هوش مصنوعی در پژوهش

Artificial Intelligence

از ایده تا استقرار بالینی

هوش مصنوعی در خدمت سلامت

ارائه‌دهنده: دکتر یاسین فدائی

دانشگاه: دانشگاه علوم پزشکی شهرکرد

تاریخ:1404/8/29

مقدمه: هوش مصنوعی در پزشکی

- نقش هوش مصنوعی در تشخیص، درمان و تصمیم‌گیری‌های بالینی روز به روز پررنگ‌تر می‌شود.
- **مزایا:** افزایش دقت در تشخیص، سرعت بالا در پردازش داده‌ها، کشف الگوهای پنهان، و کمک به تصمیم‌گیری پزشکان.
- **ریسک‌ها:** احتمال سوگیری در داده‌ها، چالش‌های تفسیرپذیری، مسائل حریم خصوصی بیماران و امنیت اطلاعات.
- هدف نهایی: نه جایگزینی پزشک، بلکه تقویت توانایی‌های پزشکان برای مراقبت بهتر از بیماران.

انواع مدل‌های هوش مصنوعی در پزشکی

مدل‌های پیش‌بینی‌کننده:

- مدل پیش‌بینی خطر سکته قلبی
- مدل تشخیص زودهنگام سرطان پستان
- مدل پیش‌بینی پاسخ به دارو
- مدل پیش‌بینی دیابت نوع ۲
- مدل پیش‌بینی نیاز به بستری در ICU
- مدل پیش‌بینی عوارض جراحی

مدل‌های پردازش تصویر:

- تشخیص ضایعات در تصاویر MRI مغز
- شناسایی شکستگی در تصاویر X-Ray
- تحلیل اسلایدهای پاتولوژی دیجیتال
- بخش‌بندی ریه در تصاویر CT
- تشخیص رتینوپاتی دیابتی از فوندوس چشم
- شناسایی اختلالات قلبی از اکوکاردیوگرافی

مدل‌های زبانی:

- استخراج اطلاعات از گزارش‌های پزشکی
- تفسیر نتایج آزمایش‌ها به زبان ساده
- خلاصه‌سازی خودکار پرونده پزشکی

نقشه راه: مدل‌های پیش‌بینی‌کننده

۱ **بیان مسئله**
شناسایی مشکل بالینی واقعی

۲ **استخراج متغیرها**
شناسایی عوامل مؤثر

۳ **اعتبارسنجی متغیرها**
ارتباط با خروجی و کیفیت

۴ **تعیین حجم نمونه**
با فرمول EPV

۵ **بررسی وضعیت داده**
گذشته‌نگر یا آینده‌نگر

۶ **آماده‌سازی داده**
پاکسازی و استانداردسازی

۷ **انتخاب مدل**
بر اساس نوع خروجی و حجم

۸ **ارزیابی مدل**
سنجش آماری و بالینی

۹ **اجرای پایلوت**
استقرار محدود و کنترل‌شده

نکته: اجرای صحیح همه مراحل، کلید موفقیت مدل‌های پیش‌بینی‌کننده در محیط بالینی است.

گام 1: بیان مساله

در پژوهش‌های علوم پزشکی، باید مسأله‌ای انتخاب شود که هوش مصنوعی واقعاً بتواند به شکل معنی‌دار و منطقی در حل آن کمک کند.

وقتی می‌خواهید یک موضوع انتخاب کنید، این ۵ سؤال کلیدی را از خود بپرسید:

- آیا داده‌های این مساله زیاد، پیچیده و متنوع هستند؟
- آیا انسان‌ها در این کار خطا، ناهماهنگی یا محدودیت دارند؟
- آیا زمان زیاد یا هزینه زیاد برای حل مساله صرف می‌شود؟
- آیا الگوهای پنهانی وجود دارد که انسان نمی‌تواند ببیند؟
- آیا محصول نهایی تحقیق می‌تواند در بالین قابل استفاده باشد؟

گام 2: استخراج متغیرهای اولیه

- **نظر متخصصین بالینی:** مشورت با پزشکان متخصص در حوزه مورد مطالعه برای شناسایی متغیرهای کلینیکی مهم و مؤثر در تشخیص و پیش‌بینی.
- **مرور مقالات و گایدلاین‌ها:** بررسی سیستماتیک مطالعات مشابه قبلی و دستورالعمل‌های بالینی معتبر برای استخراج متغیرهای استاندارد.
- **تعریف عملیاتی هر متغیر:** مشخص کردن نحوه دقیق اندازه‌گیری و ثبت هر متغیر برای اطمینان از یکنواختی داده‌ها.
- ایجاد فهرست اولیه از متغیرهای مستقل (پیش‌بینی‌کننده‌ها) و متغیر وابسته (خروجی) با جزئیات کافی.

کیفیت مدل نهایی به شدت وابسته به انتخاب صحیح و جامع متغیرهای اولیه است.

گام 3: اعتبارسنجی متغیرها

معیارهای اعتبارسنجی متغیرها:

- آیا همه متغیرها وجود دارند؟
- آیا واقعاً مرتبط با خروجی است یا خیر؟
- کیفیت کافی برای استفاده در مدل دارند یا نه؟
- قابل اندازه‌گیری و قابل اتکا هستند؟
- آیا سوگیری (Bias) وارد مدل نمی‌کنند؟

در صورت نیاز متغیرهای بومی اضافه گردد.

گام 4: تعیین حجم نمونه (EPV)

قانون طلایی حجم نمونه: EPV

EPV (Events Per Variable): برای هر متغیری که وارد مدل می‌کنیم، حداقل ۱۰ تا ۲۰ رویداد واقعی لازم است.

حجم نمونه = (EPV × تعداد متغیرها) ÷ شیوع

این فرمول، تضمین می‌کند که مدل به اندازه کافی از نمونه‌های مثبت برای یادگیری الگوهای مرتبط استفاده کرده و از پدیده Overfitting جلوگیری می‌شود.

انتخاب EPV مناسب

- مدل‌های خطی ساده: EPV = 10
- مدل‌های پیچیده (Random Forest، XGBoost): EPV = 15-20
- داده‌های نامتوازن یا با شیوع پایین: EPV = 20-25

مثال عملی محاسبه

۱. **تعداد متغیرهای مدل:** ۱۵ متغیر

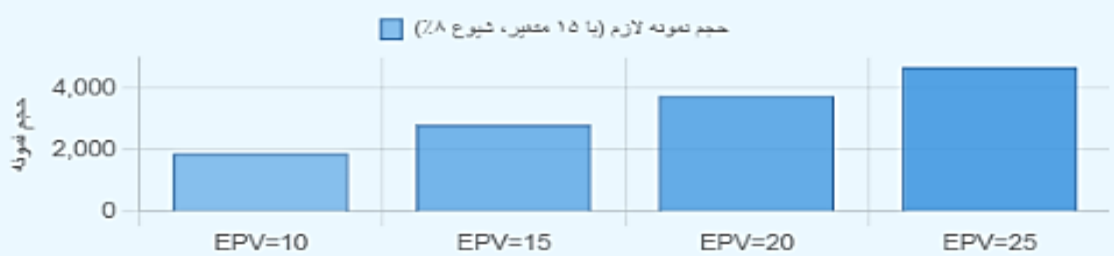
۲. **شیوع بیماری:** ۸٪ (۰.۰۸)

۳. **EPV مناسب:** ۲۰

۴. **محاسبه:** **حجم نمونه = (20 × 15) ÷ 0.08 = 3750**

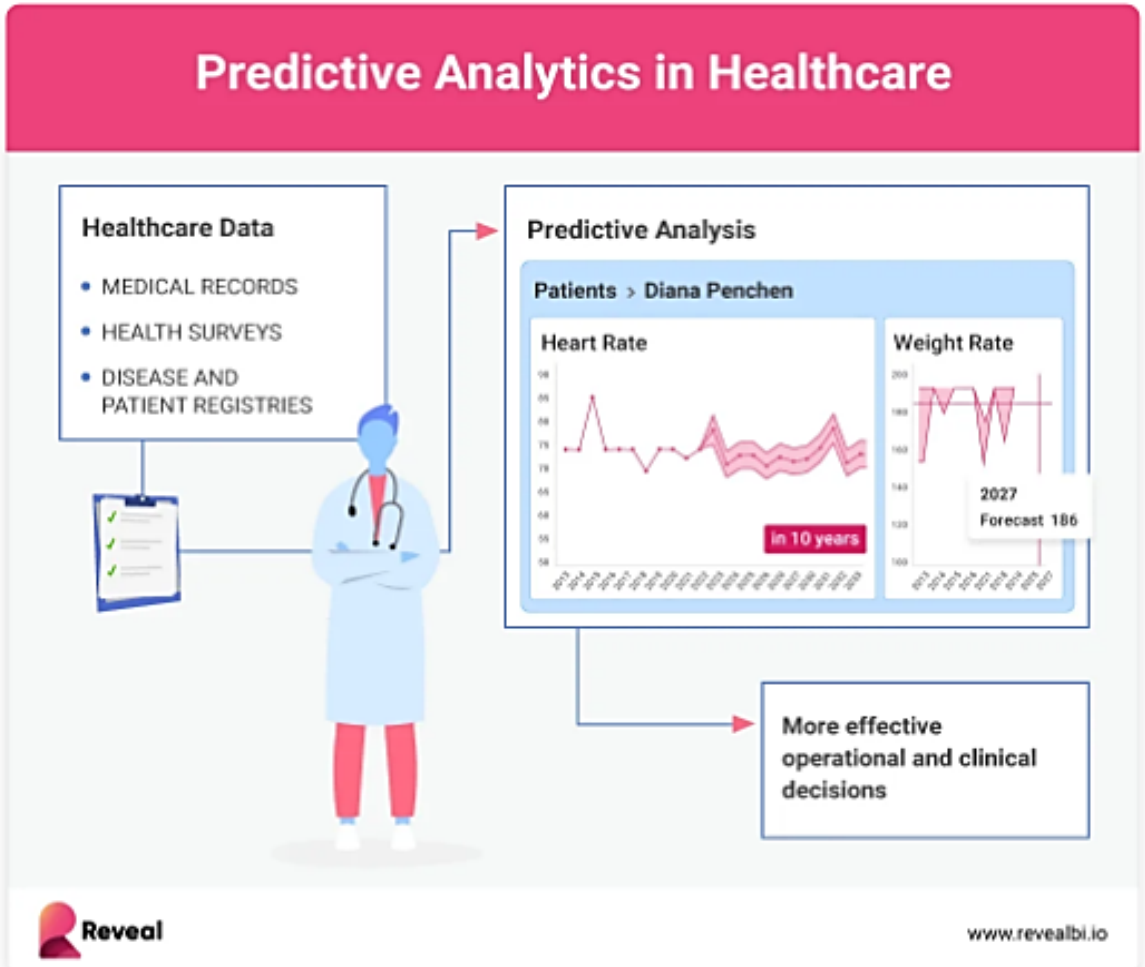
۵. **با احتساب داده‌های ناقص (۲۰-۳۰٪):**

حجم نمونه نهایی ≈ ۴۰۰۰-۵۰۰۰



هشدار: حجم نمونه ناکافی منجر به مدل‌های غیرقابل اعتماد و نتایج اشتباه می‌شود.

گام 5: بررسی وضعیت داده‌ها



- آیا داده‌های کافی با حجم مناسب و برچسب‌گذاری شده وجود دارد؟
- **مطالعه گذشته‌نگر (Retrospective):** سریع و ارزان، اما با چالش‌های بسیار مانند داده‌های ناقص، سوگیری‌های متعدد، و خطاهای ثبت.
- **مطالعه آینده‌نگر (Prospective):** زمان‌بر و پرهزینه، اما با کیفیت بالا، دقت بیشتر و کنترل کامل بر فرآیند جمع‌آوری.
- انتخاب استراتژی جمع‌آوری داده باید با توجه به زمان، بودجه، و الزامات کیفیت پروژه انجام شود.

مطالعات نشان می‌دهد که 60 تا 80 درصد زمان پروژه‌های هوش مصنوعی صرف آماده‌سازی داده می‌شود.

گام 6: آماده‌سازی داده‌ها

مراحل پیش‌پردازش داده

Missing
مدیریت داده‌های گمشده با جایگذاری میانگین/میانه/مد یا پیش‌بینی

Scaling & Normalization
جلوگیری از تسلط یک متغیر بر دیگر متغیرها با مقیاس‌بندی

Balancing
متعادل‌سازی کلاس‌ها با SMOTE، Under/Over-Sampling

Categorical (Encoding)
تبدیل متغیرهای غیرعددی به عددی با روش‌های One-Hot یا Label

Outlier
شناسایی و مدیریت داده‌های پرت با روش‌های آماری

Data Augmentation
افزایش حجم داده با ایجاد نمونه‌های مصنوعی برای آموزش بهتر

نکته کلیدی:
داده خام همیشه پر از خطاست؛ مدل خوب از داده تمیز ساخته می‌شود.

گام 7: انتخاب مدل

نوع خروجی پیش‌بینی:

- دوتایی-چندکلاسه: Logistic Regression، Random Forest، XGBoost، SVM، LightGBM، Multinomial Logistic Regression

حجم داده:

- کمتر 1000: Logistic Regression، SVM، Random Forest
- متوسط بین 1000 تا 5000: Random Forest، XGBoost
- بیشتر از 5000: XGBoost، LightGBM
- خیلی زیاد: مدل‌های مبتنی بر شبکه‌های عصبی

تفسیرپذیری:

- Logistic Regression، Cox Model، Decision Tree (بیشترین قابلیت تفسیر برای کاربرد بالینی)

گام 8: ارزیابی مدل

ارزیابی آماری

حساسیت (Sensitivity / Recall)
از بین بیماران واقعی، چند نفر را درست تشخیص داده؟

TP / (TP + FN)

ویژگی (Specificity)
از بین افراد سالم، چند نفر را درست سالم تشخیص داده؟

TN / (TN + FP)

دقت کلی (Accuracy)
چند درصد پیش‌بینی‌ها درست بوده‌اند؟

TP + TN + FP +) / (TP + TN)
(FN

دقت پیش‌بینی مثبت (PPV)
از بین کسانی که مدل بیمار گفته، چند نفر واقعاً بیمار هستند؟

TP / (TP + FP)

AUC-ROC

نموداری از Sensitivity در مقابل Specificity که توانایی تشخیص بیمار/سالم را نشان می‌دهد.
هرچه به 1 نزدیک‌تر باشد، مدل قوی‌تر است.

هشدار: مدل با دقت بالا اما تفسیرپذیری پایین، در محیط بالینی مورد اعتماد قرار نمی‌گیرد.

گام 8: ارزیابی مدل

ارزیابی بالینی

تفسیرپذیری (Interpretability)

در پزشکی هیچ مدلی بدون تفسیرپذیری وارد بالین نمی‌شود. پزشک باید بداند تصمیم مدل بر چه اساسی بوده است.

- **SHAP Values**: تعیین اهمیت هر متغیر در تصمیم‌گیری مدل
- **LIME**: توضیح محلی برای پیش‌بینی‌های مشخص
- **Rule Extraction**: استخراج قوانین تصمیم‌گیری از مدل‌های پیچیده

مقایسه با متخصصان انسانی

- آیا مدل حداقل به اندازه متخصصان عملکرد دارد؟
- آیا می‌تواند به پزشکان در تشخیص بهتر کمک کند؟
- آیا زمان تصمیم‌گیری را کاهش می‌دهد؟

گام 9: اجرای پایلوت بالینی

آزمون در محیط واقعی

- استقرار محدود و کنترل‌شده در بخش کوچکی از بیمارستان
- سنجش عملکرد واقعی مدل و شناسایی مشکلات اجرایی
- ارزیابی اثر مدل روی کیفیت مراقبت از بیمار
- بررسی میزان پذیرش مدل توسط پزشکان و کادر درمان
- ارزیابی تأثیر بر جریان کار بالینی (Workflow) و فرایندهای موجود

نکته کلیدی:

"اگر مدل در کار روزمره پزشک جا نگیرد، حتی اگر بهترین دقت را داشته باشد، استفاده نخواهد شد."

بخش دوم

مدل‌های پردازش تصویر

از دریافت DICOM تا گزارش قابل اتکا: فرآیند طراحی و پیاده‌سازی مدل‌های پردازش تصویر برای تشخیص و تحلیل تصاویر پزشکی

تشخیص ضایعات در MRI

بخش‌بندی ریه در CT

شناسایی شکستگی در X-Ray

تشخیص رتینوپاتی دیابتی

گام‌های 1-3: تعریف مسئله، جمع‌آوری داده، برچسب‌گذاری

۱

?

تعریف دقیق مسئله

مشخص کردن دقیق هدف مدل: کشف ضایعه، طبقه‌بندی یا بخش‌بندی تصویر. تعیین خروجی‌های مورد انتظار و معیارهای موفقیت.

۲

جمع‌آوری داده

استخراج از سیستم PACS بیمارستان، آرشیو پاتولوژی یا تصویربرداری جدید. تعیین فرمت تصاویر (DICOM، JPEG) و استانداردها.

۳

برچسب‌گذاری

استفاده از حداقل سه متخصص برای برچسب‌گذاری و توافق بین آنها. تهیه دستورالعمل برچسب‌گذاری استاندارد برای اطمینان از یکسان‌سازی.

نکته: کیفیت داده و برچسب‌گذاری دقیق، تأثیر بسزایی در عملکرد نهایی مدل‌های پردازش تصویر دارد.

گام 4: تعیین حجم نمونه تصویری

اصول کلی:

- در مدل‌های پردازش تصویر، حجم نمونه بر اساس فرمول‌های کلاسیک تعیین نمی‌شود
- حجم نمونه تابع پیچیدگی مدل، تنوع داده، و هدف مطالعه است

مدل‌های از پیش آموزش‌دیده:

- حداقل ۱۰۰۰-۱۵۰۰ تصویر برای هر کلاس لیبل (ResNet، EfficientNet، ViT)
- مناسب برای مواردی که داده محدود است اما نیاز به دقت بالا دارید

آموزش از صفر:

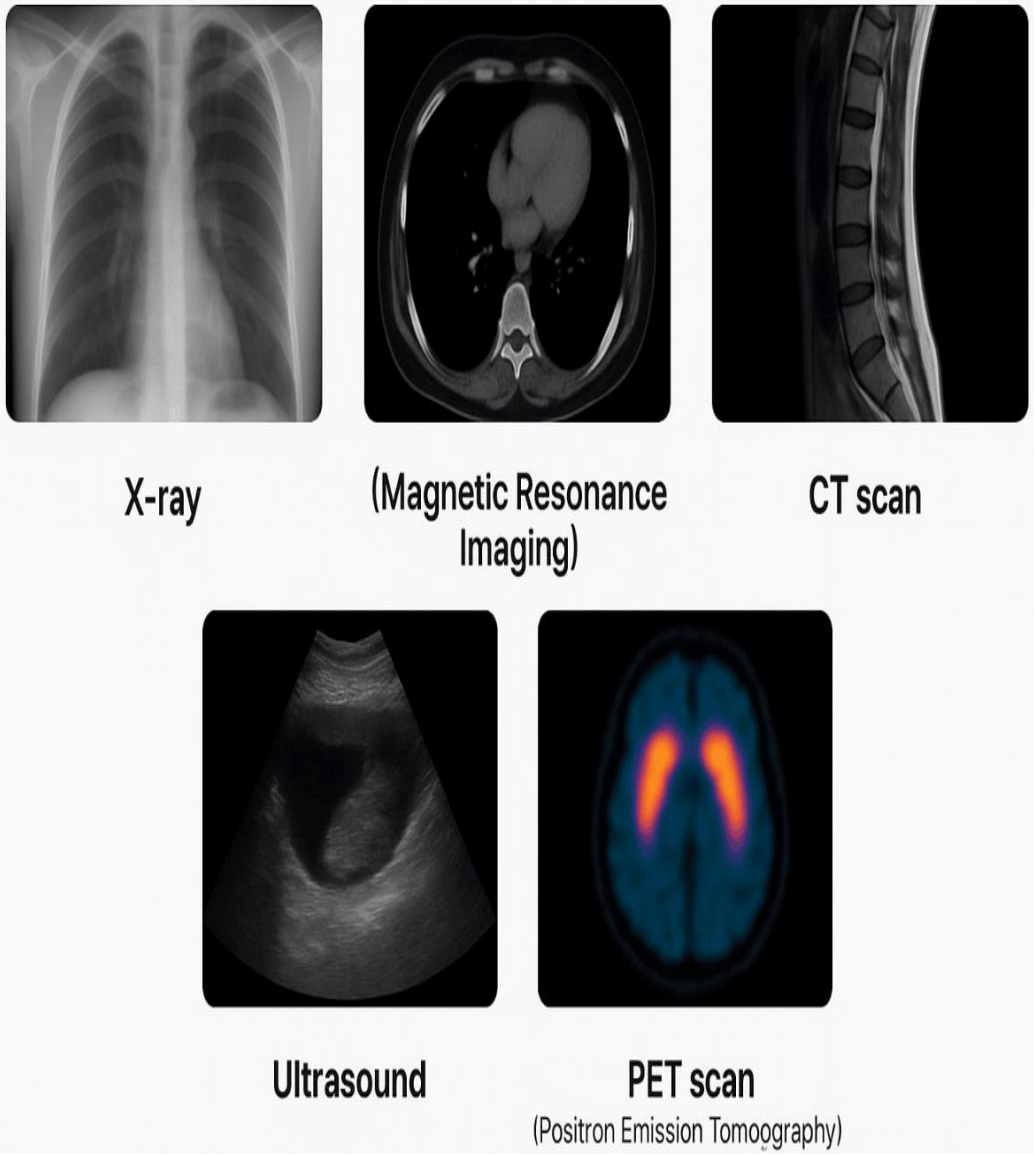
- حداقل ۵۰۰۰ تصویر برای هر کلاس لیبل
- نیازمند منابع محاسباتی قوی‌تر و زمان آموزش طولانی‌تر

موارد خاص:

- برای داده‌های ویدئویی: ۲ تا ۳ برابر حجم داده بیشتر نیاز است
- استفاده از Data Augmentation برای افزایش مصنوعی داده‌ها (Rotation، Flip، Zoom)
- پرهیز از Data Duplicates که باعث سوگیری در آموزش می‌شود

نکته: کیفیت داده‌ها همیشه مهم‌تر از کمیت صرف است. داده‌های متنوع با کیفیت بالا نتایج بهتری نسبت به داده‌های تکراری و کم‌کیفیت ارائه می‌دهند.

گام 5: پیش‌پردازش تصاویر



- **استانداردسازی ابعاد تصویر:** مدل‌های هوش مصنوعی اندازه ثابت می‌پذیرند (مثلاً 224×224 یا 512×512 پیکسل).
- **نرمال‌سازی شدت پیکسل‌ها:** تبدیل مقادیر به بازه [0,1] برای آموزش سریع‌تر و پایدارتر مدل.
- **حذف نویز و بهبود کیفیت:** فیلترهای مختلف برای تمرکز مدل روی ساختارهای واقعی بافت، نه نویز تصویر.
- **برش ناحیه‌های غیرضروری:** تمرکز بر منطقه مورد نظر (Region of Interest) برای بهبود دقت تشخیص.
- **تبدیل فرمت‌ها:** تبدیل DICOM به JPG/PNG برای کاهش حجم داده و افزایش سرعت.
- **حذف اطلاعات حساس (De-identification):** حفظ حریم خصوصی بیمار و رعایت اصول محرمانگی.

بدون پیش‌پردازش مناسب، حتی بهترین مدل‌های هوش مصنوعی نیز عملکرد ضعیفی خواهند داشت.

گام 6: انتخاب و طراحی مدل

مدل‌های طبقه‌بندی (Classification):

- ResNet: شبکه عمیق با اتصالات باقیمانده، کاهش مشکل گرادیان ناپدید شونده
- EfficientNet: توازن بین عمق، عرض و رزولوشن، کارایی بالا با پارامتر کمتر
- ConvNeXt: ترکیب مزایای CNN با ویژگی‌های Transformer
- DenseNet: اتصالات متراکم بین لایه‌ها، استفاده بهینه از اطلاعات
- Vision Transformer (ViT): بر پایه مکانیزم توجه، بدون استفاده از کانولوشن

مدل‌های تشخیص ضایعه (Detection):

- YOLOv8/9: سریعترین مدل‌های تشخیص، مناسب برای کاربرد بلادرنگ
- RetinaNet: توازن بین سرعت و دقت، با تمرکز بر مشکل عدم توازن کلاس‌ها
- Faster R-CNN: دقت بالا با معماری دو مرحله‌ای، استاندارد در تصاویر پزشکی

مدل‌های تقسیم‌بندی (Segmentation):

- U-Net: معماری انکودر-دیکودر با اتصالات پرشی، استاندارد طلایی در پزشکی
- UNet3+: ترکیب اطلاعات چندمقیاسی با قابلیت یادگیری عمیق‌تر
- Swin-UNet: ترکیب ترنسفورمر پنجره‌ای با معماری U-Net
- U-Net++: نسخه بهبودیافته U-Net با اتصالات پرشی متراکم‌تر
- DeepLabv3+: ترکیب کانولوشن با نرخ اتساع و انکودر-دیکودر

گام 7: ارزیابی مدل تصویری

- **آماری:**

دسته بندی Classification

- Sensitivity / Recall
- Specificity
- Accuracy
- AUC-ROC

تشخیص ضایعه Detection

$$\frac{|A \cap B|}{|A \cup B|} = IoU$$

- A: باکس مدل
- B: باکس واقعی (Ground Truth)
- IoU ≥ 0.5 → قابل قبول
- IoU ≥ 0.75 → خوب
- IoU ≥ 0.9 → بسیار دقیق (برای پزشکی حساس)

تقسیم‌بندی بافت Segmentation

$$\frac{|A \cap B|2}{|A| + |B|} = Dice$$

- A: سگمنت مدل
- B: سگمنت طلایی (پزشک/کارشناس)

- **بالینی:** مقایسه عملکرد مدل با عملکرد متخصصان انسانی.

گام 8: پایلوت در محیط تصویربرداری

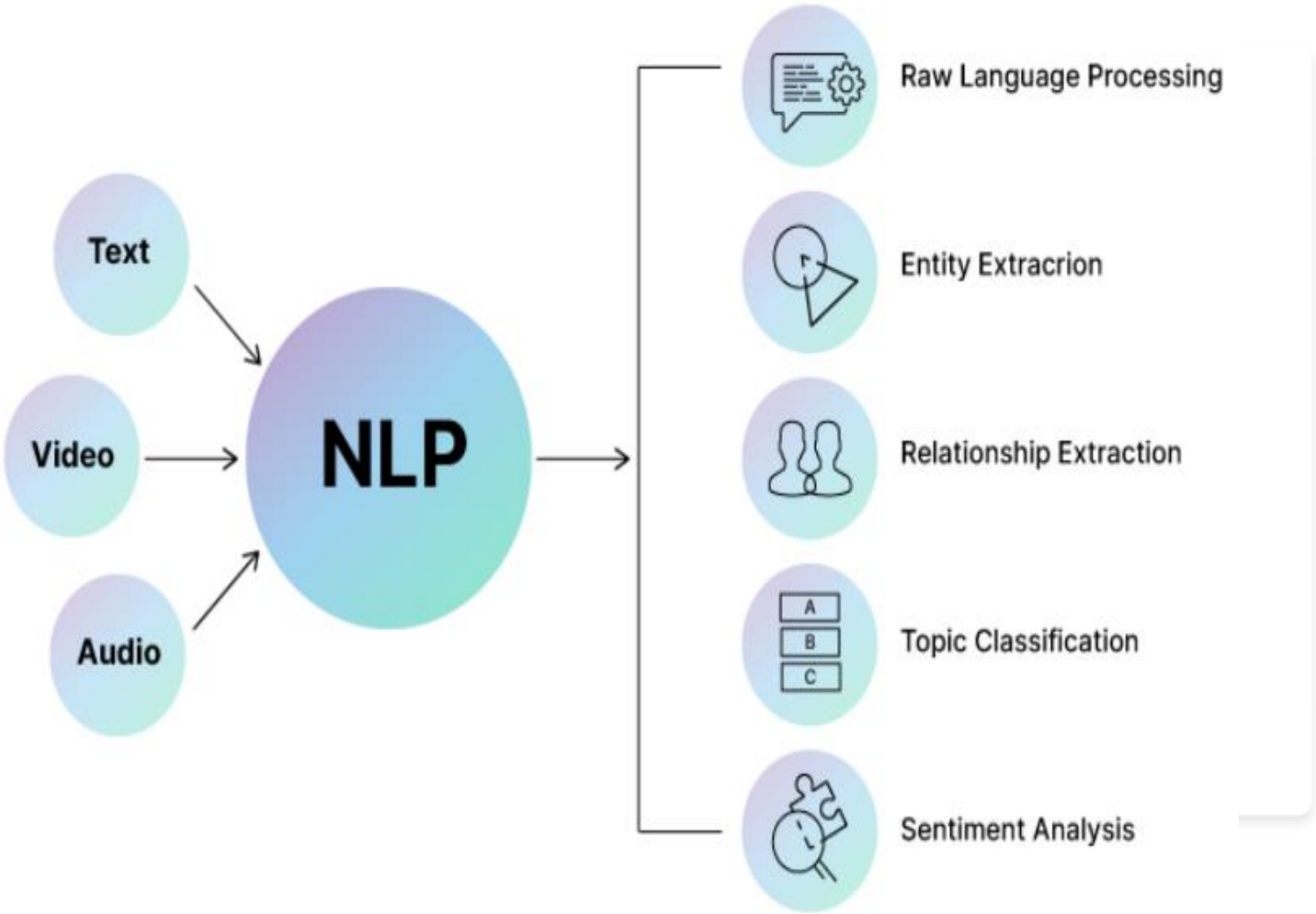
- ادغام با PACS/RIS (سیستم‌های آرشیو تصاویر و اطلاعات رادیولوژی)
- سنجش سرعت استنتاج و زمان پردازش در محیط واقعی
- ارزیابی سیستم هشدار خطا و مدیریت موارد حاشیه‌ای
- بررسی تجربه کاربری رادیولوژیست‌ها و کارکنان تصویربرداری
- سنجش تأثیر بر Workflow و فرآیند کاری بخش تصویربرداری
- بررسی تأثیر بر کیفیت و سرعت گزارش‌نویسی

نکته مهم:
حتی بهترین مدل‌های پردازش تصویر، بدون ادغام مناسب با سیستم‌های موجود و پذیرش توسط کاربران بالینی، به مرحله استقرار واقعی نخواهند رسید.

مدل‌های زبانی در پزشکی

از EMR تا تصمیم‌یار بالینی: استفاده از پردازش زبان طبیعی و مدل‌های زبانی برای استخراج، خلاصه‌سازی و تفسیر متون پزشکی

گام 1: جمع‌آوری و آماده‌سازی داده متنی



- **منابع داده:** پرونده‌های الکترونیک (EMR)، گزارش‌های بالینی، ثبت مکالمات پزشک-بیمار، مقالات علمی، گایدلاین‌ها
- **پاک‌سازی:** حذف خطاهای تایپی، اصلاح ناسازگاری‌های املایی، استانداردسازی فرمت‌ها
- **حذف اطلاعات شناسایی (De-identification):** حذف نام، شماره بیمه، آدرس و سایر اطلاعات حساس بیماران
- **نرمال‌سازی اصطلاحات:** تبدیل به استانداردهای پزشکی مانند SNOMED CT (تشخیص‌ها)، ICD-10 (کدهای بیماری‌ها)، RadLex (اصطلاحات رادیولوژی)
- **ساختاردهی:** تبدیل متون نامنظم به ساختارهای قابل تحلیل برای الگوریتم‌های پردازش زبان طبیعی

کیفیت مدل‌های زبانی پزشکی به طور مستقیم وابسته به کیفیت و پاکیزگی داده‌های متنی اولیه است.

گام 2: استراتژی ساخت مدل زبانی

Fine-tuning روی LLM موجود

آموزش مجدد مدل زبانی بزرگ موجود با داده‌های پزشکی اختصاصی.
رویکرد کم‌هزینه و رایج با نتایج قابل قبول برای اکثر کاربردها.

Instruction Tuning با سناریوهای بالینی

آموزش مدل با دستورالعمل‌ها و نمونه‌های واقعی پزشکی برای بهبود پاسخ‌دهی در موقعیت‌های بالینی مختلف.

RAG: بازیابی + تولید

اتصال مدل زبانی به پایگاه دانش پزشکی، گایدلاین‌ها و منابع به‌روز برای ارائه پاسخ‌های مستند و قابل استناد.

Training از صفر

ساخت مدل زبانی اختصاصی از پایه. نیازمند داده و منابع محاسباتی عظیم، مناسب تنها برای سازمان‌های بسیار بزرگ.

نکته: برای اکثر پروژه‌های بالینی، استفاده از راهکارهای قبلی توصیه می‌شود.

مدل های بینایی-زبانی VLM

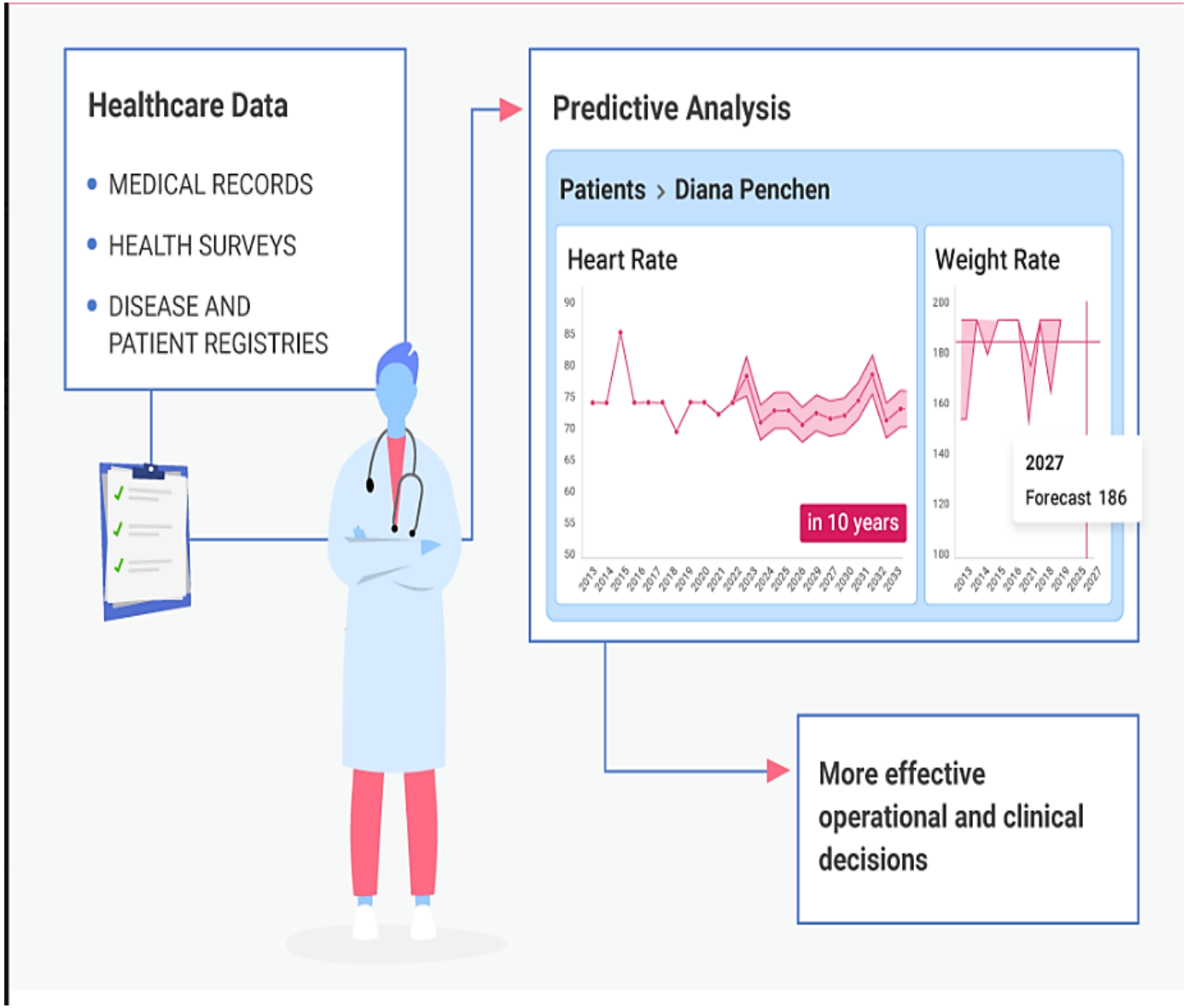
مدل‌های بینایی-زبانی (Vision-Language Models) توانایی درک و تحلیل هم‌زمان تصاویر پزشکی و متون بالینی را دارند.

چرا برای علوم پزشکی مهم هستند؟

- توضیح‌پذیری: مدل می‌تواند توضیح دهد چرا ضایعه‌ای را تشخیص داده است.
- چندمنبعی بودن: می‌تواند همزمان تصویر + علائم + سن + آزمایش را تحلیل کند.
- خطاپایین‌تر از مدل‌های تک‌تصویری در موارد مبهم.
- قابلیت ساخت سیستم‌های کمک‌تشخیص واقعی.

جمع‌بندی و توصیه‌های عملی

- از مساله بالینی واقعی شروع کنید، نه از تکنیک‌های هوش مصنوعی. مدل باید نیاز واقعی را برطرف کند.
- کیفیت داده و تفسیرپذیری، کلید ورود به بالین است. بدون این دو، حتی دقیق‌ترین مدل‌ها اعتماد پزشکان را جلب نمی‌کنند.
- اجرای پایلوت کنترل‌شده پیش از استقرار گسترده الزامی است تا اثرات جانبی ناخواسته شناسایی شوند.
- تیم چندتخصصی متشکل از متخصصین بالینی، داده و هوش مصنوعی برای موفقیت ضروری است.



> تکنولوژی به تنهایی کافی نیست، بلکه باید با دانش بالینی و اخلاق پزشکی همراه باشد.

معرفی ابزار جستجو و یافتن شکاف تحقیقاتی

- Chatgpt.com
- Grok.com
- Gemini.com
- Scopus
- Pubmed
- Scispace
- Connected Papers
- Liner
- Notebooklm
- Researchrabit
- Consensus
- Answer This

سپاسگزارم

از توجه شما به برنامه کاربردی هوش مصنوعی در پزشکی سپاسگزارم

دکتر یاسین فدائی - دانشگاه علوم پزشکی شهرکرد

Email: Fadaei.yasin@gmail.com