرب و مولکول های مرتبط) در یک سیستم بیولوژیکی متمرکز است. Lipidomics می تواند اطلاعات مهمی را در مورد پویایی غشای سلولی و مسیرهای سیگنال دهی نشان دهد.

Glycomics: این حوزه به مطالعه ساختار، عملکرد و بیولوژی کربوهیدرات ها (گلیکان ها) و نقش آنها در فرآیندهای مختلف بیولوژیکی می پردازد.

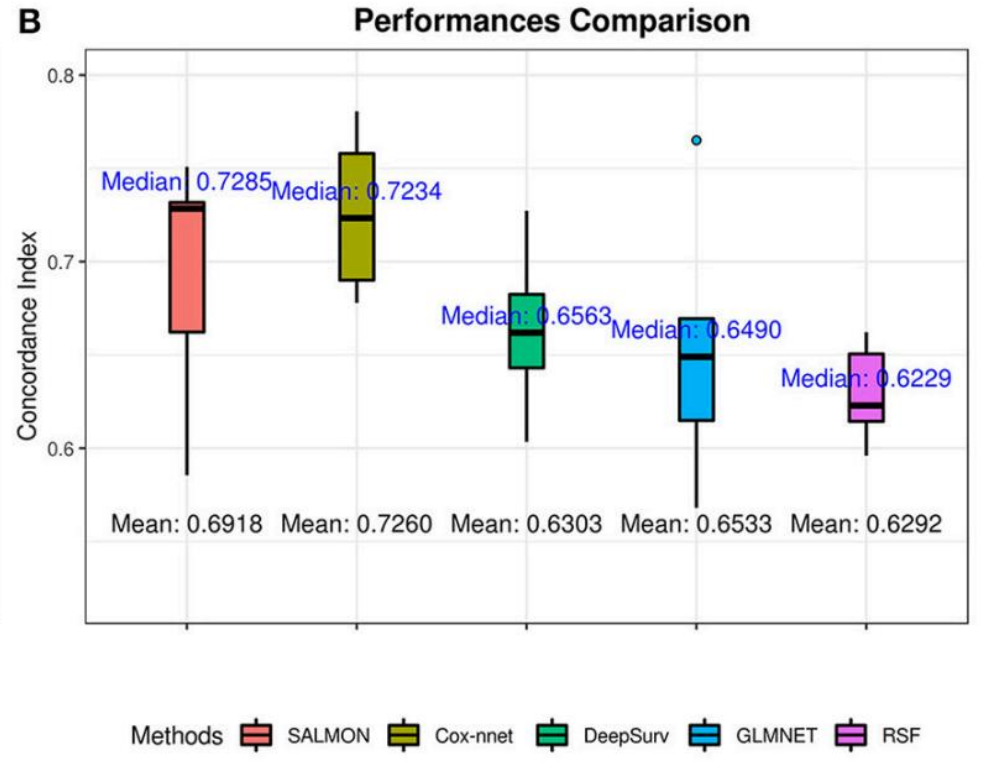
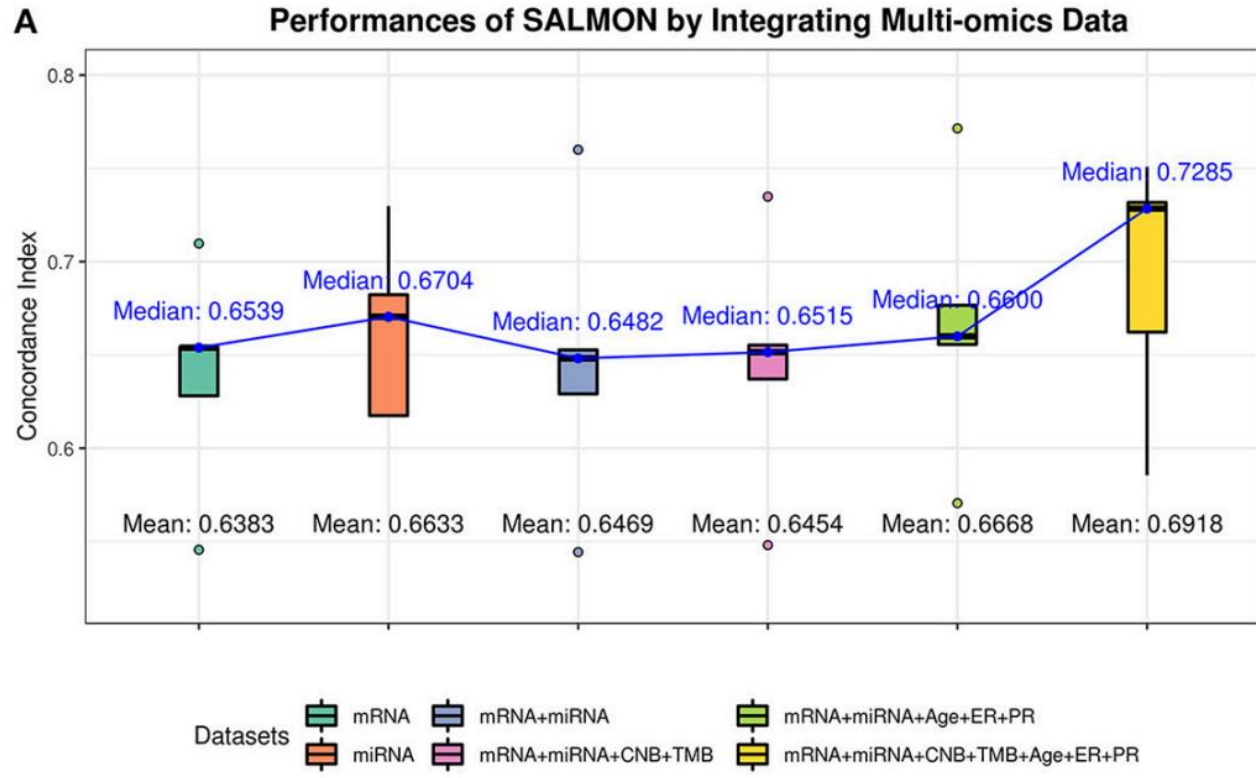
Epigenomics: این حوزه شامل مطالعه تغییرات در فعالیت ژنی است که شامل تغییراتی در توالی DNA زمینه ای نمی شود، از جمله متیلاسیون DNA و اصلاح هیستون.

داده‌های Omics به داده‌های مقیاس بزرگ تولید شده از این مطالعات اشاره دارد که می‌تواند اطلاعات زیادی در مورد سیستم‌های بیولوژیکی ارائه دهد. ادغام داده‌های مختلف omics، که اغلب به عنوان "مولتی omics" شناخته می‌شود، به محققان اجازه می‌دهد تا درک جامع‌تری از فرآیندهای بیولوژیکی و نحوه تعامل انواع مولکولی در یک موجود زنده به دست آورند.

رویکردهای Omics کاربردهای گسترده‌ای در زمینه‌هایی مانند پزشکی، کشاورزی، بیوتکنولوژی و علوم محیطی دارند و به پیشرفت در پزشکی شخصی، تحقیقات بیماری‌ها و درک تعاملات پیچیده بیولوژیکی کمک می‌کنند.



- Concordance index = AUC >>> 0.5 = ineffective
- Concordance index > 0.7 good predictive model



mRNA = messenger RNA
 miRNA – micro RNA
 CNB = copy number burden

TMB = tumor mutation burden
 ER = estrogen receptors
 PR = progesterone receptors

هوش مصنوعی در سرطان روده بزرگ

پولیپ روده بزرگ و سرطان کولورکتال

◦ سومین کنسر شایع در امریکا

◦ ۱۴۷ هزار مورد جدید در سال ۲۰۲۰ (incidence)

◦ کاربرد هوش مصنوعی در غربالگری، تشخیص و درمان (CRC)

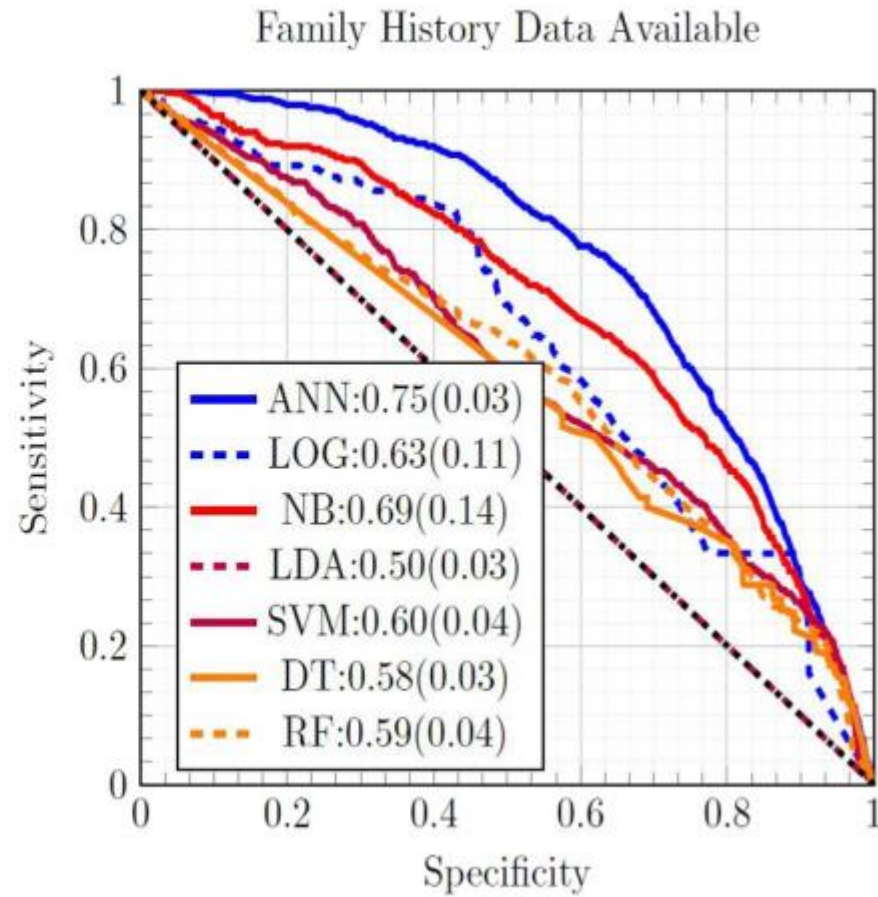
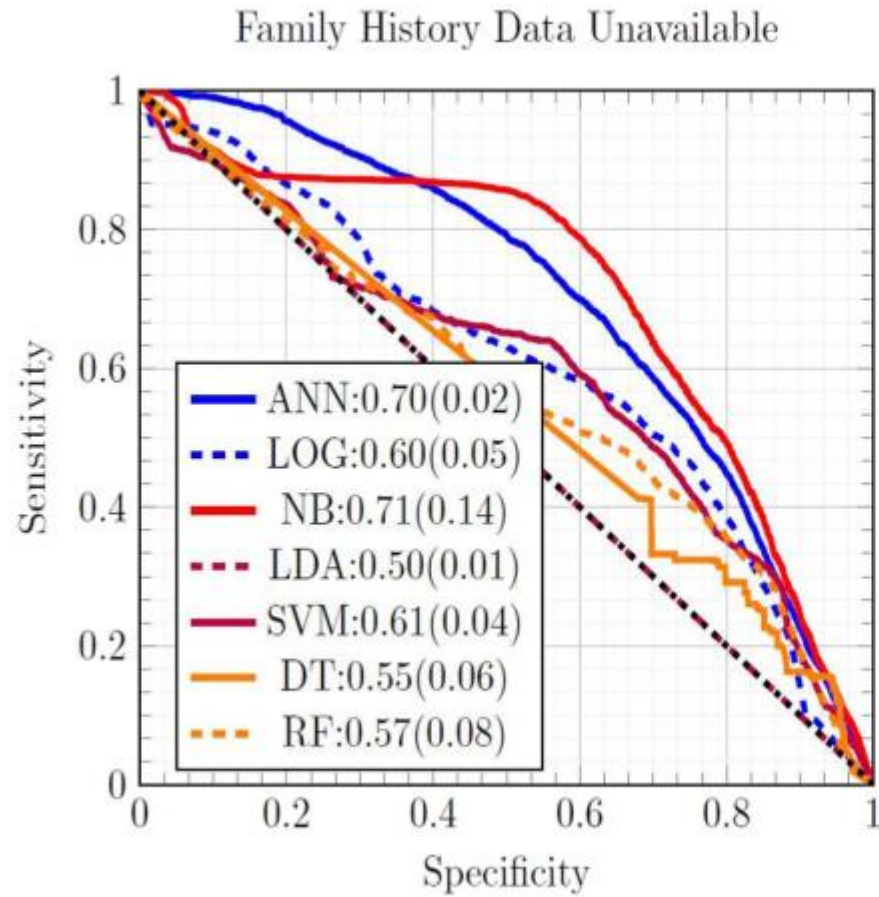
غربالگری عمومی برای CRC با استفاده از یادگیری ماشین

- غربالگری در حال حاضر بر اساس سابقه خانوادگی و سن افراد
- انجمن سرطان آمریکا: افراد در سنین ۴۵ تا ۷۵
- خصوصاً افرادی که در خانواده مورد CRC داشتند
- با این وجود، برای افراد ۱۸ تا ۴۵ و بدون سابقه خانوادگی احتمال خطر وجود دارد
- پیش بینی بر اساس مطالعات روتین (و نه بیومارکرها)
- داده ها از ((National Health Interview Survey (NHIS)) اخذ به روش پرسشنامه ای
- تا ۱۴ سال چک کردن برای کنسر ادامه داشته
- hypertension, ulcers, a stroke, any liver comorbidity, arthritis, bronchitis, coronary heart disease, myocardial infarction, and/or emphysema >> binary variable (0 -1)
- not diabetic, pre-diabetic/borderline, and diabetic 0, 0.5, 1
- Age, BMI, gender, Not Hispanic/Spanish origin, smoking status, family history

غربالگری عمومی برای CRC با استفاده از یادگیری ماشین

- برای غربالگری عمومی (mass screening) هفت روش مقایسه شده
- linear discriminant analysis, support vector machine, naive Bayes, decision tree, random forest, logistic regression, and artificial neural network
- بهترین نتیجه برای شبکه عصبی مصنوعی
 - دو لایه
 - تابع فعالسازی سیگموئید
 - گرادیان نزولی افقی
- نتایج
 - $AUC = 0.75$
 - Sensitivity = 0.63
 - Specificity = 0.82

تأثير سوابق خانوادگی



- دقت بهتر در گروه سنی جوانتر
- هدف مقاله هم تمرکز روی گروه جوانتر و بدون سابقه خانوادگی بود

	Family History Data Unavailable							Family History Data Available							
	ANN	LOG	NB	SVM	LDA	DT	RF	ANN	LOG	NB	SVM	LDA	DT	RF	
Ages 18-49	EM	78(06)	69(14)	69(14)	69(12)	50(05)	55(06)	58(08)	82(11)	74(20)	63(14)	62(15)	50(11)	55(12)	60(14)
	Gauss	76(07)	83(07)	60(09)	57(09)	50(05)	57(07)	63(06)	76(11)	78(22)	50(11)	56(11)	50(11)	56(12)	61(15)
	1-hot	71(06)	85(26)	68(06)	73(05)	69(14)	51(05)	54(15)	81(10)	68(17)	50(11)	80(10)	78(12)	73(20)	72(16)
	Cauchy	75(06)	67(15)	54(06)	54(06)	51(05)	60(07)	59(09)	70(11)	61(13)	53(11)	63(12)	52(11)	54(11)	56(12)
	Mean	76(06)	76(08)	51(05)	54(06)	50(05)	59(11)	62(08)	73(11)	71(19)	51(11)	53(11)	64(15)	50(11)	52(15)
	Delete	81(06)	51(06)	51(06)	62(11)	50(06)	65(07)	67(06)	83(12)	80(25)	51(14)	64(18)	50(14)	60(15)	63(18)
Ages 50-75	EM	64(05)	52(02)	58(05)	55(02)	56(04)	54(03)	54(03)	64(04)	58(05)	57(06)	54(03)	63(10)	55(03)	59(06)
	Gauss	58(02)	58(05)	57(02)	52(02)	62(08)	51(02)	53(05)	59(04)	58(08)	57(04)	52(03)	50(03)	54(03)	54(05)
	1-hot	59(02)	50(01)	58(05)	52(02)	62(08)	54(02)	55(04)	58(05)	52(03)	59(07)	50(03)	63(10)	56(06)	57(06)
	Cauchy	60(04)	59(02)	52(02)	56(02)	50(01)	53(02)	54(02)	59(06)	57(03)	53(04)	54(04)	50(03)	55(05)	54(04)
	Mean	60(03)	61(08)	58(02)	52(02)	50(01)	51(02)	54(04)	61(05)	62(09)	57(04)	53(03)	61(05)	51(03)	54(06)
	Delete	61(03)	52(02)	54(05)	52(02)	62(09)	53(04)	56(05)	58(04)	51(03)	57(03)	53(03)	63(09)	53(05)	56(06)
All ages	EM	70(02)	60(05)	71(14)	61(04)	50(01)	55(06)	55(04)	75(03)	63(11)	69(14)	60(04)	50(03)	58(03)	57(06)
	Gauss	73(13)	73(17)	72(12)	54(02)	74(17)	53(01)	56(12)	67(15)	66(18)	69(12)	56(07)	78(20)	53(03)	56(14)
	1-hot	69(11)	53(02)	66(11)	53(02)	50(01)	60(06)	54(03)	56(04)	53(03)	65(11)	54(03)	50(03)	59(08)	61(05)
	Cauchy	73(13)	63(02)	68(12)	61(07)	50(01)	58(05)	57(03)	67(13)	63(05)	58(05)	55(05)	50(03)	55(03)	56(04)
	Mean	73(13)	76(18)	71(11)	54(03)	78(20)	54(02)	59(13)	69(12)	76(18)	68(11)	54(03)	79(20)	54(05)	59(14)
	Delete	74(13)	53(02)	61(04)	59(07)	61(08)	51(02)	55(04)	67(12)	52(03)	54(04)	53(03)	77(20)	56(04)	59(09)

ColonFlag

- شناسایی افراد با ریسک بالای احتمال به کلورکتال کنسر بر پایه
 - اطلاعات پایه بیمار
 - تست خون complete blood cell count (CBC)
- طراحی نرم افزار کولن فلگ برای شناسایی high risk adenomatous در کولونوسکوپی
- جامعه مورد مطالعه، افراد بدون علامت بین ۵۰ تا ۷۵ ساله، تحت کولونوسکوپی غربالگری قرار گرفتند. همه افراد حداقل یک نتیجه CBC در سال قبل از کولونوسکوپی داشتند.
- بر اساس سن، جنس، پارامترهای گلبول قرمز، سلول‌های التهابی و پلاکت‌ها، الگوریتم ColonFlag امتیازی از ۰ تا ۱۰۰ تولید میکند
- At specificity =0.95, odds ratio (OR) = 2
- از آزمایش خون ساده روتین استفاده میکند برای شناسایی افراد با ریسک بالا برای پولیپ‌ها بدخیم

مدل پیشگو برای سرطان colorectal

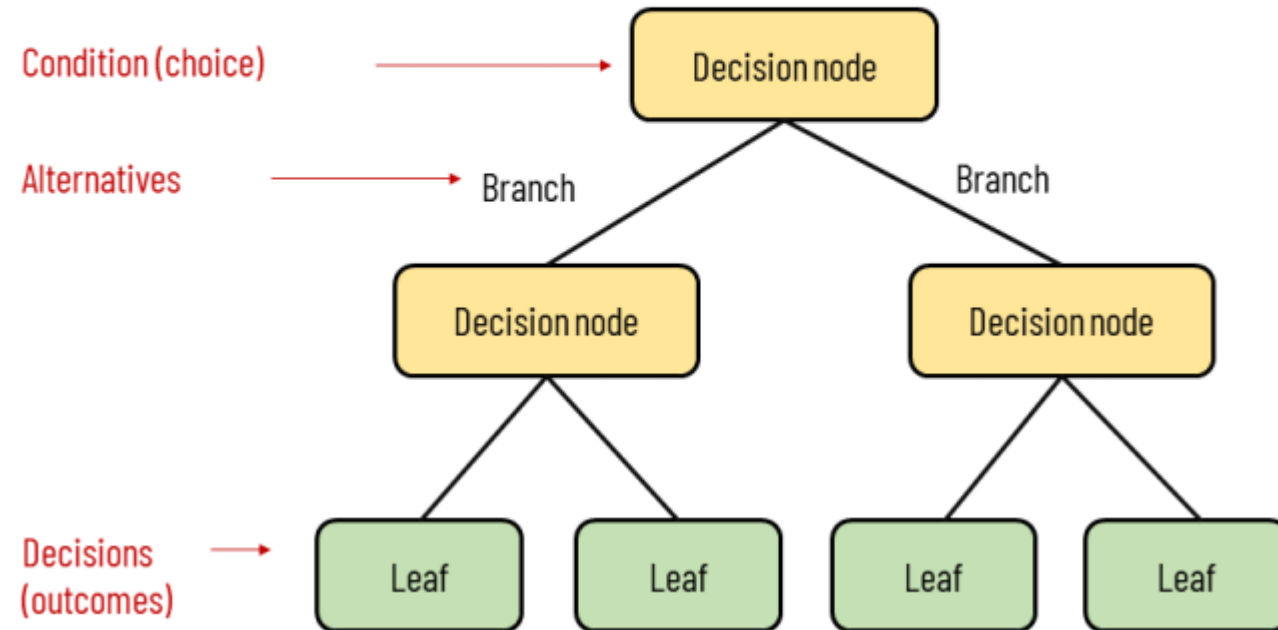
- پیشبینی کنسر با استفاده از سن، جنسیت و آزمایش خون ساده
- دو جامعه آماری ۳۱۲۵ و ۵۰۶۱ مبتلا و گروه کنترل ۶۰۰۰۰ و ۲۵۰۰۰ نفری
- آزمایش خون سه تا شش ماه قبل از تشخیص
- مدل طبقه بندی یادگیری ماشین بر اساس **ensembles of decision trees**

◦ نتایج:

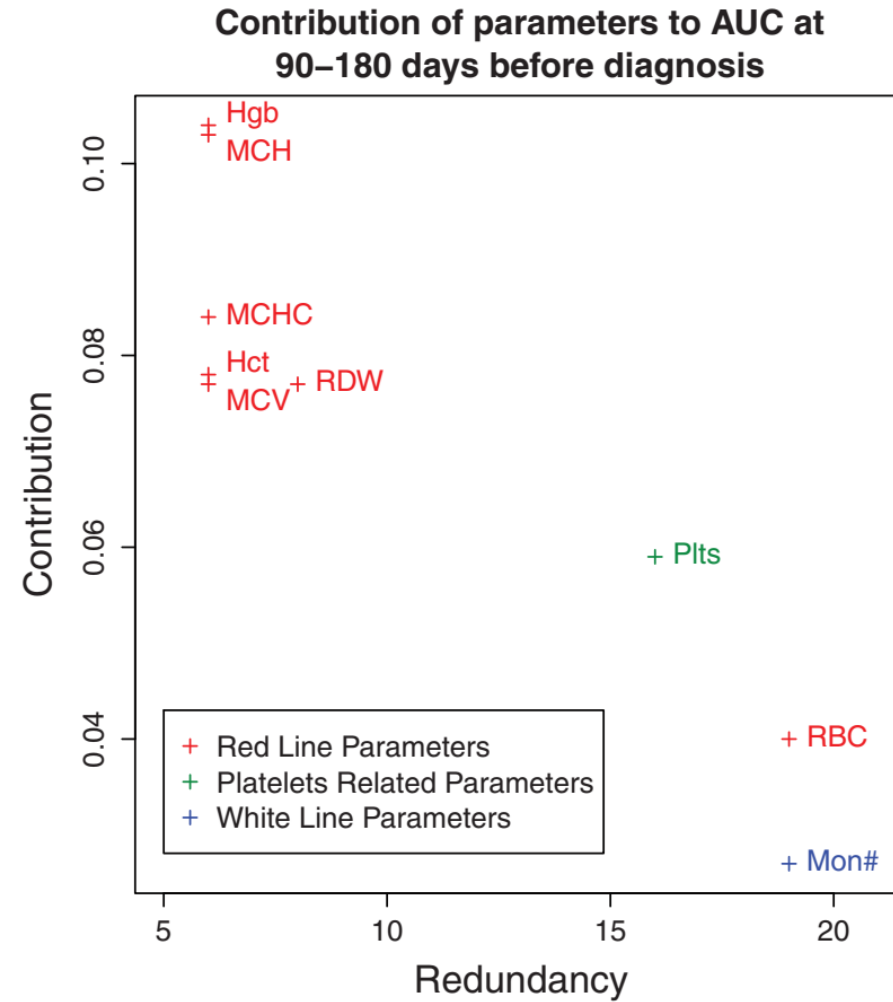
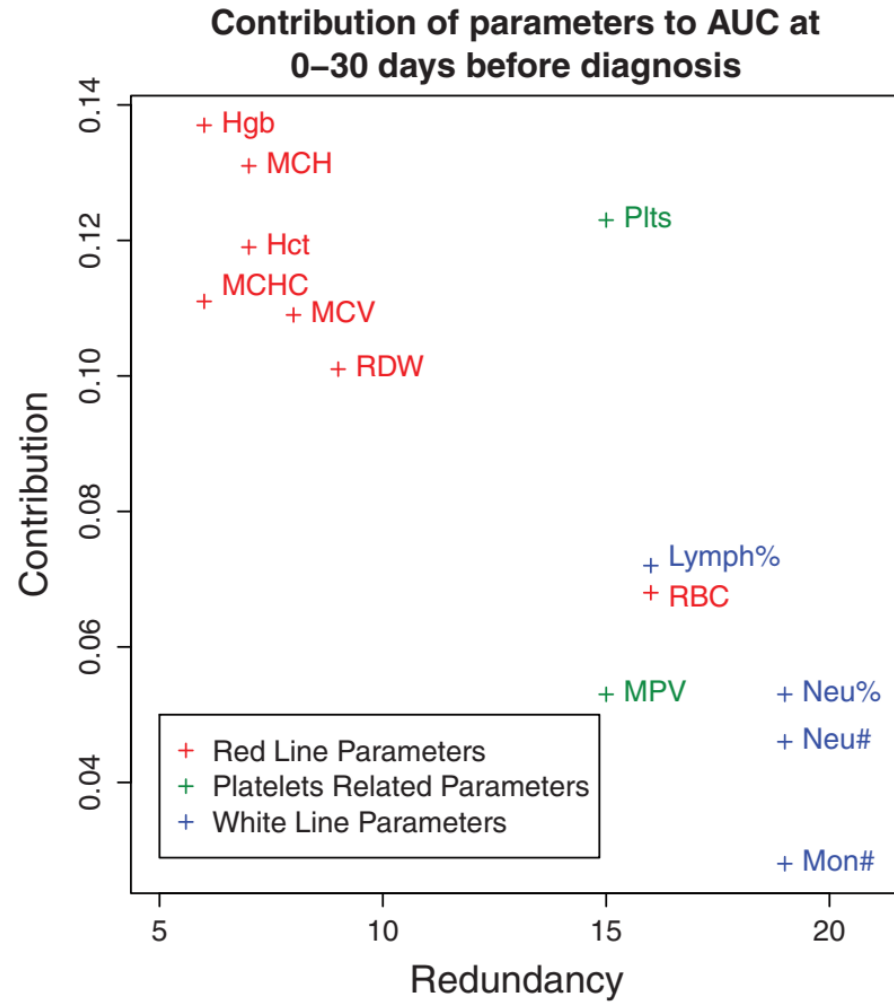
◦ **AUC = 0.82,**

◦ **false-positive rate = 0.5%, Odds Ratio (OR) = 40**

Elements of a decision tree



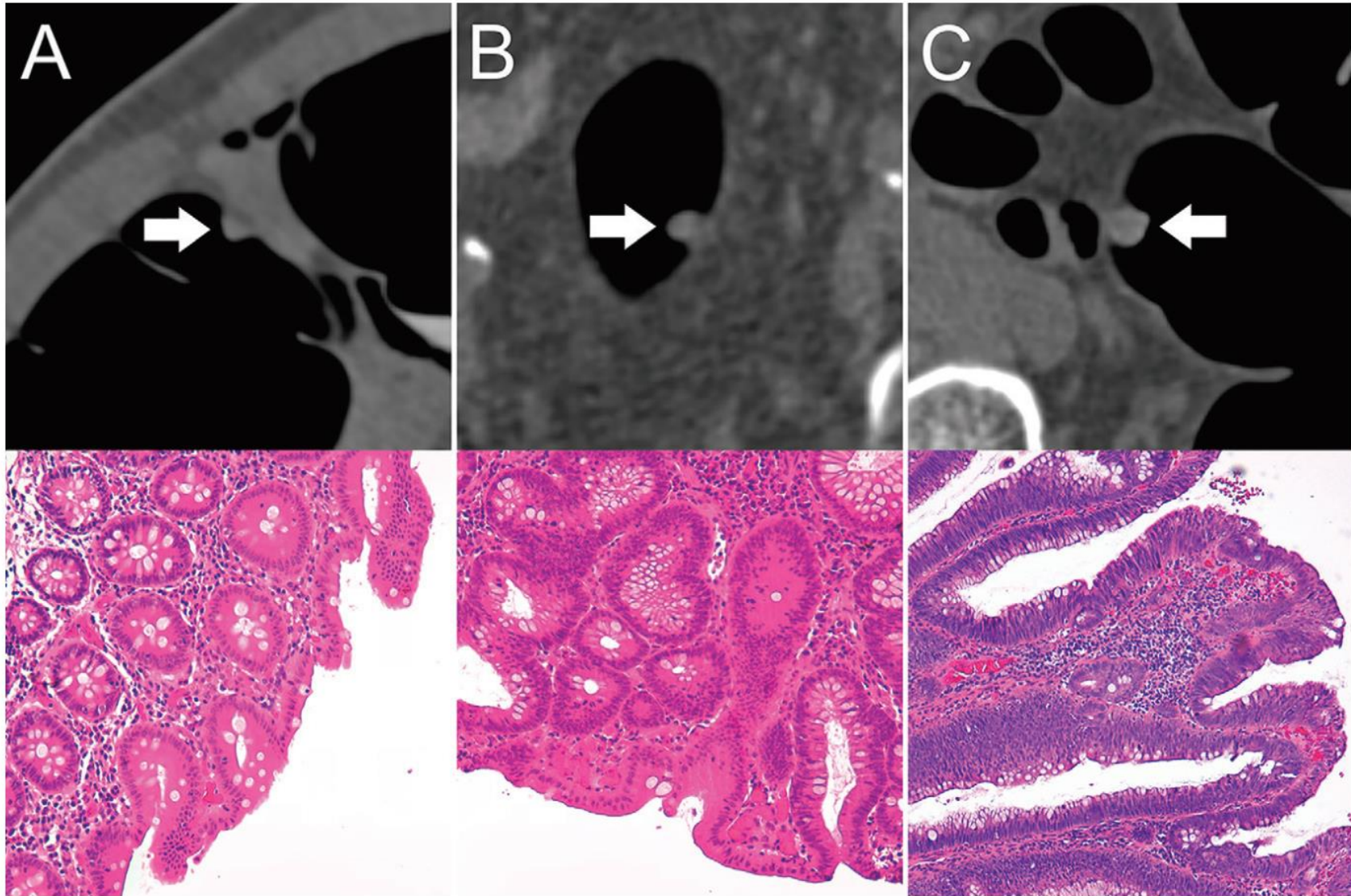
اثر پارامترهای ورودی الگوریتم در نتایج خروجی

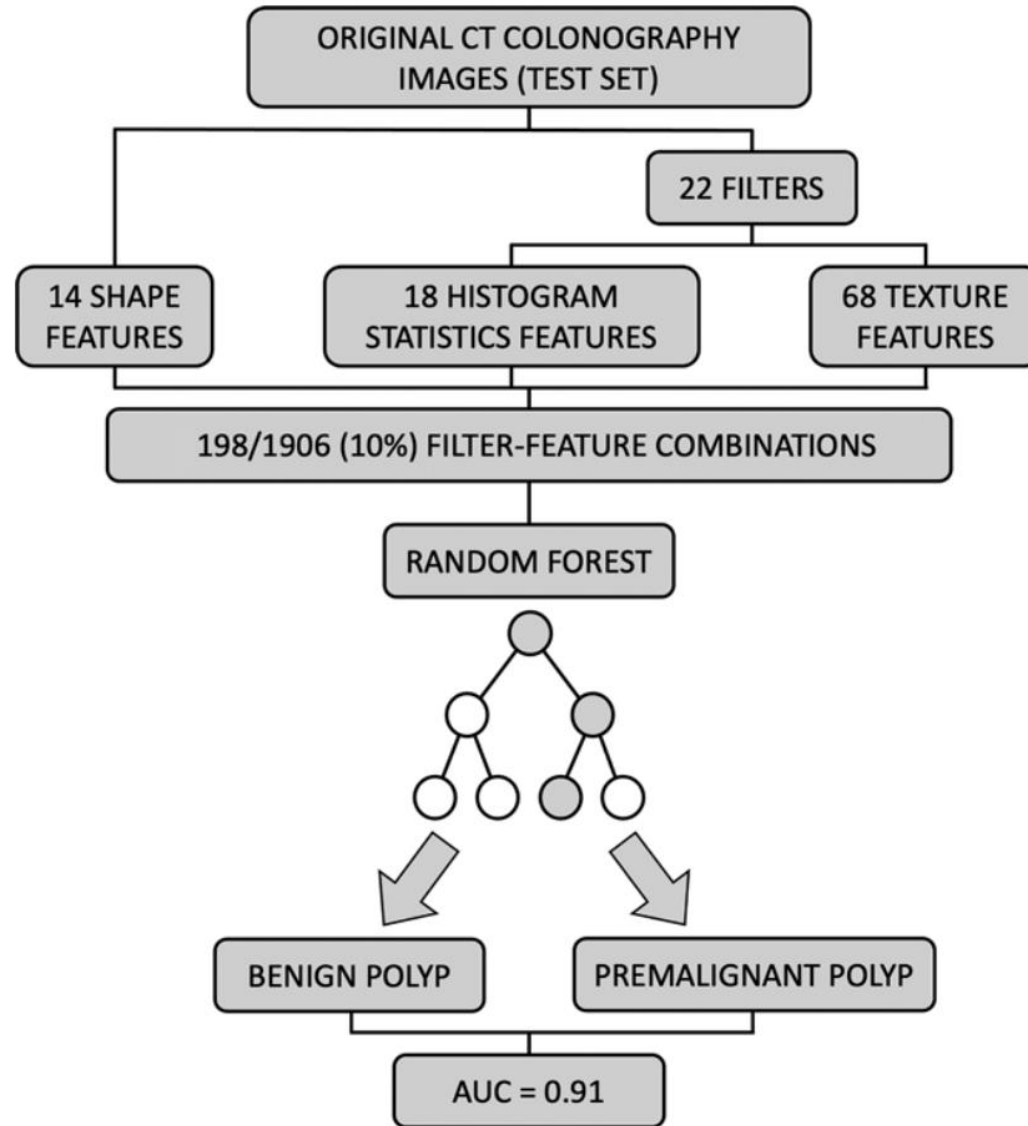


طبقه بندی بر اساس CT colonography

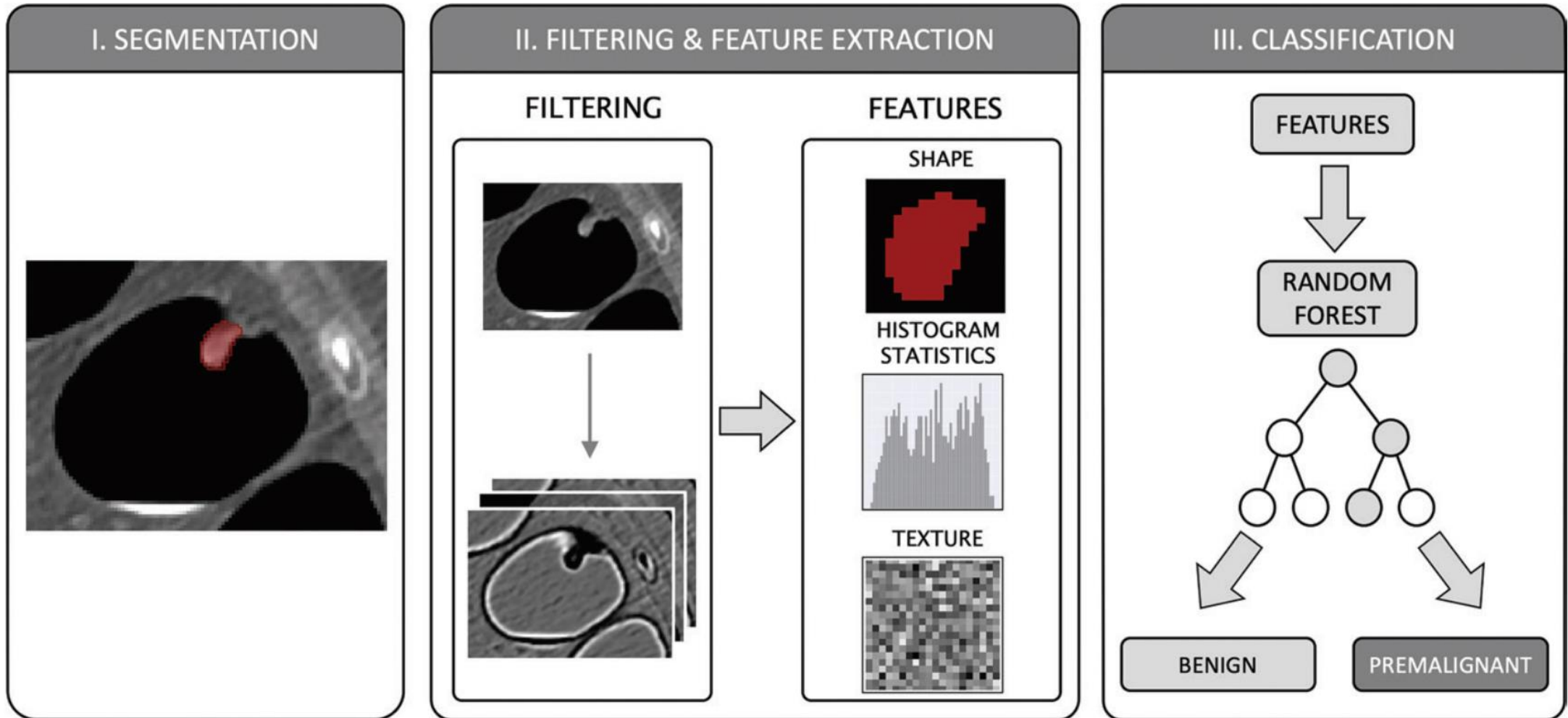
- طبقه بندی پولیپ های کولورکتال خوش خیم و premalignant (adenoma)
- بر اساس ۲۲ فیلتر ویژگیهای مختلف از تصاویر CT colonoscopy استخراج شد
- ۱۹۰۶ ویژگی استخراج شد، ۱۹۸ برای ساخت مدل (ویژگی های با ضریب همبستگی بالا ۰.۸ حذف شدند)
- طبقه بندی بر اساس random forest
- آموزش بر اساس دادگان ۶۳ بیمار
- تست بر اساس دادگان ۵۹ بیمار
- لیبل دهی بر اساس هیستوپاتولوژی
- نتایج
 - $AUC = 0.91$
 - $Sensitivity = 0.82$
 - $Specificity = 0.85$
- موثرترین ویژگی استخراجی first-order gray level histogram

Grosu, S., Wesp, P., Graser, A., Maurus, S., Schulz, C., Knösel, T., ... & Kazmierczak, P. M. (2021). Machine learning–based differentiation of benign and premalignant colorectal polyps detected with CT colonography in an asymptomatic screening population: a proof-of-concept study. *Radiology*, 299(2), 326-335.





خلاصه فرایند

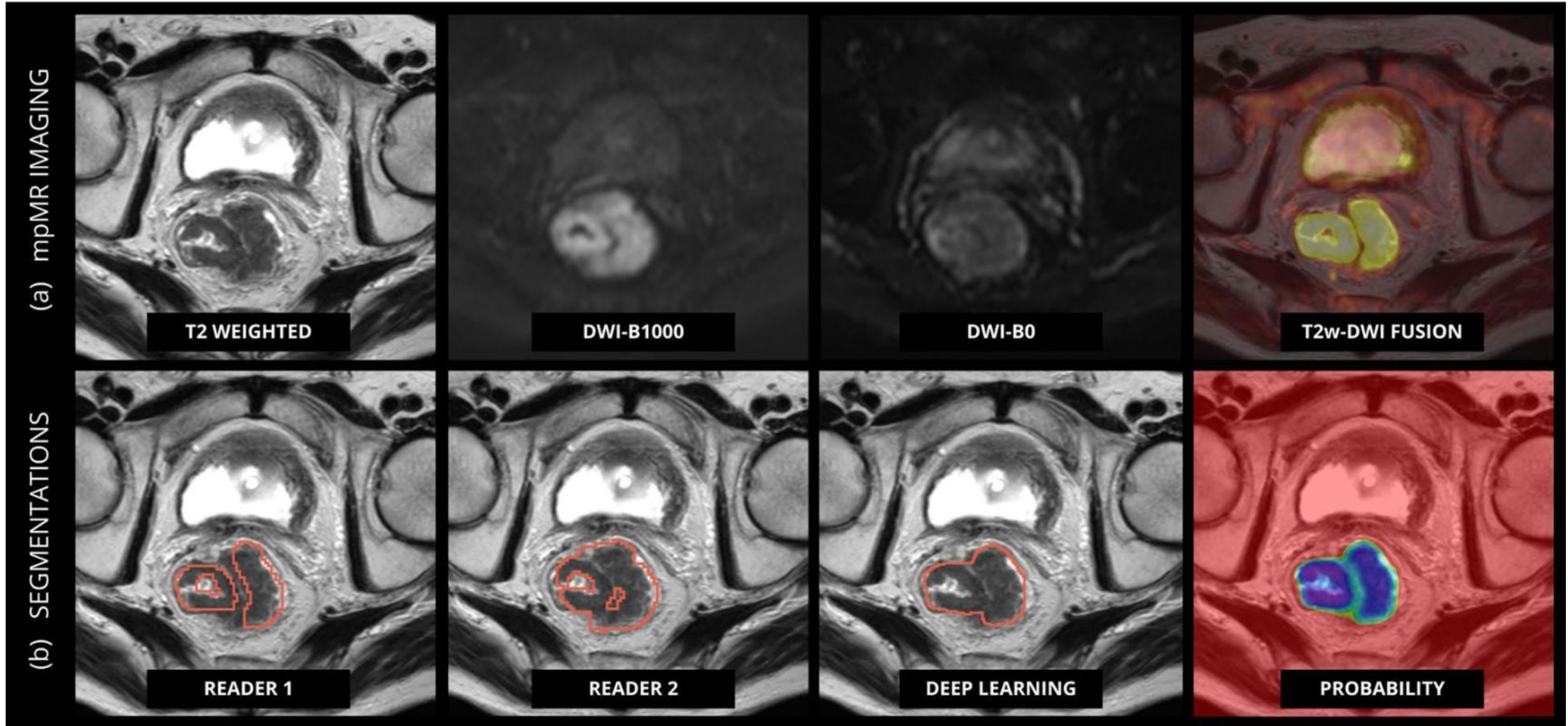


Deep Learning for Fully-Automated Localization and Segmentation of Rectal Tumors

- مکانیابی و سگمنت کردن اتوماتیک تصاویر MRI از تومورهای روده بزرگ
- آنالیز حجمی تصاویر MRI
- داده های ۱۴۰ بیمار
- تصویر هر تومور به وسیله دو متخصص سگمنت شد
- هدف: کلاسبندی هر Voxel به تومور و غیرتومور
- روش: شبکه convolutional neural network (CNN)
- نتایج
- **AUC = 0.99 (SD=0.05)**
- دقت و سرعت بسیار بالاتر

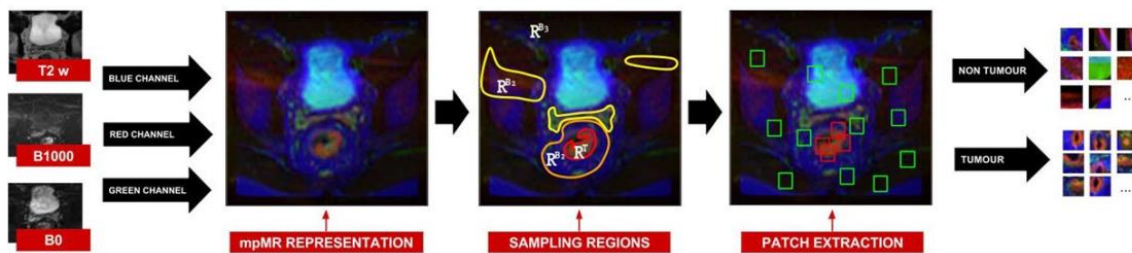
Trebeschi, S., van Griethuysen, J. J., Lambregts, D. M., Lahaye, M. J., Parmar, C., Bakers, F. C., ... & Aerts, H. J. (2017). Deep learning for fully-automated localization and segmentation of rectal cancer on multiparametric MR. *Scientific reports*, 7(1), 5301.

- Example of Multiparametric MR in a rectal cancer patient
- By repeating the procedure for each voxel of each image, we could generate a probability map, where $p(v)$ is the probability of voxel v to represent tumour tissue.

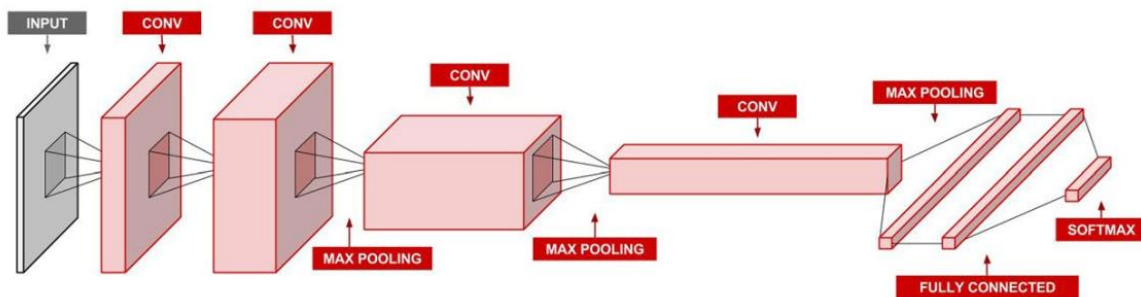


مراحل

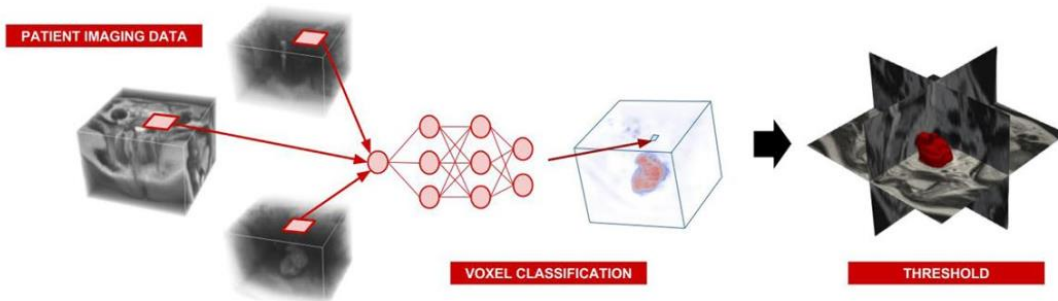
(a) PATCH EXTRACTION



(b) TRAINING

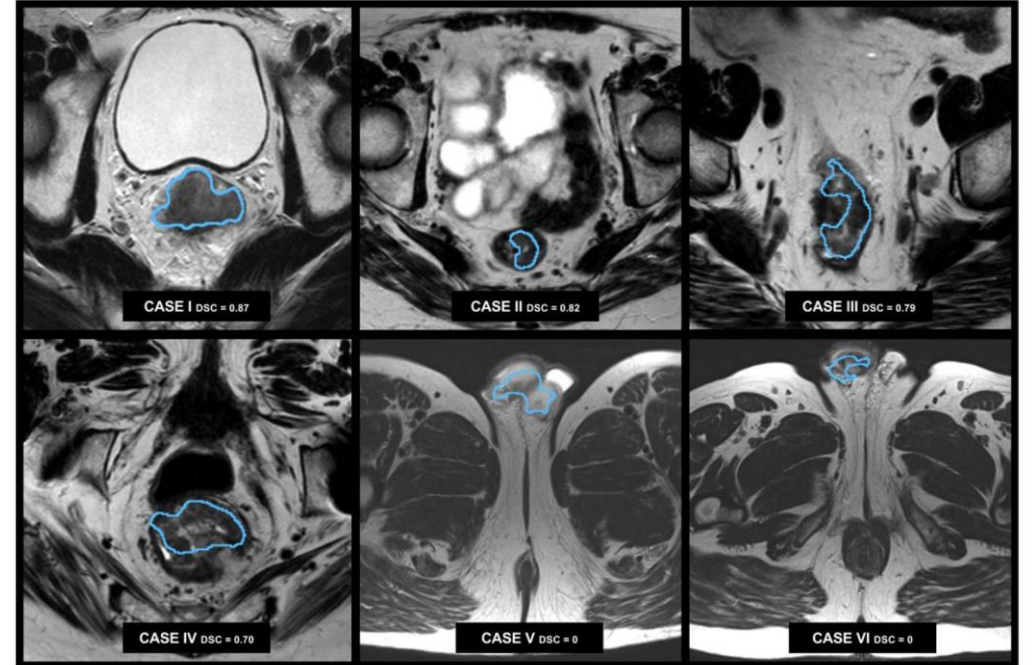
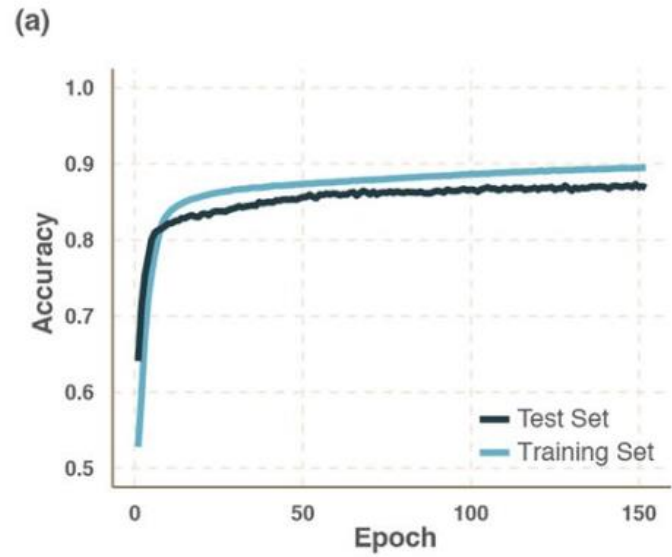
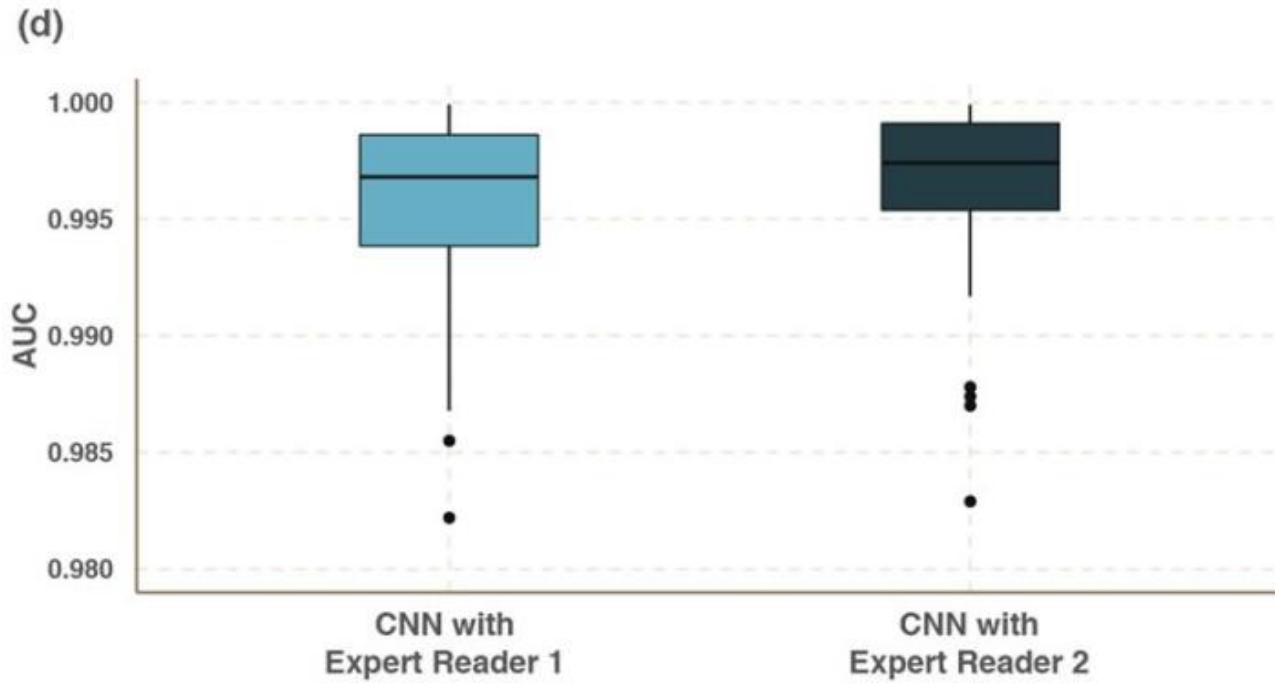


(c) VALIDATION



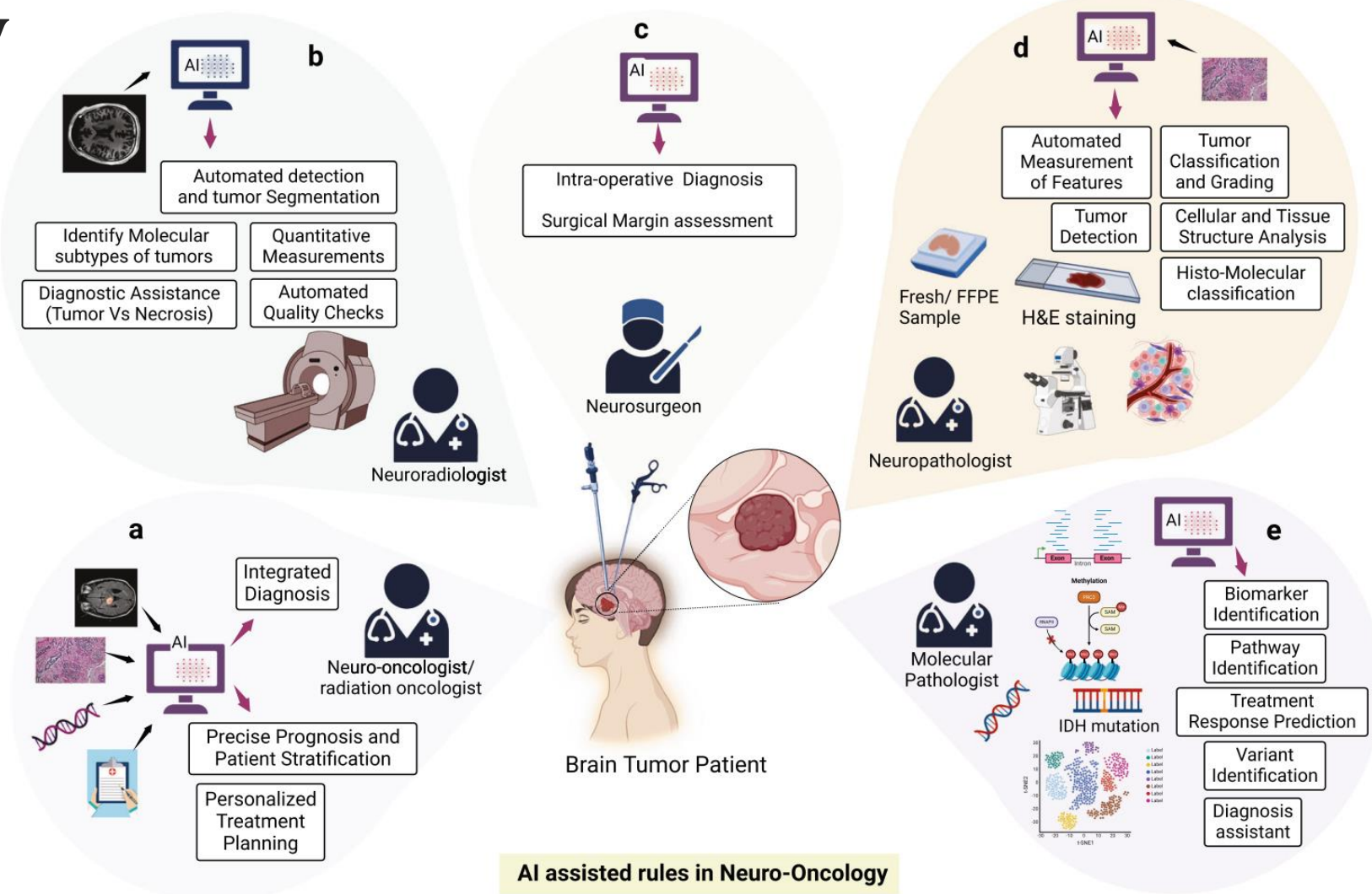
- Patch extraction: در آموزش مدل، به صورت تصادفی تعداد مساوی واکسل از ناحیه سالم و تومور برداشته میشود. این مطلب برای متعادل بودن نمونه های سالم و غیرسالم برای مرحله آموزش است. نواحی مختلف عکس سختی متفاوتی برای طبقه بندی دارند.
- شبکه: یک شبکه نه لایه برای طبقه بندی آموزش دید.

نتایج



هوش مصنوعی در سرطان مغز

Neuro-Oncology



سرطانهای Central Nervous System

◦ نرخ بروز سالانه ۱۵ نفر در ۱۰۰۰۰۰ نفر، غالباً در مردان

◦ علیرغم پیشرفتهای درمان بسیار چالش برانگیز است (نظیر glioblastoma multiform (GBM)

◦ مدیریت سرطان های CNS اولیه (درجا) یا متاستاتیک به دلیل تکثیر سریع، نئوواسکولاریزاسیون شدید، تهاجم به نقاط دوردست و پاسخ ضعیف به شیمی درمانی به دلیل سد خونی مغزی، چالش برانگیز است.

◦ مدیریت بالینی شامل مشاهده اولیه، درجه بندی، دسترسی به عمق نفوذ، تقسیم بندی و محل تومور، ارزیابی هیستوپاتولوژیک و شناسایی نشانگرهای مولکولی است.

◦ در نتیجه، پزشکان باید به صورت دستی تمام داده‌ها را برای اعتبارسنجی جمع‌آوری کنند تا یک برنامه درمانی را تدوین کنند.

در این راستا، هوش مصنوعی در تشخیص و مدیریت بدخیمی های CNS مفید است

Tumor grades classification and grading via convolutional neural networks and genetic algorithms

- طبقه بندی گریدهای مختلف گلیوما بر اساس تصاویر MRI
- روش: شبکه های کانولوشنی (CNN) با قابلیت تغییر بر اساس GA
- استفاده از bagging برای کاهش واریانس خطا
- نتایج:

- مطالعه یک، جداسازی سه گرید گلیوما: $ACC = 91\%$
- مطالعه دو، جداسازی Glioma, Meningioma, and Pituitary tumor: $ACC = 94\%$

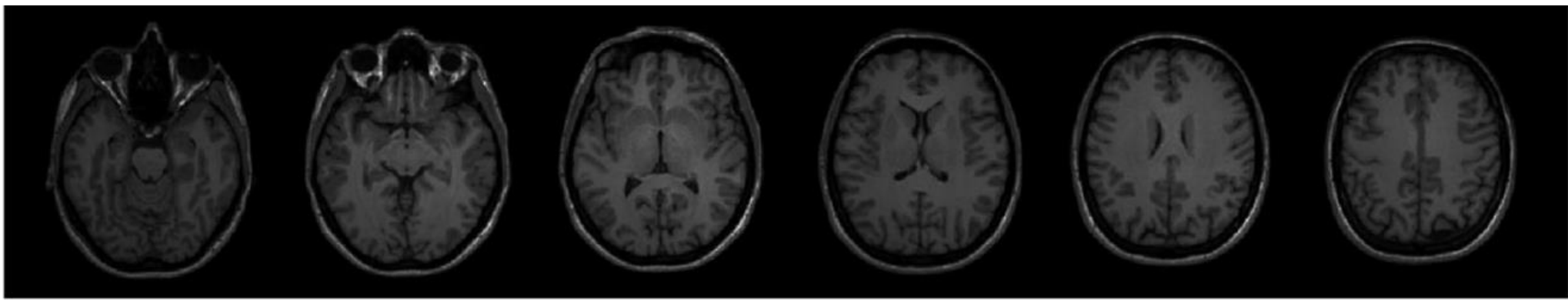


Fig. 1 – An example of the six selected sections from the axial MR of a normal person.

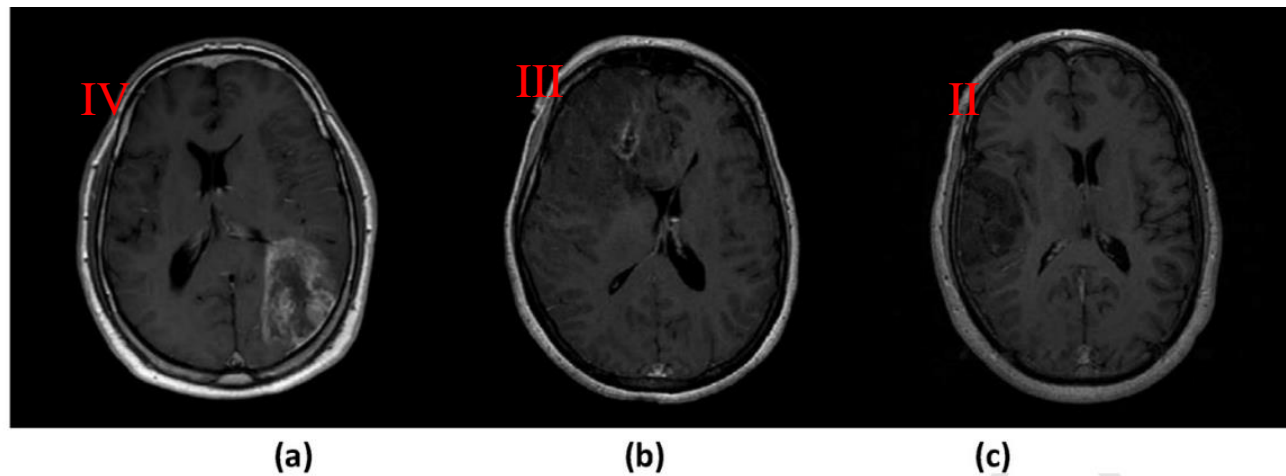


Fig. 2 – Examples of three different grades of Gliomas axial brain images (a) Glioma Grade IV, (b) Glioma Grade III, and (c) Glioma Grade II.

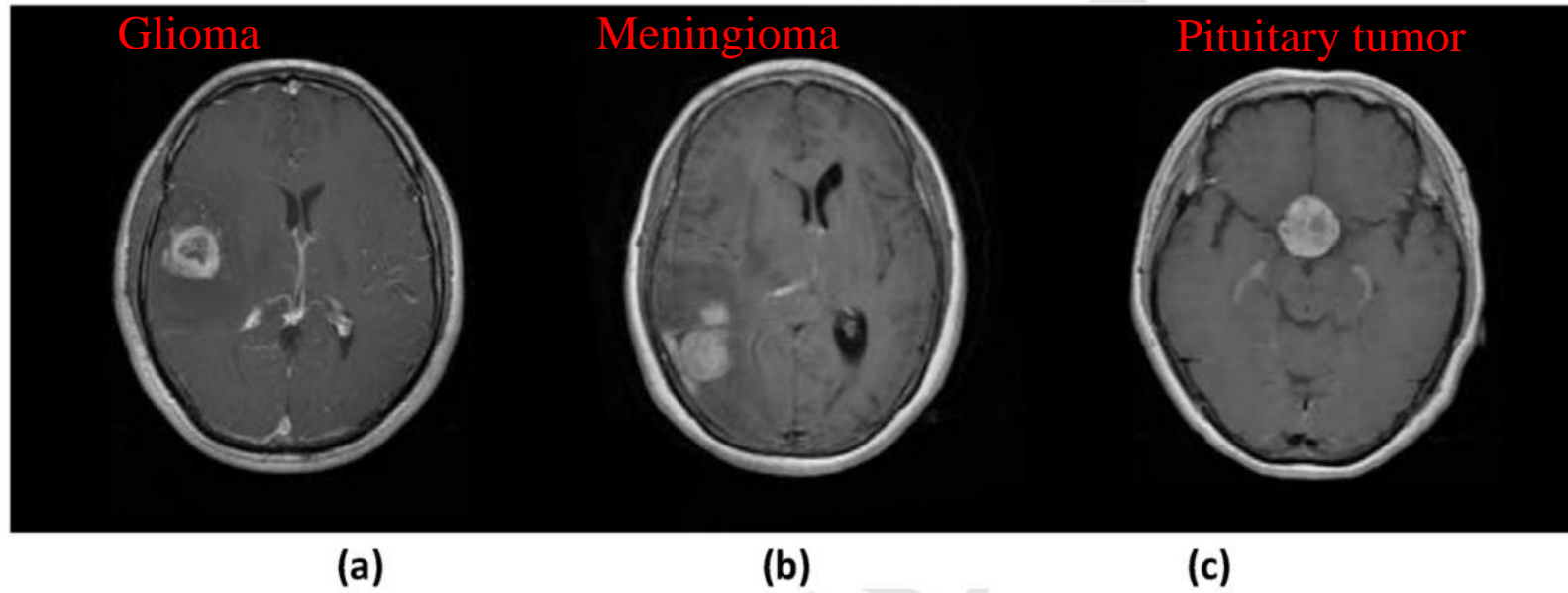


Fig. 3 – Examples of axial brain images from the public dataset provided by Cheng et al. (a) Glioma, (b) Meningioma, and (c) Pituitary tumors.

Table 1 – The parameters employed to evolve the best CNN structure and their associated values.

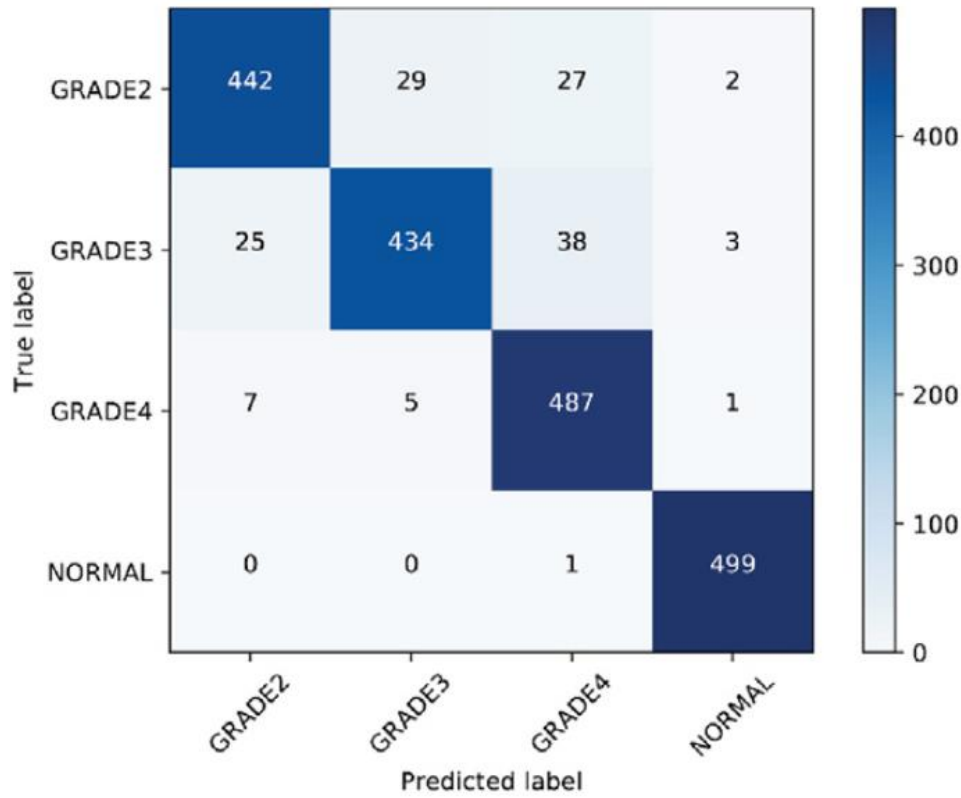
Number of convolutional + max pooling layers	2, 3, 4, 5, 6
Number of fully connected + dropout layers	1, 2, 3
Number of filters	16, 24, 32, 48, 64, 96, 128
Kernel sizes	2, 3, 4, 5, 6, 7
Number of fully connected neurons	128, 192, 256, 384, 512
Activation functions	ReLU, Leaky ReLU, ELU, SELU
Feedforward optimizers	SGD, ADAM, ADAMAX, NADAM,
Learning rate	1e-4, 1e-3, 1e-2
Dropout rate	0.1, 0.2, 0.3, 0.4, 0.5

• بهینه سازی ساختار شبکه CNN بر اساس جدول پارامترهای فوق و فلوجارت الگوریتم ژنتیک

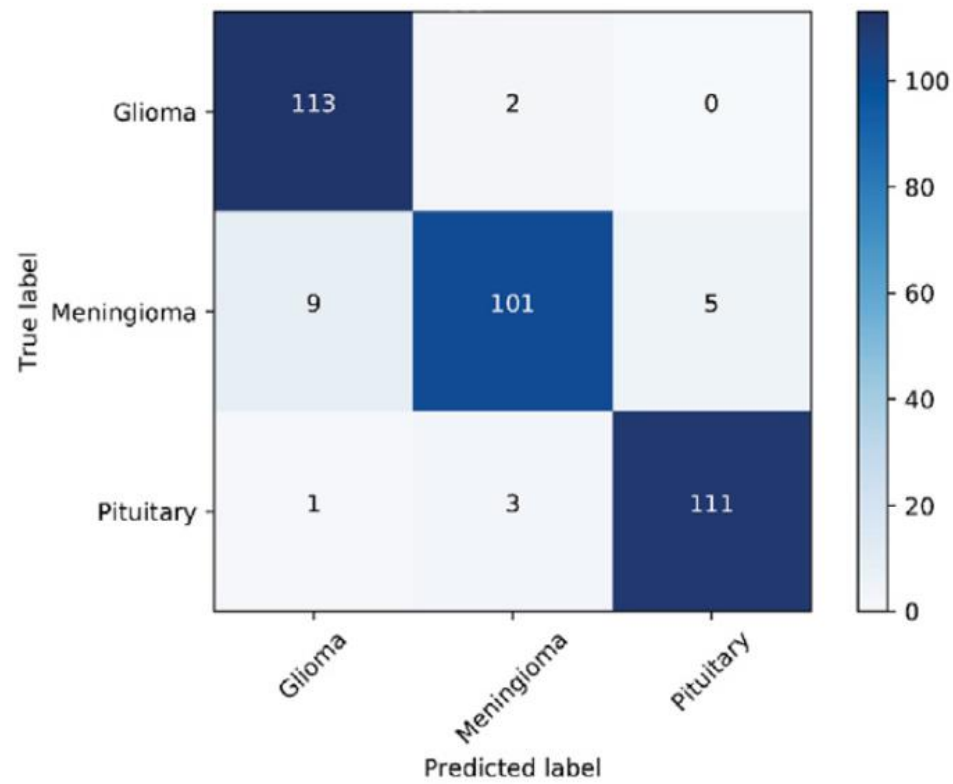


Fig. 6 – Genetic Algorithm Flowchart.

نتایج



(a)



(b)

Fig. 10 – Confusion matrices for (a) Case Study I and (b) Case Study II.

Table 2 – Evaluation of the proposed method.

	TP	FP	TN	FN				
Case Study I								
Normal	499	6	1494	1				
Grade II	442	32	1468	58				
Grade III	434	34	1466	66				
Grade IV	487	66	1434	13				
Case Study II								
Glioma	113	10	220	2				
Meningioma	101	5	225	14				
Pituitary	111	5	225	4				
	TPR	TNR	PPV	NPV	FPR	FNR	FDR	ACC
Case Study I								
Normal	0.998	0.996	0.988	0.999	0.004	0.002	0.012	0.997
Grade II	0.884	0.979	0.932	0.962	0.021	0.116	0.068	0.955
Grade III	0.868	0.977	0.927	0.957	0.023	0.132	0.073	0.950
Grade IV	0.974	0.956	0.881	0.991	0.044	0.026	0.119	0.961
Case Study II								
Glioma	0.983	0.957	0.919	0.991	0.043	0.017	0.081	0.965
Meningioma	0.878	0.978	0.953	0.941	0.022	0.122	0.047	0.945
Pituitary	0.965	0.978	0.957	0.983	0.022	0.035	0.043	0.974

Differentiation Between Glioblastoma, Brain Metastasis and Subtypes Using Radiomics Analysis

- تشخیص افتراقی Glioblastoma از Brain Metastasis بعلت رویکردهای درمانی مختلف بسیار مهم است
- تشخیص این دو از روی تصاویر MRI برای متخصصین دشوار است
- هدف: بر اساس یادگیری ماشین و آنالیز radiomics تشخیص افتراقی انجام شود
- سوژه ها: ۲۱۲ گلیوبلاستوما، ۲۲۷ متاستاز مغزی
- تصاویر MRI یک و نیم و سه تسلا
- ارزیابی: آنالیز شامل پیش پردازش تصویر، تقسیم‌بندی ناحیه تومور، و استخراج ویژگی‌ها از جمله: اطلاعات بالینی بیماران (سن، جنسیت، وزن)، محل تومور، آماره مرتبه اول و دوم، ویژگی‌های مورفولوژیکی، ویژگی‌های موجک. سپس کاهش ابعاد، طبقه‌بندی با استفاده از الگوریتم‌های مختلف یادگیری ماشینی از جمله ماشین بردار پشتیبان SVM، KNN، درخت‌های تصمیم‌گیری و ensemble classifiers
- بهترین نتایج برای SVM
- جدا کردن گلیوبلاستوما از متاستاز مغز: $Acc = 0.85$, $sens = 0.86$, $specif = 0.85$, $AUC = 0.96$
- جدا کردن گلیوبلاستوما از زیرتایپ‌های متاستاز مغز: $Acc = 0.85, 0.89, 0.75, 0.90$

	Number	Age (years)	Gender
Total	439	60 ± 13	230M/209F
Glioblastoma	212	60 ± 12	136M/76F
Brain metastasis	227	60 ± 13	94M/133F
Breast origin	61	54 ± 15	0M/61F
Lung origin	121		
Other origin	45	62 ± 11	9M/19F

M = male; F = female.

760 features

Artzi et al.: Radiomics Classification of Brain Tumors

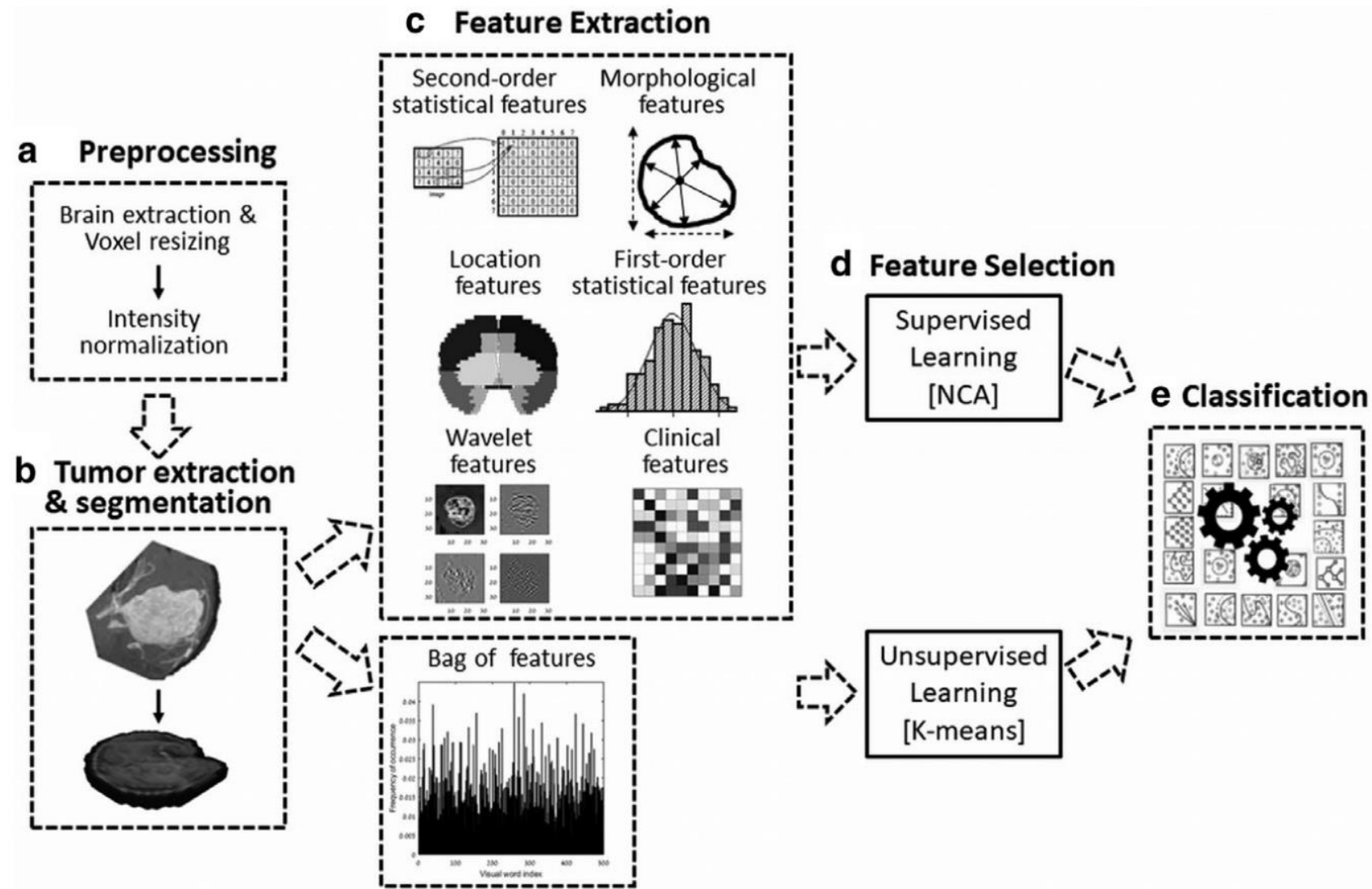


FIGURE 1: Analysis of workflow.

جمع بندی

محدودیت ها

○ بیش برآزش (overfitting): وقتی خطای آموزش کوچک است، ولی در حالت تعمیم یافته (generalization) خطا بزرگ میشود. مساله تعمیم پذیری همواره مهم است زیرا مدل باید در مواجهه با دیتای جدید هم خطای قابل قبولی ارائه دهد.

○ مساله جعبه سیاه: انسان نمی تواند فرآیند تجزیه و تحلیل نتایج به دست آمده توسط یادگیری ماشین را درک کند. مساله جعبه سیاه باعث میشود طراح یا کاربر نتواند رفتار سیستم را پیشبینی کند.

○ اختلاف بین امکانات، به ویژه در تصویربرداری پزشکی (تغییر دامنه و تطبیق دامنه): مساله تغییر دامنه وقتی رخ میدهد که یک الگوریتم بر اساس یک نوع داده ثبت شده توسط یک نوع دستگاه از سازنده‌ای مشخص توسعه داده شده و با استفاده از آن داده ای ثبت شده از دستگاه دیگر طبقه بندی میشود. تغییرات در توزیع آماری دادگان بخاطر اختلاف در پروتکل‌های ثبت؛ نوع دستگاه، تغییر جامعه بیماران رخ میدهد و باعث کاهش عملکرد الگوریتم‌های هوش مصنوعی میشود.

جمع بندی

- توسعه بیش از ۶۰ نوع تجهیزات پزشکی مبتنی بر هوش مصنوعی با تاییدیه **FDA**
- با توجه به توسعه این حوزه، هوش مصنوعی ابزار اصلی فناورانه در مطالعات آنکولوژی خواهد بود
- فناوری های هوش مصنوعی برای تجزیه و تحلیل **omics** در آینده به تدریج افزایش می یابد
- هوش مصنوعی مزایای بسیاری دارد ولی همچنان به مشکلات بسیاری نیز مبتلاست
- بهترین رویکرد تعامل انسان و هوش مصنوعی برای توسعه درمان های بهتر برای سرطان است

Table 1. List of AI-equipped medical devices approved by the US FDA.

No.	FDA Approval Number	Product Name (Company)	Description	Regulation Medical Specialty	Decision Date	Regulatory Class (Submission Type)
1	K140933	AliveCor Heart Monitor (AliveCor)	An ECG recording device using machine learning techniques to detect abnormal heart rhythms.	Cardiovascular	08/15/2014	Class II (510(k))
2	K143468	QbCheck (Qbtech)	A non-invasive test using AI for diagnosis and treatment of ADHD in children.	Neurology	03/22/2016	Unclassified (510(k))
3	K160016	Steth IO (StratoScientific)	An acoustic device using AI to aid in the identification of abnormal heart and lung sounds.	Cardiovascular	07/15/2016	Class II (510(k))
4	K163253	Arterys Cardio DL (Arterys)	A software using deep learning to visualize and quantify cardiovascular MR images.	Radiology	01/05/2017	Class II (510(k))
5	K161328	CANTAB Mobile (Cambridge Cognition)	An iPad-based memory-assessment tool for older adults.	Neurology	01/13/2017	Class II (510(k))
6	K162627	EnsoSleep (EnsoData)	An AI sleep scoring and analysis solution that automates event detection during sleep.	Neurology	03/31/2017	Class II (510(k))
7 *	K162574	AmCAD-US (AmCad BioMed Corporation)	A software to visualize and quantify ultrasound image data with backscattered signals.	Radiology	05/30/2017	Class II (510(k))
8 *	DEN170022	QuantX (Quantitative Insights)	An AI-equipped diagnosis system to aid in accurate diagnosis of breast cancer.	Radiology	07/19/2017	Class II (De Novo)
9	K172311	BioFlux Device (Biotricity)	A remote patient monitoring platform with AI.	Cardiovascular	12/15/2017	Class II (510(k))
10	K171056	WAVE Clinical Platform (Excel Medical Electronics)	A patient surveillance and predictive algorithm platform using AI.	Cardiovascular	01/04/2018	Class II (510(k))

Table 1. Cont.

No.	FDA Approval Number	Product Name (Company)	Description	Regulation Medical Specialty	Decision Date	Regulatory Class (Submission Type)
11 *	K173542	Arterys Oncology DL (Arterys)	An AI-based, cloud-based medical imaging software that automatically measures and tracks lesions and nodules in MRI and CT scans.	Radiology	01/25/2018	Class II (510(k))
12	DEN170073	ContaCT (Viz.AI)	An AI algorithm to analyze CT scans and detect signs of stroke.	Radiology	02/13/2018	Class II (De Novo)
13	K170540	DM-Density (Densitas)	A machine learning application that provides on demand automated breast density assessments at point-of-care.	Radiology	02/23/2018	Class II (510(k))
14	P160007	Guardian Connect System (Medtronic MiniMed)	A continuous glucose monitor with AI assistance.	Clinical Chemistry	03/08/2018	PMA
15	DEN180001	IDx-DR (IDx)	A software program that uses an AI algorithm to analyze retinal images.	Ophthalmic	04/11/2018	Class II (De Novo)
16	K173931	MindMotion GO (MindMaze)	A gamified neurorehabilitation therapy platform using AI.	Physical Medicine	05/17/2018	Class II (510(k))
17	K180455	NeuralBot (Neural Analytics)	A lucid robotic ultrasound system for brain blood flow assessment.	Radiology	05/22/2018	Class II (510(k))
18	DEN180005	OsteoDetect (Imagen Technologies)	A computer-aided detection and diagnostic software that uses an AI algorithm to analyze two-dimensional X-ray images for signs of distal radius fracture.	Radiology	05/24/2018	Class II (De Novo)
19	K173821	LungQ (Thirona Corporation)	A lung quantification software to analyze chest CT scans.	Radiology	06/05/2018	Class II (510(k))
20	DEN170043	DreaMed Advisor Pro (DreaMed Diabetes)	An AI-powered technology to seamlessly treat patients remotely with its virtual diabetes management service.	Clinical Chemistry	06/12/2018	Class II (De Novo)

Table 1. Cont.

No.	FDA Approval Number	Product Name (Company)	Description	Regulation Medical Specialty	Decision Date	Regulatory Class (Submission Type)
21	K172983	HealthCCS (Zebra Medical Vision)	An AI-powered software that can be used to evaluate calcified plaques in the coronary arteries.	Radiology	06/13/2018	Class II (510(k))
22	K173780	EchoMD Automated Ejection Fraction Software (Bay Labs)	A system that enables fully automated AI echocardiogram analysis.	Radiology	06/14/2018	Class II (510(k))
23	K180647	BriefCase (Aidoc Medical)	An AI algorithm to detect and triage abnormal findings in non-enhanced head CT images.	Radiology	08/01/2018	Class II (510(k))
24	DEN180042	Irregular Rhythm Notification Feature (Apple)	An application to detect irregular heart rhythms in pulse rate data collected by the Apple Watch photoplethysmograph sensors.	Cardiovascular	09/11/2018	Class II (De Novo)
25	DEN180044	ECG App (Apple)	Applications to detect atrial fibrillations and sinus rhythms in ECG data from an Apple Watch and display the results on an iPhone.	Cardiovascular	09/11/2018	Class II (De Novo)
26	K173872	FibriCheck (Qompium)	A smartphone application for the detection of heart rhythm disorders such as atrial fibrillation.	Cardiovascular	09/28/2018	Class II (510(k))
27	K181771	RightEye Vision System (RightEye)	A cloud-based system that uses objective eye movement measurements to aid in the evaluation of Parkinson's disease.	Neurology	09/28/2018	Class II (510(k))
28 *	K182034	Arterys MICA (Arterys)	An AI-based platform for analyzing medical images such as MRI and CT.	Radiology	10/17/2018	Class II (510(k))

Dataset	Cancer	URL
AIDA-E	Stomach	https://isbi-aida.grand-challenge.org/
BraTS	Brain	https://www.med.upenn.edu/sbia/brats2018/data.html
BreakHis	Breast	https://web.inf.ufpr.br/vri/databases/breast-cancer-histopathological-database-breakhis/
BACH	Breast	https://iciar2018-challenge.grand-challenge.org/Dataset/
CAMELYON	Lymph node	https://camelyon17.grand-challenge.org/Data/
CCSD	Cervix	https://www.kaggle.com/c/intel-mobileodt-cervical-cancer-screening/data
Histopathologic Cancer Detection	Lymph node	https://www.kaggle.com/c/histopathologic-cancer-detection/data
ISIC 2018	Skin	https://challenge.kitware.com/#challenge/5aab46f156357d5e82b00fe5
INbreast	Breast	http://medicalresearch.inescporto.pt/breastresearch/index.php/Get_INbreast_Database

INbreast	Breast	http://medicalresearch.inescporto.pt/breastresearch/index.php/Get_INbreast_Database
KiTS	Kidney	https://kits19.grand-challenge.org/data/
LIDC-IDRI	Lung	https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI#940027f1a8a845d0a61a1b5b5083567e
MED-NODE	Skin	http://www.cs.rug.nl/imaging/databases/melanoma_naevi/
mini-MIAS	Breast	http://peipa.essex.ac.uk/info/mias.html
mitos-atypia	Breast	https://mitos-atypia-14.grand-challenge.org/dataset/
PROMISE12	Prostate	https://promise12.grand-challenge.org/Home/
PAIP	Liver	https://paip2019.grand-challenge.org/Dataset/
TCIA	Miscellaneous	https://www.cancerimagingarchive.net/
18F-FDG	Head and Neck	https://www.kaggle.com/c/pet-radiomics-challenges/data

با تشکر از توجه شما

